Ecology, 93(8) Supplement, 2012, pp. S14–S30

Assessing the effects of spatial contingency and environmental filtering on metacommunity phylogenetics

Pedro R. Peres-Neto,[1,4] Mathew A. Leibold,[2] and Stéphane Dray[3]

[1]*Canada Research Chair in Spatial Modeling and Biodiversity, Department of Biological Sciences, Université du Québec à Montréal, Case Postale 8888, Succursale Centre-Ville Montréal, Québec H3C 3P8 Canada*
[2]*Section of Integrative Biology, 1 University Station C0930, University of Texas, Austin, Texas 76712 USA*
[3]*Laboratoire de Biométrie et Biologie Evolutive, Université de Lyon, Université Lyon 1, CNRS, UMR 5558, 43 Boulevard du 11 Novembre 1918, 69622 Villeurbanne, France*

Abstract. Patterns in biodiversity and species coexistence are the result of multiple interacting processes including evolutionary history, trait variation, species interactions, dispersal, environmental variation, and landscape heterogeneity. Exploring patterns of biodiversity across space is perhaps the best integrative method (in contrast to the scarcity of temporal data) to interpret the influence of these multiple and interactive effects in determining community assembly, but it is still underdeveloped. Two emerging fields, metacommunity ecology and community phylogenetics, have been making relevant, though rather independent, progress toward understanding how communities are assembled in space. Our main goals were twofold. First, we described a heuristical framework to merge these two fields into "metacommunity phylogenetics." The main goal of this framework is to provide a way to think about how niche properties of species arranged across the environment and different spatial scales influence the process of community assembly. Second, we developed an analytical framework to link niche properties based on trait and phylogenetics to environmental and spatial variation. In order to assess the performance of the framework, we used extensive computer simulations of community assembly to show that the procedure is robust under a variety of scenarios.

Key words: community phylogenetics; comparative biology; gradient analysis; metacommunity; multiscale analyses; multivariate analyses; spatial analysis; stochastic simulation; trait-based community assembly.

Introduction

Local communities consist of species that are subsets of those that exist in the larger regional species pools of the biotas in which they occur. What factors determine which subsets are selected to coexist in which sites? Community assembly theory (Diamond 1975, Keddy 1992, Brown et al. 2000, Emerson and Gillespie 2008) seeks to identify rules that reflect different possible processes that determine how local communities are selected from large species pools. Although the processes are many, and they interact with each other, they have two key elements: regional processes that regulate the arrival of organisms into the various local communities (e.g., environmental heterogeneity, landscape connectivity, dispersal limitation) and local processes that regulate the success of species following either their own arrival or the arrival of other species (e.g., niche differentiation, local environment, microhabitat heterogeneity).

Two emerging fields, metacommunity ecology (Leibold et al. 2004) and community phylogenetics (Webb et al. 2002), have been making relevant progress toward understanding the importance of local and regional processes, though rather independently (but see Leibold et al. 2010). Metacommunity ecology is interested in how dispersal interacts with local community assembly (Leibold et al. 2004) and has largely focused on environmental heterogeneity and spatial regulation of species distributions (e.g., Cottenie 2005, Beisner et al. 2006, Sharma et al. 2011). Because these studies are based on patterns of species distributions (and not community structure per se; but see Leibold and Mikkelson 2002, Presley et al. 2010), they tell us little about how community assembly actually results from the interactions between local and regional assembly processes. Moreover, community studies based on species identity can be largely context dependent in the sense that species compositions change across regional pools (landscapes) and through time.

Alternatively, community phylogenetics is focused on how species traits and phylogenetic relatedness are associated with mechanisms of community assembly with the goal of understanding how deeply the footprint of evolution extends into current ecological processes and associated patterns. One major goal of community phylogenetics (but see Cavender-Bares et al. 2009 for

Manuscript received 16 March 2011; revised 7 November 2011; accepted 16 November 2011. Corresponding Editor: J. Cavender-Bares.

[4] E-mail: peres-neto.pedro@uqam.ca

other approaches and goals) is to understand how species that share similar traits or evolutionary histories should coexist (Webb et al. 2002, Strauss et al. 2006, Cavender-Bares et al. 2009), and determine the conditions that are likely to make similar species coexist more or less often than expected by chance, a pattern referred as to phylogenetic clustering (often evoked as the result of environmental filtering) and evenness (often evoked as the result of density-dependent interactions), respectively.

Taken together, current methods of metacommunity analysis do not sufficiently examine patterns in relation to variation among species (i.e., community structure rather than species distributions), while community phylogenetics does not sufficiently examine patterns in relation to variation among sites in their contributions to niche-related patterns, and particularly by ignoring spatial and multi-scale environmental effects. We thus seek to merge the two fields into metacommunity phylogenetics. We start by formalizing the links and interplays between the ecological (traits) or evolutionary (phylogenetic history) attributes of species, and local environment and the spatial heterogeneity (e.g., due to the spatial structure of environmental features and/or spatial signatures due to dispersal dynamics and species-interactions) across the sites of species assemblages. This treatment provides a heuristic framework to think about how species niche properties are structured across environmental gradients and across spatial scales, ultimately influencing the process of community assembly. We then introduce an analytical approach that decomposes the total phylogenetic (or trait) variation into a mean (compositional) and a variance (dispersion) component that are subsequently linked to environmental and spatial variation at different scales.

The goal of our proposed framework, shared by many studies, is to explore how environmental and spatial variation (spatial scale) influence community structure by selecting species (or allowing species to coexist) on the basis of their trait and phylogenetic associations. Although different trait/phylogenetic patterns are associated to different predictions about the processes underlying community assembly, we now know that both of these patterns can be found within the same metacommunity, often acting at different spatial scales (Cavender-Bares et al. 2006, Swenson et al. 2007, Willis et al. 2010) and different environments (Graham et al. 2009, Willis et al. 2010, Machac et al. 2011; such that some communities are phylogenetically/trait clustered, whereas others are evenly dispersed or even random). By proving a heuristical and analytical framework that formalizes the links between community trait/phylogenetic patterns and environmental and spatial variation, our hope is that ecologists start looking at these links in a more systematic way. Moreover, because local communities within metacommunities are likely to be composed by a mix of patterns, our approach can help resolve these conflicting patterns if they are related to environmental or spatial variation.

While our approach certainly has some analytical and/or heuristical commonalities with previous ones (Ackerly and Cornwell 2007, Leibold et al. 2010, Ives and Helmus 2011, Pavoine et al. 2011; see *Discussion* for further details), we present some interesting novel features including the coupling of evolutionary stochastic simulations within a modeling framework, how to consider spatial variation in the analysis of phylogenetic/trait associations and the decomposition of phylogenetic/trait means and variances.

A multi-scale framework for metacommunity phylogenetics

Different species (and individuals) have different movement and dispersal capabilities, thus responding differently to the way in which optimum, non-optimum, and unfavorable environments are spatially organized. Moreover, because species have different dispersal capacities and perceptual ranges (in the case of some animals), they also differ in the way in which they respond to the spatial organization of environmental features (Lima and Zollner 1996). The links among these elements (i.e., spatial organization of environmental features and differences in dispersal capacities) with local and regional species interactions result in the complex spatial patterns observed in species distributions and metacommunities. Although the geographic multi-scale nature of environmental controls on species distributions is well recognized in modern ecology (Wiens 1989, Menge and Olson 1990, Levin 1992), there has been little effort to tackle the relevant geographic scales of phylogenetic and trait structure. In the same way that ecologists use direct gradient analysis to explore the processes underlying species distributions on the basis of environmental and spatial predictors (Cottenie 2005, Legendre et al. 2005, Peres-Neto et al. 2006 and references therein), our approach is to model phylogenetic or trait community structure (compositional and variation components; see Fig. 1) in relation to local environment and multi-scale spatial predictors (see Fig. 2 and next sections).

Trait and phylogenetic community structure have been used either as a predictor (e.g., Lavorel and Garnier 2002, Cadotte et al. 2009, Ives and Helmus 2011, Lavorel et al. 2011) or as response variable (as in our framework; e.g., Lavorel and Garnier 2002, Leibold et al. 2010, Willis et al. 2010, Ricotta and Moretti 2011). As in many other areas, whether a variable is considered as a response or as a predictor depends on the specific research goals: Environment may select the species that can inhabit a particular local community on the basis of their traits or phylogenetic relationships ("select the type of species": phylogeny/trait as response), which in turn may affect ecosystem functioning (phylogeny/trait as predictor; Lavorel et al. 2011).

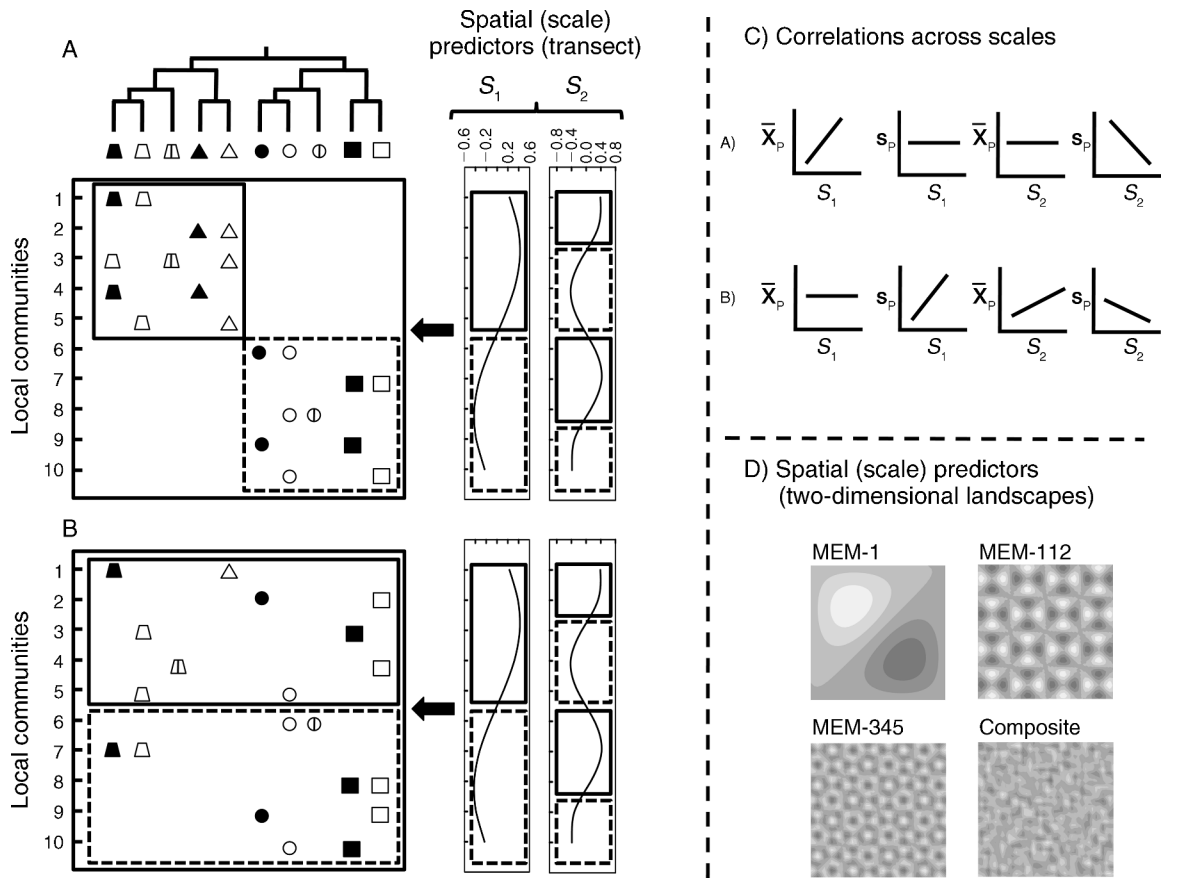

FIG. 1. A representation of the multi-scale metacommunity phylogenetics framework. Two different metacommunities are represented in which local communities are organized according to different levels of phylogenetic structuring across two different spatial scales. Local communities are distributed across a one-dimensional transect to facilitate visualization of these patterns. (A) Local communities differ in their phylogenetic composition across the largest scale S_1, whereas in (B), local communities differ in their levels across phylogenetic similarity. The opposite patterns between the two metacommunities (A and B) are observed at the smaller spatial scale S_2 (see section on *A multi-scale framework for metacommunity phylogenetics* for further details). Although the scales are continuous, to facilitate visualization we identified similar patterns of community organization by solid vs. dashed boxes. For instance, in (A), for S_1, the two boxes represent major changes in the mean (compositional) components, whereas for S_2, it represents changes in the variance component (dashed boxes encompass phylogenetically clustered communities, whereas the solid boxes designate evenly dispersed communities). (C) Relationships between scale variation (S_1 vs. S_2) and the mean ($\bar{\mathbf{X}}_P$) and variance ($\mathbf{s}_P$) components of metacommunity structure shown in rows corresponding to the scenarios in panels (A) and (B), respectively. (D) A two-dimensional representation of spatial scales using the Moran's eigenvector maps (MEMs; see section on *Representing spatial scales and patchiness* for further details; note that scales in panels (A) and (B) are one-dimensional MEMs) based on a 40×40 regular lattice. Three MEMs are presented to illustrate how they depict spatial scaling and spatial heterogeneity. These MEMs are then used in our framework as spatial predictor variables (S) to depict scale-dependent variation in metacommunity structure. The composite map represents a linear combination of the three presented MEMs and serves to demonstrate how they can depict very detailed spatial structure in data.

Although trait and phylogenetic summary statistics regarding composition and variation components have been used as response variables to assess several different patterns and hypotheses regarding community structure, they have not all received the same attention. Several studies have modeled community-weighted trait means (CWM, mean component) as a function of the environment as a way to understand how environmental gradients select trait composition at local communities (trait gradient analyses; e.g., Ackerly and Cornwell 2007). Similar attention has been given to the factors that drive the trait variation component (i.e., summary statistics of functional diversity; e.g., Barnett and Beisner 2007, Pedruski and Arnott 2011). Although the component of phylogenetic variation (i.e., diversity) has been widely used as a predictor (e.g., Cadotte et al. 2009), it has only recently gained attention as a response variable. For example, Willis et al. (2010) found that patterns of phylogenetic dispersion (i.e., variation component) were a function of the variance in light availability within communities (see also Graham et al. 2009 and Machac et al. 2011 for additional examples). Although biogeographical studies are often interested in understanding how clade (phylogenetic) composition

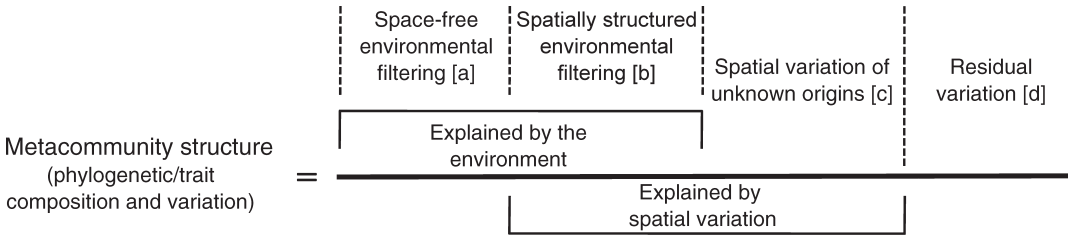

FIG. 2. A representation of our modeling scheme of metacommunity structure considering both compositional and variation components. These two components are modeled against local environmental and spatial predictors. The contributions of each set of predictors are interpreted according to four elements or fractions: [a] the component explained by environment that is not spatially structured or space-free environmental filtering, [b] the component explained by environment that is spatially structured interpreted as spatially structured environmental filtering, [c] the component in metacommunity structured explained by space independently of the environmental variation (because this component can have multiple origins [see section *Environmental filtering vs. spatial contingency*], it is referred to as "of unknown origins"), and finally, fraction [d] refers to the nonspatially structured unexplained (residual) variation in metacommunity structure.

varies across space and environments (e.g., progression rule hypothesis related to sequential colonization and speciation; Parent and Crespi 2006), the use of phylogenetic composition in relatively "smaller scale" community ecology either as a response or a predictor has been scarce compared to phylogenetic variation and trait components (but see Webb et al. 2008 and Leibold et al. 2010). There are at least two important reasons why to consider phylogenetic composition as a response variable. First, geographic differences in regional species pool (often related to phylogenetic composition) can be as important in driving community structure as environmental factors (see McPeek and Brown 2000, Lessard et al. 2011 and references therein) and this can be assessed by regressing the phylogenetic composition in relation to spatial and environmental components. As we show below (see Fig. 1 and the *Discussion*), it is important to assess the degree in which phylogenetic composition is driven by environmental and spatial variation at different scales (Fig. 1) because broader scale phylogenetic source pools (composition) can have further consequences into the way to that local communities are assembled (Swenson et al. 2007, Lessard et al. 2011). Second, in the lack of traits (or all possible ecologically important traits structured traits), phylogeny serves as a proxy for the magnitude of expected trait (niche) differences among species if those traits are phylogenetically structured (Cavender-Bares et al. 2009); therefore, in the same way that CWM is used to evaluate trait composition, phylogenetic composition serves as a proxy to infer about the composition of phylogenetically related trait differences across communities (see also *Representing phylogenetic variation* below).

Patterns of phylogenetic/trait community structure in relation to the environment have been explored in several studies (e.g., Ackerly and Cornwell 2007, Graham et al. 2009, Willis et al. 2010, Machac et al. 2011). For example, Graham et al. (2009) and Machac et al. (2011) found that the levels of phylogenetic clustering and evenness are affected by environmental features, in which some environments may be more prone to environmental filtering, whereas others to density dependent interactions. Comparable patterns have been found regarding phylogenetic structure across space. For instance, Swenson et al. (2007) found that phylogenetic evenness was more important at small spatial scales; and Willis et al. (2010) found that phylogenetic clustering occurs at greater spatial scales, whereas at smaller scales phylogenetic patterns were highly variable, varying from clustered, randomly, to evenly distributed. Because the interpretation (strength and direction) of the links between trait/phylogenetic structure are relatively simple in contrast to spatial variation, our goal in this section is to provide an intuitive way to understand how phylogenetic/trait community structure can be observed and described within a multi-scale spatial approach.

Our multi-scale template for metacommunity trait/phylogenetics is based on the integration of three elements: (1) the hierarchical selective filtering scheme used by Keddy (1992) based on deletion rules imposed by niche-relationships, (2) the decomposition of spatial scales into different levels of patchiness (Borcard et al. 2004, Dray et al. 2006, Griffith and Peres-Neto 2006, Jombart et al. 2009), and (3) the arrangement of species properties (trait or/and phylogenetic variation) across a heterogeneous mosaic (Fig. 1). Spatial scaling as referred to here correlates with the degree of spatial dependence of a variable (e.g., phylogenetic community structure) on a landscape (see Plate 1). The degree of spatial dependence relates to spatial autocorrelation in the sense that large-scale variation represents a greater degree of spatial dependence for a particular variable of interest compared to intermediate and small-scale structures. Here, we use a definition of spatial scale that relates directly to a common definition of patchiness as cyclic variation in data (phylogenetic/trait community structure) in the form of waves of alternating low and high values (Dale and Fortin 2009; see Borcard and Legendre 2002 as well); the smaller the frequency, the patchier the variation in data (see Fig. 1 for graphic

examples), the smaller the autocorrelation and finally the smaller the scale.

Here, the scale-dependency of metacommunity organization may be a function of a variety of factors such as how processes that act at greater spatial scales constrain phylogenetic and trait pools that are available to structure local communities at smaller scales and vice versa. For instance, large-scale environmental features (e.g., environmental filtering via differences in climate and elevational gradients) may affect the distribution of sub-clades across the landscape, which in turn, may affect the levels of species relatedness and the potential for limiting similarity at lower spatial scales (e.g., local environment, microhabitat). Another factor is the difference among species in terms of their dispersal abilities because dispersal limitation may act as an important mechanism in segregating species spatially, thereby reducing density dependent interactions and exclusion (e.g., Seidler and Plotkin 2006). Taken together, the complexity of spatial patterns observed in metacommunities is related to the way in which species respond to the organization of their environments across different spatial scales (e.g., differences in dispersal capacity and/or perceptual range) and how the strength of species interactions change across these scales (different pools of species are present at different spatial scales).

In Fig. 1A and B we present two examples of simple, but different multi-scale scenarios regarding the ways in which species may be organized phylogenetically into local communities across two spatial scales. Each metacommunitity is composed of 10 local communities that were sampled across a transect. We used one-dimensional spatial scales across a transect rather than multiple scales at two-dimensional landscapes because it is easier to perceive these patterns at a single spatial dimension (but see Fig. 3D for examples of different spatial scales on a 40×40 lattice landscape). For both metacommunities, we present two spatial scales, a broad (S_1, large cycles, greater autocorrelation) and an intermediate (S_2, intermediate cycles, intermediate autocorrelation) scale pattern. In the first metacommunity (Fig. 1A), S_1 relates to a major phylogenetic compositional difference across local communities (i.e., positive values of the scale descriptor relate to one clade and negative values to another clade), whereas S_2 is more strongly associated to a variance difference in phylogenetic relatedness, i.e., some local communities are composed of more similar (local communities 1, 2, 6, 7, and 8; related to positive values of the scale descriptor) and others of more dissimilar species (local communities 3, 4, 5, 9 and 10; related to negative values of the scale descriptor; Fig. 1C). In the second metacommunity (Fig. 1B), we reversed these patterns, in which S_1 is related to different levels of phylogenetic similarity within communities and S_2 is related to a weaker phylogenetic compositional pattern and an even weaker clustering signal (Fig. 1C); for instance, species in local communities 1, 2, 6, 7, and 8 (Fig. 1B) always belong to one single sub-clade, whereas species in local communities 3, 4, 5, 9, and 10, local communities have a tendency to vary in their phylogenetic composition.

Given that real landscapes are composed of multiple scales, there may be multiple ways in which phylogenetic composition and variation may be related to each other in a scaled-dependent fashion. Indeed, both of these patterns (clustering and even phylogenetic/trait dispersion), as well as random arrangements of species, can be found within the same metacommunity, often acting at different spatial scales (Cavender-Bares et al. 2006, Swenson et al. 2007, Willis et al. 2010, Helmus and Ives 2012) and different environments (Graham et al. 2009, Savage and Cavender-Bares 2012).

Environmental filtering vs. spatial contingency

Even in simple experimental landscapes, evidence shows that different levels of habitat connectivity and spatial distribution of resources impose different metacommunity patterns (Gonzalez 2000, Cadotte and Fukami 2005). Natural landscapes are extremely complex in the sense that it is unlikely that two different metacommunities inhabit landscapes that are identical in spatial attributes. As a consequence, local phylogenetic and trait structure may be highly contingent on how the environment and other spatially structured factors and processes are structured across spatial scales. A highly diverse pattern may emerge from landscape to landscape in the sense that two exact pools of species may find different solutions for their coexistence based on different levels of landscape heterogeneity, a phenomena that we refer here as spatial contingency. Indeed, many criticisms regarding contingency and complexity have been leveled at community ecology in recent years (Ricklefs 1987, Simberloff 2004, but see Vellend 2010), and community ecology needs to start dealing with these aspects. While the complex interactions between regional and local processes may induce us to perceive that resultant patterns are highly context dependent and contingent, the properties of recurrent scaling across different landscapes and their metacommunities may provide us a common currency by which the components of phylogenetic and trait variation are determined according to their respective scaling patterning. As Feigenbaum mentioned: "The only things that can be universal, in a sense, are scaling things" (in Wiens 1989:385).

Our approach estimates and tests the contribution of local environment (i.e., abiotic predictors) and spatial predictors (see Fig. 2 and next sections) to phylogenetic and trait community structure (compositional and dispersion components; see Fig. 1). It follows that parts of the environmental contribution to metacommunity structure may be spatially organized whereas others may be spatially independent. Here, we classify the environmental variation that is independent of space as "space-free environmental filtering" (fraction [a] in Fig. 2). This

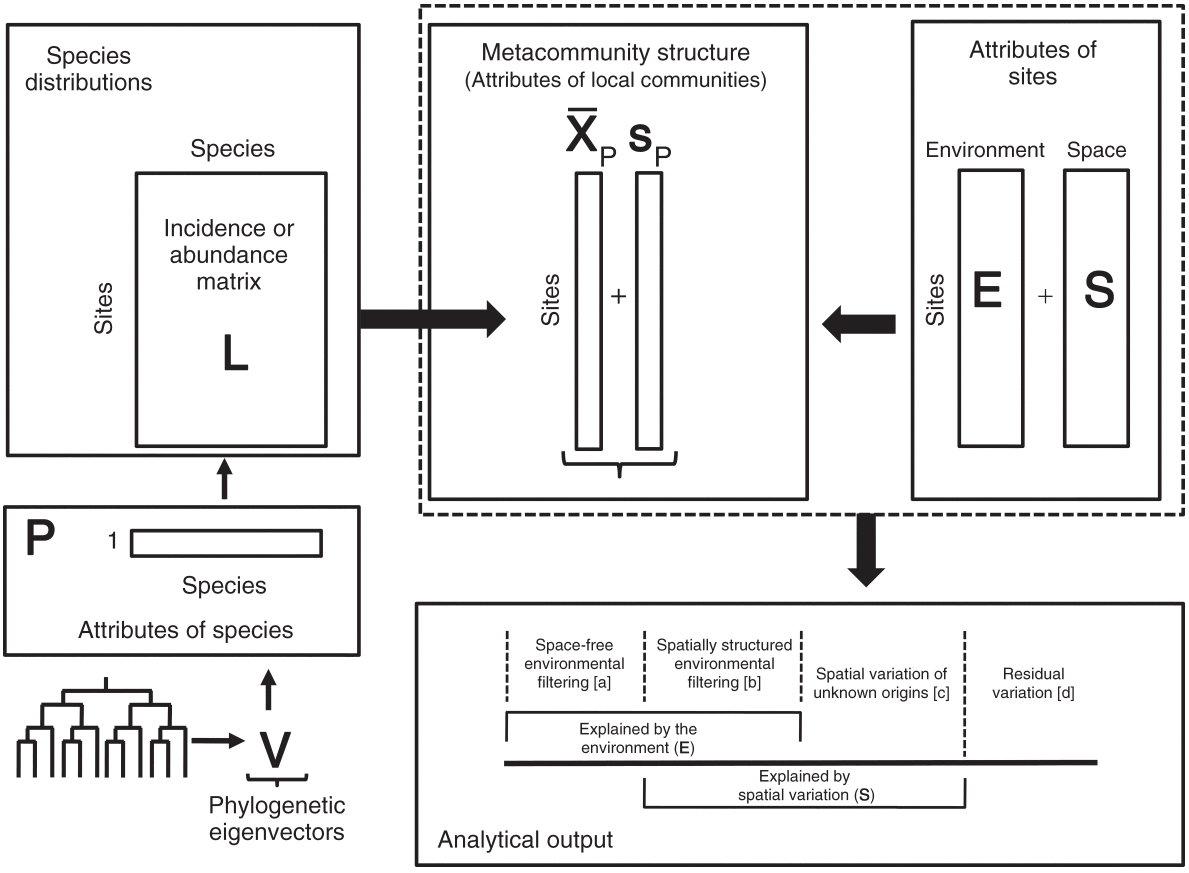

FIG. 3. Links between the types of data (matrices) that are used in our metacommunity phylogenetic framework (see *Methods* for more details). The ultimate goal is to understand how the attributes of sites (space and environment) affect metacommunity structure in terms of phylogenetic variation. The dashed box represents the model components in which attributes of local communities are modeled against attributes of sites. **P** is a vector of species attributes (phylogenetic or trait component) that is linked to a species' distribution matrix **L** to generate a compositional component ($\bar{\mathbf{X}}_P$) and a variation component ($\mathbf{s}_P$). These components are then modeled against environmental (**E**) and spatial (**S**) predictor matrices, and components of variation related to each set of predictors is calculated accordingly.

variation is perhaps the one that can be most directly translated to different landscape settings because they are spatial (scale) free. Next, there are the phylogenetic/trait components of metacommunity structure that are spatially structured. This component ([bc]; Fig. 2) has two fractions. The first is related to the measured environmental variation across local communities, which we classify as "spatially structured environmental filtering" (fraction [b]; Fig. 2); and the second (due to the spatial variation independent of our measures of local environment) is classified as "spatial variation of unknown origins" (fraction [c], i.e., [bc] = [b] + [c]; Fig. 2). Finally, there is the non-environmental (in the sense of the environmental features measured) and nonspatially structured variation in phylogenetic structure that we classify as "residual" (fraction [d] in Fig. 2). Therefore, by estimating the contributions of these fractions, we can gain specific knowledge about how environmental and spatial factors/processes drive metacommunity structure (see Fig. 4).

Although the spatial predictors in our framework are quite powerful in the sense that they can describe multiple scale and patchiness levels (see section *Representing spatial scales and patchiness*), as in any other spatial modeling framework, the origins of the spatial variation in community structure remain unknown because immediate causes of such patterns are inferred rather than tested (Cottenie 2005, Peres-Neto and Legendre 2010; see also *Discussion*). They could be due to several unmeasured factors and processes such as missing environmental predictors, intra- and interspecific interactions, dispersal limitation, problems in experimental design (i.e., distribution of sites), among others. As a consequence, the two spatial components (known and unknown origins) are likely to be contingent on the landscape in the sense that it is unlikely that they will be similar across different landscape settings. However, the spatial component of unknown origins provides a template for pattern-based inference of the relevant processes structuring metacommunities (McIn-

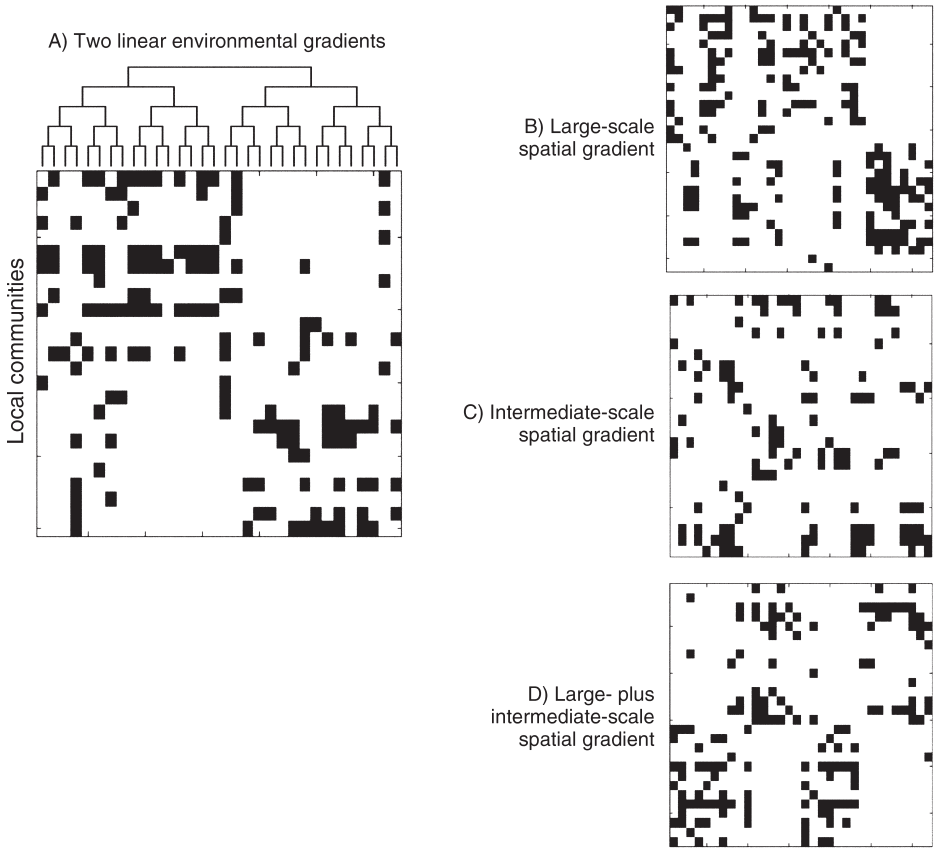

FIG. 4. Examples of simulated species distribution data sets (see *Results: Phylogenetic gradient analyses*) for scenarios (A) I, (B) V, (C) VI, and (D) VII. Black squares represent species presence. Note the amount of noise in the data, though the data shown were among the most structured in each scenario to illustrate the patterns in data according to each scenario.

tire and Fajardo 2009). Moreover, environmental filtering and spatial heterogeneity may interact such that at some levels they may jointly facilitate coexistence of similar species, whereas at other levels, they may hinder such coexistence. For instance, local environment and spatial connectivity (a source of spatial heterogeneity) may interact such that sites with similar local conditions may have similar species if they are highly connected, but differ if they are disconnected because they will be inhabited by species drawn from different subregional species pools (i.e., pools at larger scales within the same metacommunity).

METHODS

Our basic analytical framework for metacommunity phylogenetics is presented in Fig. 3. It links the phylogenetic information among species (vector $\mathbf{P}$, $k \times 1$, where k is the number of species), with the distributional information for these species across local communities (matrix $\mathbf{L}$, sites $\times$ species) containing either presence/absence (i.e., incidence) or abundance information (or also biomass or density). $\mathbf{P}$ and $\mathbf{L}$ are linked by the following operation: $\mathbf{P}_L = \mathbf{1}_n\mathbf{P}_{std}^{\top} \bullet \mathbf{L}$ ($\bullet$ stands for the Hadamard element-wise multiplier, $\mathbf{1}_n$ is a sites $\times$ 1 vector of ones, and std indicates the standardized form of $\mathbf{P}$), resulting in a sites $\times$ species ($n \times k$) matrix containing the total variation in phylogenetic (or trait) structure across local communities. $\mathbf{P}_L$ is then further decomposed into a phylogenetic (or trait) mean $\bar{\mathbf{X}}_P$ (community phylogenetic composition) and a vector variance $\mathbf{s}_P$ (community phylogenetic dispersion) component (see Appendix A for detailed calculations). Finally, these two components (composition and dispersion) are modeled against environmental and spatial (scale) predictors (Figs. 2 and 3).

These components are akin to beta (mean compositional phylogenetic/trait variation across communities) and alpha (phylogenetic/trait variation within communities) phylogenetic diversity (or among- and within-community trait components in Ackerly and Cornwell [2007] framework; see also Ricotta and Moretti [2011]). Here, we mathematically formalize the decomposition of the total variation in trait/phylogeny into these two components (Appendix A), thus allowing a straightforward measure to assess which component is more prevalent for any given metacommunity. Differences in $\bar{\mathbf{X}}_P$ across communities translate into differences in phylogenetic composition (akin to phylogenetic beta diversity), whereas differences in $\mathbf{s}_P$ translate into differences in phylogenetic dispersion (akin to phyloge-

netic alpha diversity). Phylogenetic dispersion (variance) assesses how similar or dissimilar species are according to their phylogenetic position within communities, independent of phylogenetic composition. Two communities can have different phylogenetic compositions but similar levels of phylogenetic clustering among species, and vice versa (Fig. 1), though some levels of correlation are to be expected between these two components in real metacommunities. The variation across local communities regarding $\bar{\mathbf{X}}_{\mathbf{P}}$ and $\mathbf{s}_{\mathbf{P}}$ is then further partitioned into an environmental and a spatial component (Fig. 3, analytical output; see also Peres-Neto and Legendre 2010). Although the approach is quite simple to grasp, it presents several interesting algebraic developments (these are presented in Appendix A), including the appropriate redistribution of the total phylogenetic variation $\mathbf{P}_{\mathrm{L}}$ into a mean (compositional) and variance (dispersion) component across sites (i.e., local communities), and the development of appropriate permutation test procedures. The complete algebra, permutation test, and a small example for which all components are calculated are presented in Appendix A. In the main body of the paper, we focus on the performance of the framework. The procedure can be directly adapted to trait variation (Appendix A), but we focus our discussion on phylogenetic effects, as they are more complicated to resolve (see next section).

Representing phylogenetic variation

There are several statistics that have been suggested to measure the degree of phylogenetic structure based on patterns of phylogenetic clustering and eveness (see Kembel 2009 and references therein) and different ways of representing phylogenetic composition (Pillar and Duarte 2010, Pavoine et al. 2011, Ricotta and Moretti 2011). However, none of these methods allows decomposing the total variation in metacommunity phylogenetic (and trait for that matter) structure ($\mathbf{P}_{\mathrm{L}}$) using a single unified approach that includes the two separate components of interest (i.e., $\bar{\mathbf{X}}_{\mathbf{P}}$ and $\mathbf{s}_{\mathbf{P}}$). Here we present one such solution to the problem by combining an eigenvector decomposition of the phylogenetic distance matrix (Diniz-Filho et al. 1998, Kühn et al. 2009, Pavoine et al. 2011) with stochastic phylogenetic simulations based on the principles of comparative biology (Garland et al. 1993, Revell 2010). The basis of the procedure is described in this section, whereas the analytical details are provided in Appendix B.

The underlying principal behind our method is the use of phylogenetic stochastic simulations based on the fact that community phylogenetics assumes that niche-related traits (e.g., resource use, habitat optima, physiological tolerance, dispersal capacity) are phylogenetically structured (i.e., their major differences are phylogenetically dependent). However, assuming that niche-related traits are phylogenetically structured does not mean that the trajectory that these traits underwent during evolutionary history is best estimated by their phylogenies. Therefore, the estimated phylogeny may be seen as an average expectation of niche differentiation based on these possible evolutionary scenarios (Martins 1995). Here we used stochastic Brownian simulation to estimate different possible phylogenetic scenarios to model how species may have differed in their niche as a test of how these would fit the data (i.e., community structure) better than the expected average. This was done by contrasting $\mathbf{P}_{\mathrm{EV}}$ (one single evolutionary scenario based on the mean expected variation across species in their niche) vs. $\mathbf{P}_{\mathrm{EV\text{-}B}}$ (multiple scenarios based on the stochastic simulations to explore other scenarios of niche evolution; Brown) using our framework (see Appendix B for calculation details). In this case, we assumed that the best explained scenario (mean or via stochastic simulation) is the closest to the true evolutionary variation. The phylogenetic community structure was then modeled against local environmental features and spatial predictors. We used the Brownian model (a well-used model in evolutionary biology) to contrast different outcomes within the same evolutionary model, though other evolutionary models (e.g., Ornstein-Uhlenbeck) would be just as easy to consider in our framework.

Here we suggest an estimation of separate sets of phylogenetic vectors (either $\mathbf{P}_{\mathrm{EV}}$ or $\mathbf{P}_{\mathrm{EV\text{-}B}}$) for space and environment (or for any other source of variation of interest). The estimation of different phylogenetic vectors is in accord with the analysis of multiple traits in the sense that not all traits will respond similarly to different sources of variations and the ecological gradients they represent. This separate estimation is also in agreement with the fact that different phylogenetically structured traits or niche axes may respond differently to different environmental gradients, and is akin to the fact that in comparative biology traits (or niche axes) can have different phylogenetic structures. In this case, we end up with two response variables (one for each source of variation) and our framework is just as straightforward given that a multiple regression with multiple response variables become a redundancy analysis (RDA) and, therefore, the exact same algebra presented in our Appendix A and interpretation regarding variation partitioning applies (see Peres-Neto et al. 2006 for details).

A simulation study to assess the performance of the framework

In this paper, we make use of two sets of simulations. The first, explained in this section, aims at assessing the statistical robustness of the framework in terms of Type I error and statistical power under different permutation test procedures. We also compare the two methods of representing phylogenetic variation ($\mathbf{P}_{\mathrm{EV}}$ and $\mathbf{P}_{\mathrm{EV\text{-}B}}$). We additionally considered the case of trait variation. Here, only a single predictor $\mathbf{E}$ was considered because there are a number of issues to be tested; once we demonstrated the overall performance in this simpler

case, we then used the most robust combination of methods (i.e., permutation procedure and method to represent phylogenetic variation) to evaluate our complete framework based on the variation partitioning of community structure between predictors **E** and **S** in a variety of situations (see section *Putting it all together: phylogenetic gradient analyses*).

The first set of simulations was based on species distribution matrices (100 sites × 50 species) that were constrained, for simplicity, by one environmental predictor **E** and one phylogenetic vector **P**. The phylogenetic component was based on a stochastic pure-birth phylogeny with 50 terminal species using the "ape" phylogenetics package for R (Paradis et al. 2004). In order to generate a species distribution matrix **L** that is constrained by both an environmental variable and a phylogenetically patterned vector, we adapted the protocol described in Dray and Legendre (2008) as follows:

1) Generate a vector **E** (100 sites by 1) containing uniformly distributed random values between 0 and 100. E_i is the environmental value for site i.
2) Generate a vector **P**. Three different ways of generating a phylogenetic vector **P** were considered: (a) a Brownian evolved "trait" vector **t**, (b) three Brownian evolved traits that were each multiplied by a different uniformly distributed random value between 0 and 1 and then summed, and (c) three randomly selected phylogenetic eigenvector that were each multiplied by a different uniformly distributed random value between 0 and 1 and then summed. Each of the **P** vectors was then transformed to vary between −1 and 101. P_j is the phylogenetic value for species j for any of the three methods of generating phylogenetically structured data. In the case of traits, we generated a vector **T** (50 species by 1) containing uniformly distributed random values between −1 and 101. T_j is the trait value for species j. Because the number of species were smaller than the number of sites, it was necessary to apply a slightly greater interval (instead of 0 and 100 for **E** in point 1) to generate the range of explained variation simulated.
3) Generate a vector **h** (50 species × 1) containing uniformly distributed random values between 0.5 and 1. The value h_j is the height of species j at its optimum (i.e., maximum abundance of a species).
4) Generate a vector **σ** (50 species × 1) containing normally distributed random values with standard deviation 10 and mean tolerance (μ_{tol}). The value σ_j is the level of tolerance (i.e., niche breadths) of species j. By varying μ_{tol} across different simulation scenarios, we were able to manipulate the strength of the relationship between phylogenetic variation across species and environmental variation across sites (the greater μ_{tol}, the weaker the relationship).
5) Generate a unimodal response curve for the jth species as follows:

$$\mathbf{L}_{ij} = \mathbf{h}_j \exp\left[\frac{-(\mathbf{E}_i - \mathbf{P}_j)^2}{2\boldsymbol{\sigma}_j^2}\right]$$

where $\mathbf{L}_{ij}$ is the abundance value for the jth species at the ith site. In the case of traits, we simply substituted $\mathbf{P}_j$ with $\mathbf{T}_j$.

In order to assess the Type I error of the procedure, we considered two situations. In the first case, we considered the situation in which **P** was important, but not environment: **E** was replaced by a vector of normally distributed random values N(0, 1) after **L** was generated (results presented in Table 1). In the second case, we considered the reverse situation in which **E** was important, but not phylogenetic variation: **P** was replaced by a vector of normally distributed random values N(0, 1) after **L** was generated (results presented in Table 2). In order to assess the power of the framework and the contrast between methods for representing phylogenetic variation (i.e., $\mathbf{P}_{EV}$ vs. $\mathbf{P}_{EV\text{-}B}$), we generated 1000 matrices **L** for each combination between the three ways of generating phylogenetically structured data and 13 values of μ_{tol} (results presented in Table 3). Both power and Type I error rates were estimated as the proportion of cases out of 1000 in which the null hypothesis was correctly or not correctly rejected, respectively.

In the case of traits, we also considered two situations for assessing the Type I error of the framework. A first case in which **T** was important, but not environment: **E** was replaced by a vector of normally distributed random values N(0, 1) after **L** was generated (results presented in Table 4). A second case in which **E** was important, but not trait variation: **T** was replaced by a vector of normally distributed random values N(0, 1) after **L** was generated (results presented in Table 4). Power was based on the original **E** and **T** used to generate **L**. Power and Type I error were then estimated as the proportion of cases out of 1000 in which the null hypothesis was correctly or not correctly rejected, respectively.

We assessed the performance of three different data permutation test procedures (explained in Appendix A; and in Dray and Legendre 2008, and Pavoine et al. 2011), each based on 999 permutations. These procedures differ in terms of strategies of matrix permutation. The row strategy permutes the rows of **E** (or predictor), the column strategy permutes the columns of **P** (species), and the row and column strategy permutes both columns of **P** and rows of **E** as well.

Representing spatial scales and patchiness

Our framework is based on a matrix of spatial predictors **S** (Figs. 2 and 3) that can be directly embedded into a regression model and variation partitioning schemes (i.e., the rows of **S** represent sites) to depict scale-dependent variation in metacommunity

TABLE 1. Simulation results for the case in which phylogeny, but not environment, was important in species distributions, using different methods, tolerance levels (μ_{tol}), and permutation strategies (row, column, or row and column).

Simulation method and tolerance level	Type I error rate estimate			R^2 (average variation explained across simulations)		
	Row	Column	Row and column	Row	Column	Row and column
One Brownian trait						
$\mu_{tol} = 1$	0.052	0.286	0.028	0.002	0.003	0.001
$\mu_{tol} = 3$	0.036	0.384	0.058	0.001	0.003	0.001
$\mu_{tol} = 5$	0.040	0.588	0.040	0.001	0.002	0.001
$\mu_{tol} = 7$	0.044	0.668	0.038	0.000	0.001	0.000
$\mu_{tol} = 9$	0.042	0.740	0.050	0.000	0.001	0.000
Three Brownian traits						
$\mu_{tol} = 1$	0.058	0.330	0.052	0.002	0.004	0.002
$\mu_{tol} = 3$	0.046	0.418	0.060	0.001	0.003	0.001
$\mu_{tol} = 5$	0.040	0.566	0.044	0.000	0.002	0.001
$\mu_{tol} = 7$	0.034	0.706	0.052	0.000	0.001	0.000
$\mu_{tol} = 9$	0.038	0.714	0.042	0.000	0.000	0.000
Three phylogenetic eigenvectors						
$\mu_{tol} = 1$	0.050	0.750	0.052	0.002	0.010	0.004
$\mu_{tol} = 3$	0.068	0.738	0.057	0.002	0.006	0.002
$\mu_{tol} = 5$	0.044	0.800	0.042	0.001	0.003	0.001
$\mu_{tol} = 7$	0.034	0.806	0.038	0.000	0.001	0.000
$\mu_{tol} = 9$	0.022	0.804	0.032	0.000	0.001	0.000

Notes: Permutation strategy refers to the type of permutation test used (see last paragraph in section *A simulation study to assess the performance of the framework*). Type I error estimates and average R^2 are only shown for the phylogenetic stochastic simulation because it presented greater power in most scenarios (see Table 3). Type I error rate estimates are based on the rate of rejection out of 1000 simulated data sets.

structure. There are different ways of generating spatial predictors (see Legendre et al. 2005 for different methods) and here we suggest the use of Moran's eigenvector maps (MEM; Fig. 1; Dray et al. 2006, Griffith and Peres-Neto 2006). Spatial eigenvector mapping is based on translating the spatial arrangement of data points into predictor variables that capture spatial effects at different spatial scales. The resulting eigenvectors, themselves, are then used directly as synthetic explanatory spatial variables in our variation partitioning scheme. This modeling approach is semi-parametric as it casts the spatial autocorrelation as some unknown function estimated from the data at hands. Therefore, by regressing the components of community phylogenetic structure (mean and variance) against spatial eigenvectors, we can estimate the levels of spatial

TABLE 2. Simulation results regarding Type I error rates for the two methods of representing phylogenetic variation for the case in which the environmental component was important to species distributions but not the phylogenetic one.

Tolerance, μ_{tol}	Permutation test		
	Row	Column	Row and column
Type I error rate estimates			
1	1.000	0.014	0.042
3	1.000	0.016	0.048
5	1.000	0.012	0.056
7	1.000	0.020	0.042
9	1.000	0.016	0.036
Amount of explained variation (average R^2 across simulations)			
1	0.033	0.023	0.025
3	0.022	0.016	0.019
5	0.015	0.008	0.012
7	0.004	0.004	0.006
9	0.002	0.003	0.003

Notes: Permutation strategy refers to the type of permutation test used (see last paragraph in section *A simulation study to assess the performance of the framework*). Type I error estimates and average R^2 are only shown for the phylogenetic stochastic simulation because it presented greater power in most scenarios (see Table 3). Type I error rate estimates are based on the rate of rejection out of 1000 simulated data sets.

TABLE 3. Simulation results regarding power for the two methods of representing phylogenetic variation.

Method and tolerance	Power estimate			R^2 (average variation explained across simulations)			r (average correlation with original simulated trait)					
							Phylogenetic stochastic simulation			Phylogenetic eigenvectors		
	Row	Column	Row and column	Row	Column	Row and column	Row	Column	Row and column	Row	Column	Row and column
One Brownian trait												
$\mu_{tol} = 1$	1.000	0.770	0.880	0.379	0.316	0.358	0.697	0.636	0.676	0.479	0.449	0.514
$\mu_{tol} = 5$	1.000	0.784	0.896	0.360	0.298	0.336	0.698	0.634	0.679	0.482	0.464	0.503
$\mu_{tol} = 10$	1.000	0.808	0.904	0.306	0.257	0.289	0.704	0.652	0.683	0.489	0.461	0.531
$\mu_{tol} = 15$	1.000	0.844	0.918	0.249	0.214	0.236	0.712	0.662	0.695	0.506	0.462	0.513
$\mu_{tol} = 20$	1.000	0.890	0.948	0.186	0.172	0.185	0.709	0.681	0.703	0.490	0.480	0.526
$\mu_{tol} = 25$	1.000	0.900	0.968	0.149	0.137	0.145	0.712	0.688	0.710	0.497	0.493	0.536
$\mu_{tol} = 30$	1.000	0.880	0.968	0.106	0.100	0.105	0.708	0.681	0.699	0.504	0.479	0.516
$\mu_{tol} = 35$	1.000	0.886	0.958	0.078	0.074	0.076	0.709	0.680	0.695	0.493	0.504	0.509
$\mu_{tol} = 40$	1.000	0.894	0.960	0.055	0.052	0.054	0.709	0.680	0.695	0.493	0.504	0.509
$\mu_{tol} = 45$	1.000	0.918	0.966	0.040	0.039	0.039	0.697	0.685	0.693	0.484	0.467	0.512
$\mu_{tol} = 50$	1.000	0.902	0.952	0.030	0.028	0.029	0.705	0.678	0.690	0.490	0.489	0.522
$\mu_{tol} = 55$	1.000	0.894	0.960	0.022	0.020	0.021	0.706	0.676	0.694	0.493	0.503	0.521
$\mu_{tol} = 60$	1.000	0.898	0.948	0.016	0.015	0.016	0.699	0.677	0.688	0.503	0.480	0.508
Three Brownian traits												
$\mu_{tol} = 1$	1.000	0.770	0.868	0.387	0.318	0.355	0.699	0.635	0.667	0.491	0.456	0.502
$\mu_{tol} = 5$	1.000	0.822	0.878	0.356	0.318	0.337	0.696	0.656	0.676	0.481	0.471	0.522
$\mu_{tol} = 10$	1.000	0.836	0.918	0.307	0.264	0.296	0.700	0.657	0.697	0.477	0.468	0.535
$\mu_{tol} = 15$	1.000	0.862	0.906	0.246	0.221	0.225	0.704	0.675	0.682	0.497	0.480	0.515
$\mu_{tol} = 20$	1.000	0.906	0.956	0.189	0.181	0.188	0.707	0.686	0.707	0.485	0.489	0.531
$\mu_{tol} = 25$	1.000	0.910	0.948	0.145	0.136	0.141	0.710	0.691	0.701	0.490	0.485	0.529
$\mu_{tol} = 30$	1.000	0.928	0.958	0.107	0.106	0.105	0.706	0.702	0.695	0.488	0.502	0.501
$\mu_{tol} = 35$	1.000	0.910	0.948	0.080	0.074	0.075	0.715	0.686	0.694	0.502	0.483	0.514
$\mu_{tol} = 40$	1.000	0.918	0.956	0.056	0.054	0.055	0.715	0.686	0.694	0.502	0.483	0.514
$\mu_{tol} = 45$	1.000	0.918	0.964	0.041	0.039	0.041	0.704	0.685	0.699	0.492	0.498	0.519
$\mu_{tol} = 50$	1.000	0.900	0.940	0.030	0.028	0.029	0.705	0.674	0.689	0.492	0.479	0.524
$\mu_{tol} = 55$	1.000	0.884	0.926	0.021	0.020	0.021	0.700	0.672	0.676	0.486	0.485	0.499
$\mu_{tol} = 60$	1.000	0.884	0.946	0.016	0.015	0.015	0.705	0.669	0.688	0.470	0.485	0.521
Three phylogenetic eigenvectors												
$\mu_{tol} = 1$	1.000	0.770	0.858	0.527	0.431	0.477	0.855	0.768	0.809	0.897	0.792	0.859
$\mu_{tol} = 5$	1.000	0.756	0.826	0.479	0.404	0.417	0.844	0.763	0.782	0.894	0.790	0.839
$\mu_{tol} = 10$	1.000	0.792	0.864	0.408	0.366	0.372	0.840	0.787	0.801	0.901	0.818	0.872
$\mu_{tol} = 15$	1.000	0.812	0.868	0.335	0.300	0.306	0.833	0.779	0.799	0.909	0.830	0.878
$\mu_{tol} = 20$	1.000	0.788	0.880	0.266	0.239	0.249	0.834	0.765	0.799	0.922	0.829	0.890
$\mu_{tol} = 25$	1.000	0.838	0.880	0.208	0.204	0.196	0.822	0.786	0.790	0.925	0.861	0.895
$\mu_{tol} = 30$	1.000	0.792	0.890	0.155	0.143	0.150	0.809	0.754	0.788	0.930	0.841	0.902
$\mu_{tol} = 35$	1.000	0.832	0.884	0.113	0.108	0.114	0.808	0.772	0.782	0.935	0.869	0.908
$\mu_{tol} = 40$	1.000	0.824	0.870	0.088	0.080	0.078	0.808	0.772	0.782	0.935	0.869	0.908
$\mu_{tol} = 45$	1.000	0.780	0.862	0.060	0.053	0.063	0.791	0.730	0.768	0.944	0.846	0.898
$\mu_{tol} = 50$	1.000	0.794	0.850	0.045	0.042	0.042	0.792	0.739	0.748	0.948	0.857	0.891
$\mu_{tol} = 55$	1.000	0.796	0.892	0.035	0.031	0.033	0.793	0.733	0.765	0.952	0.860	0.915
$\mu_{tol} = 60$	1.000	0.744	0.876	0.026	0.022	0.025	0.780	0.705	0.759	0.952	0.835	0.905

Notes: Permutation strategy refers to the type of permutation test used (see last paragraph in section *A simulation study to assess the performance of the framework*). Power estimates and average R^2 are only shown for the phylogenetic stochastic simulation because it presented greater power in most scenarios. Power estimates are based on the rate of rejection out of 1000 simulated data sets.

structure in these components, thus serving as an estimate of this unknown function. The signs of regression of coefficients on MEMs (or direction of the correlation; Fig. 1C) is arbitrary because the signs of the eigenvectors themselves are arbitrary, and therefore, the correlation can change from positive to negative depending on how they were extracted. Indeed, when dealing with spatial scales based on MEMs or other methods to represent spatial scales (e.g., spatial polynomials; see Legendre et al. 2005), the sign is not important for the interpretation, only the level. Finally, given that these spatial eigenvectors are orthogonal and represent different levels of patchiness, we can directly contrast the contribution (in terms of R^2) of each of these spatial predictors against the phylogenetic mean ($\bar{\mathbf{X}}_\mathbf{P}$) and variance ($\mathbf{s}_\mathbf{P}$) components of metacommunity structure (as in Fig. 1C; see Jombart et al. 2009 for further scale analyses). A complete treatment on how Moran's eigenvector maps are built and embedded into a variation partitioning scheme are described elsewhere

TABLE 4. Simulation results regarding Type I error and power for trait variation.

Tolerance, μ_{tol}	Type I error (random trait)				Type I error (random environment)				Power			
	Row	Column	Row and column	R^2	Row	Column	Row and column	R^2	Row	Column	Row and column	R^2
1	0.836	0.056	0.030	0.008	0.052	0.678	0.052	0.023	1.000	1.000	1.000	0.800
5	0.852	0.056	0.044	0.007	0.074	0.682	0.040	0.019	1.000	1.000	1.000	0.766
10	0.884	0.040	0.036	0.006	0.060	0.734	0.030	0.016	1.000	1.000	1.000	0.682
15	0.884	0.058	0.024	0.005	0.044	0.804	0.040	0.012	1.000	1.000	1.000	0.586
20	0.916	0.052	0.046	0.005	0.056	0.832	0.038	0.011	1.000	1.000	1.000	0.487
25	0.904	0.044	0.018	0.004	0.048	0.892	0.054	0.007	1.000	1.000	1.000	0.386
30	0.934	0.042	0.044	0.003	0.044	0.892	0.052	0.005	1.000	1.000	1.000	0.301
35	0.936	0.048	0.048	0.002	0.064	0.908	0.040	0.005	1.000	1.000	1.000	0.228
40	0.946	0.048	0.054	0.002	0.048	0.940	0.024	0.003	1.000	1.000	1.000	0.168
45	0.966	0.050	0.032	0.001	0.062	0.936	0.038	0.003	1.000	1.000	1.000	0.123
50	0.964	0.058	0.038	0.001	0.052	0.942	0.044	0.002	1.000	1.000	1.000	0.090
55	0.990	0.068	0.022	0.001	0.070	0.964	0.036	0.001	1.000	1.000	1.000	0.067
60	0.984	0.058	0.048	0.000	0.058	0.954	0.036	0.001	1.000	1.000	1.000	0.050

Notes: Estimates are based on the rate of rejection out of 1000 simulated data sets. Permutations strategy refers to the type of permutation test used (see last paragraph in section *A simulation study to assess the performance of the framework*). Type I error (random trait) relates to the case in which the environmental component was important to species distributions but not trait variation. Type I error (random environment) relates to the case in which trait variation was important to species distributions but not environmental variation. R^2 is the average sample coefficient of determination across 1000 sample matrices.

(Dray et al. 2006, Griffith and Peres-Neto 2006, Peres-Neto and Legendre 2010; some additional details are given in Appendix A).

Putting it all together: phylogenetic gradient analyses

Ecologists are now quite familiar with variation partitioning and multi-scale analyses to explore and test hypotheses about the structure of multispecies distributions in relation to environmental and spatial predictors and the rationale for analyzing phylogenetic and trait structure are just as straightforward. Our proposed framework simply formalizes the analyses of phylogenetic and trait structure across environment and spatial variation. Moreover, several studies have already explored the links between phylogenetic/trait structure and environmental and spatial variation (e.g., Ackerly and Cornwell 2007, Swenson et al. 2007, Graham et al. 2009, Willis et al. 2010, Machac et al. 2011, to name a few), but the hypotheses to explain these patterns have not been well explored yet. The following patterns are examples of questions that may be of particular interest in future studies: (1) Do harsh environments tend to select for phylogenetically/trait clustered communities, whereas benign ones select for phylogenetically/trait even communities? (2) What are the environmental and spatial conditions that select for random arrangements of species at local communities? (3) Why is phylogenetic/trait clustering generally associated with large spatial scales, whereas even spreading is generally associated with small spatial scales? (4) Is there a spatial or environmental association with phylogenetically vs. non-phylogenetically structured traits? Our framework simply provides an analytical structure to formalize and test the significance of these links (patterns), allowing ecologists to continue tackling these questions by exploring patterns and testing hypotheses about phylogenetic/trait community structure in a more consistent and robust way.

Here, we use a second set of simulations that allowed us to measure the performance of our variation partitioning framework to metacommunity trait/phylogenetics based on two predictors (**E** and **S**). These simulations will also provide an instinctive way to understand how the nature of different predictor sets (here environment and space) and the relationships between them affect their contribution to the response set (here community structure). We applied a two-gradient protocol recently employed by Ives and Helmus (2011) while assessing the performance of generalized linear mixed models for phylogenetic analyses of community structure. In contrast to our first set of simulations, Ives and Helmus (2011) used matrices that are quite challenging in terms of statistical robustness (small number of sites = 31 relative to number of species = 32). The simulation details and the generation of the simulation scenarios are provided in Appendix C. We considered seven scenarios (I–VII) in which we manipulated the relative contribution of the different fractions of variation in our partitioning scheme (see Fig. 2): (I) [E] and [S] are equally important but independent (Fig. 4A), (II) [E] and [S] are equally important but correlated, (III) [E] and [S] are "independent" but [S] is less important, (IV) [E] and [S] are not important (i.e., Type I error assessment), (V) [S] represents a large spatial scale (as in S_1; Fig. 1A, see Fig. 4B), (VI) [S] represents an intermediate spatial scale (as in S_2; Fig. 1A, see Fig. 4C), and (VII) [E] and [S] represent a large and an intermediate spatial scale, respectively (Fig. 4D).

Because there were two gradients (sources of variation), we simulated two phylogenetic vectors **P**, one for each gradient (see *Representing phylogenetic variation* section), followed by a redundancy analysis (RDA).

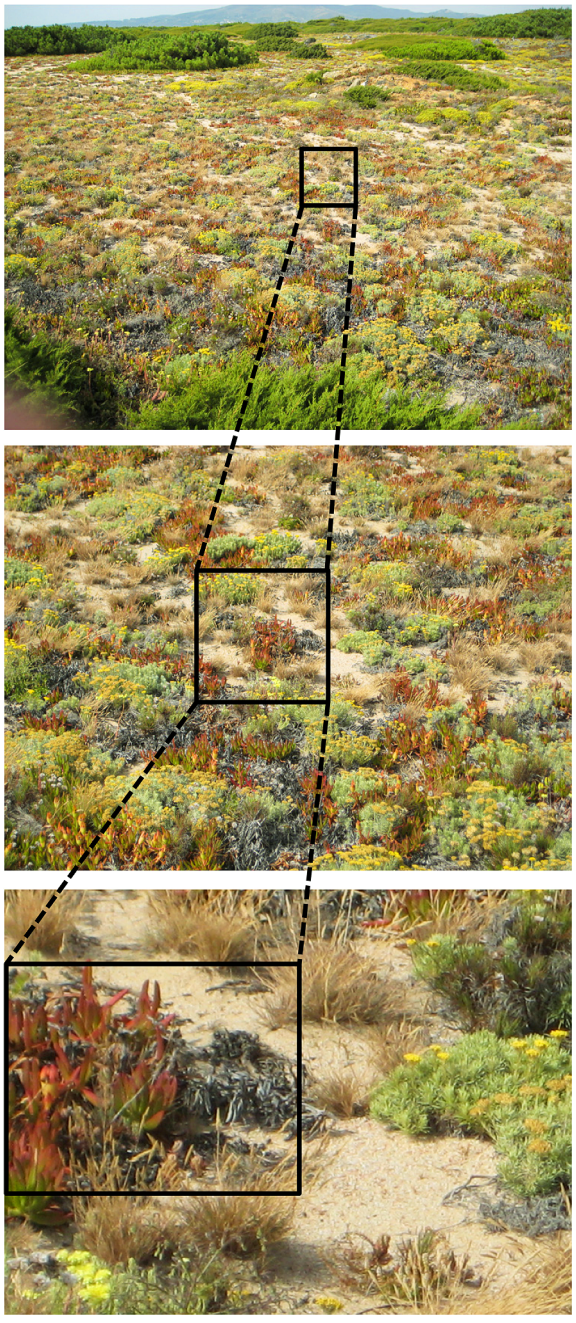

PLATE 1. Spatial-scale dependency in a subtropical Mediterranean (Portugal) plant community. Photo credit: P. R. Peres-Neto.

RESULTS

Type I and II errors

It is clear from our simulation study that the permutation procedure in which the null distribution is based on mixing two null distributions, each based on permuting rows and columns of the environmental matrix **E** and phylogenetic vector **P** separately (see Appendix A), is the only one that yields valid Type I error rates (Tables 1 and 2) in all cases. When species with similar phylogenetic relationships tend to be found with similar abundances regardless of the environment (Table 1), the test based on column permutations is in fact testing this very null hypothesis and therefore cannot provide an unbiased assessment of the hypothesis that the environment is driving species that are phylogenetically closer to co-occur. Note, however, that the row permutation strategy does furnish a correct Type I error when environment is unimportant but phylogeny is (Table 1). In all cases the amount of explained variation is extremely low. Table 2 presents the case in which environment and a non-phylogenetically structured variable (e.g., trait) are important. In this case, both the column and row and column permutation strategies provide a valid test. Given that the row approach does not affect the relationship between phylogeny and distribution, it in fact always detects the effect of environment (Table 2).

Because only the row and column permutation approach provides a valid test across the two situations, we will only interpret the power levels for this strategy. The phylogenetic stochastic simulation as a way to represent phylogenetic variation demonstrates satisfactory power levels across all levels of tolerance (i.e., niche breadth), including the ones that generate quite small R^2 levels (e.g., average tolerance = 60, average $R^2 = 0.025$ across 1000 simulations, and power = 0.876). Power was highest for the stochastic simulation approach when phylogenetic structured data was generated by Brownian simulations (1 and 3 Brownian traits; Table 3). The power was slightly smaller for $\mathbf{P}_{\text{EV-B}}$ in contrast to $\mathbf{P}_{\text{EV}}$ when dealing with data generated using a mix of three eigenvectors. Therefore, because both procedures have correct Type I error levels (i.e., $\mathbf{P}_{\text{EV}}$ and $\mathbf{P}_{\text{EV-B}}$), we suggest that one should use whichever provides the highest R^2. This slight difference in performance of the stochastic simulation ($\mathbf{P}_{\text{EV-B}}$) when analyzing data generated solely on the basis of phylogenetic eigenvectors was to be expected as the correlations between simulated values and tested ones for this case is greater for the eigenvector case (contrast average correlation with simulated trait, Table 3, between the stochastic and phylogenetic eigenvectors). Regardless of the differences, the final conclusion is that the row and column strategy provides a valid test and that both approaches to describe phylogenetic variation are quite powerful in detecting phylogenetic structure in a metacommunity. In the case of traits, the conclusions are the same in which only the row and column permutation strategy provides a valid test (Table 4).

Phylogenetic gradient analyses

The results of the simulation study to assess the performance of our variation partitioning framework to metacommunity phylogenetics are found in Table 5. Because there were two gradients (sources of variation) involved (i.e., **x** and **y**), we simulated two phylogenetic

TABLE 5. Simulation results regarding the assessment of our variation partitioning framework for metacommunity phylogenetics.

Parameter	Simulation scenarios						
	I	II	III	IV	V	VI	VII
Estimates							
Variance ratio	0.826	0.942	0.392	0.696	0.753	0.739	0.714
[E, S]	0.478	0.589	0.349	0.032	0.457	0.44	0.396
E	0.285	0.549	0.270	0.016	0.256	0.264	0.194
S	0.262	0.524	0.115	0.017	0.255	0.239	0.202
[a]	0.216	0.061	0.234	0.015	0.202	0.201	0.194
[b]	0.069	0.488	0.036	0.001	0.054	0.063	0.000
[c]	0.193	0.036	0.079	0.016	0.201	0.176	0.202
[d]	0.522	0.416	0.651	0.968	0.543	0.560	0.604
Rejection rates							
[E, S]	0.841	0.765	0.796	0.055	0.820	0.821	0.816
[E]	0.687	0.750	0.701	0.051	0.597	0.632	0.562
[S]	0.679	0.751	0.418	0.061	0.637	0.627	0.597
[a]	0.597	0.286	0.701	0.052	0.567	0.597	0.562
[c]	0.567	0.250	0.303	0.054	0.597	0.572	0.597

Notes: Variance ratio is the variation (sum-of-squares) due to the compositional ($\bar{\mathbf{X}}_\mathbf{P}$) component of phylogenetic variation over the total amount; the remaining variation is due to the variation component ($\mathbf{s}_\mathbf{P}$). Estimates represent the average contribution of sources of variation ([E, S] represents environment and space combined, E is environment, and S is space) and fractions (see Fig. 2 for their meaning). All tests were based on the row and columns permutation strategy (see last paragraph in section *A simulation study to assess the performance of the framework*). Rejections rates indicate power in all cases except scenario IV (see section *Putting it all together: phylogenetic gradient analyses* for an explanation of each scenario).

vectors **P**, one for each gradient (see *Representing phylogenetic variation* section).

Scenario I shows that the shared amount of variation between the two environmental variables is low, as one would expect given that they were generated independently (though given that the number of sites are quite small, some level of correlation is to be expected). Worth mentioning is the fact that the power of [E] in scenario I is equivalent to the model II in Ives and Helmus (2011) given that they only used **x** in their procedure, though, as here, the response was generated using both **x** and **y**. In scenario II, because **y** was generated as a function of ***X***, the shared component ([b]) increased in comparison to scenario I, reducing the power of the procedure in detecting their independent contributions, i.e., fractions [a] and [c]. As expected, in scenario III, the power to detect the independent fraction was, as expected, greater for **x** [a] than **y** [c] as the contribution of latter has been estimated with a greater amount of error. Scenario IV shows that our framework has correct Type I error for all contributions and fractions. Scenario V and VI show that the procedure is capable of detecting environmental variation that is independent of space. Scenario VII shows that, for two scales, the shared variation [b] is zero given that MEMs are orthogonal in nature and that it correctly detects that both scales had similar influences on species distributions. Taken together, these simulations show that the procedure can be quite robust even under extremes amount of noise in data (Fig. 4).

DISCUSSION

The major goal of our paper was to provide a conceptual foundation to merge community phylogenetic and metacommunity ecology, as well as present a concise and robust quantitative framework to unify two historically divergent but important domains, i.e., trait- and evolutionary-mediated interactions, and local and regional processes. Its flexibility allows the analysis of a range of patterns such as phylogenetic/trait clustering vs. eveness, phylogenetic compositional variation, assessing the number of relevant gradients in structuring phylogenetic or trait structure, and linking metacommunity phylogenetics to comparative biology using stochastic simulations.

There are now a few complementary methods available to tackle some of the issues discussed here (e.g., Pillar and Duarte 2010, Pavoine et al. 2011), though their robustness and flexibility remain to be explored (but see Ives and Helmus 2011). More particularly, our framework shares some features with the one proposed by Ackerly and Cornwell (2007); although they used abundance-weighted averages of traits per community, they did not use a weighted least square procedure and they made use of parametric tests (akin to our row permutation procedure given that they correlated values across sites). Our simulations showed that their procedure yielded large Type I error rates. Pillar and Duarte (2010) also developed methods linking phylogeny, traits, and metacommunity patterns. However, they used a distance-based approach, which does not allow the identification of scale dependent spatial structure and it has reduced power (Legendre et al. 2005). The approach used by Pavoine et al. (2011) has a number of communalities with ours, especially the algebra. However, they only focused on the component of variation related to phylogenetic/trait community composition and not dispersion. Moreover, in their scheme, phylogenetic axes (eigenvectors) were analyzed separately, whereas ours allow both a phylogenetic-scale

decomposition, as well as a way to represent phylogenetic variation into one single vector, hence facilitating interpretation (Fig. 1C). Finally, we developed a way to consider stochastic simulations, which has the ability to dramatically increase the power to detect phylogenetic structure and can be easily embedded as well in any of the current analytical tools.

There are certainly issues with our procedure that need to be resolved. First, our procedure is deeply rooted in variation partitioning and there have been recent criticisms to the method (e.g., Gilbert and Bennett 2010; although we see that the problem is largely due to model specification, i.e., they used linear regressions to model species unimodal responses). Second, the assumption of multiple regressions is that the variance component is constant across all observations, which may be clearly not the case when the phylogenetic/trait dispersion component is large in contrast to the mean one. This also applies to other procedures based on community-weighted means (CWM). Therefore, as the other previous approaches, ours may also suffer from precision regarding estimates of contributions by different predictors, especially when a large proportion of the total variation is due to dispersion rather than mean variation across communities. There are certainly approaches that can accommodate heteroscedasticity, modeling both mean and variance components simultaneously (rather than eliminating variance differences as in some approaches; see Aitkin 1987), though they remain to be extended to a variation partitioning scheme. We plan to explore this issue in greater depth with future studies.

A particular novelty of the framework is the use of macroevolutionary stochastic simulations to link metacommunity ecology and ecological phylogenetics. Although this framework is well rooted in comparative biology (e.g., Garland et al. 1993), it has not made its way into community phylogenetics so far. In this way, our framework links macroevolution to macroecology, opening a door to many other venues such as reconstructions of past community trajectories in similar ways that comparative biology has been used to infer about the way species traits track their environments through evolutionary time (Hansen 1997). A next logical step would be then to consider trait and phylogenetic variation together under a comparative biology framework (Martins 1995), to separate their commonalities as a way to understand how communities are organized regarding phylogenetically labile vs. conserved functional traits contribute to community structure. This is especially important when some ecological traits are more evolutionarily labile than others (see Savage and Cavender-Bares 2012). Finally, evolutionary stochastic simulations allow us to explore how the components of niche variation that are phylogenetically structured relate to different environmental variables and spatial scales.

Work to date shows that species distributions can have important spatial and environmental determinants involving metacommunity processes (e.g., Cottenie 2005, Beisner et al. 2006) and that phylogenetic and trait relationships can be associated with patterns of local coexistence in ways that are thought to reflect ecological processes such as habitat differentiation vs. density dependent niche differentiation (see Cavender-Bares et al. 2009). These inferences, however, are limited by possible confounding factors. For example, spatial pattern that is normally interpreted as reflecting patchiness, mass effects, or neutral dynamics in metacommunity analyses (e.g., Cottenie 2005) could be due to missing but important environmental predictors that are themselves spatially structured (Peres-Neto and Legendre 2010) or historical biogeographic legacies (Leibold et al. 2010). Spatial modeling is a powerful tool because it casts light on these unmeasured spatialized factors or processes, or both.

Although community phylogeneticists should welcome a spatial framework, one should not be tempted to interpret spatial variation as the legacy of dispersal dynamics (Peres-Neto and Legendre 2010). Due to the fact that measuring dispersal dynamics at the metacommunity level is a daunting task (Jacobson and Peres-Neto 2010), one should construe this variation as spatial variation based on unknown sources. Given the advances of GIS techniques, a way to question the origins of these unknown factor/processes that lead to spatial structures is to map the spatial patterns of community structure. Matching spatial and scale patterns on maps is indeed a powerful tool (McIntire and Fajardo 2009) and should be routinely applied. In our framework, one can easily plot predicted phylogenetic/trait community values as an attempt to understand their origins.

Our hope is that our heuristical and analytical frameworks will provide a useful instrument to think about trait and phylogenetic variation under a metacommunity framework. Our simulation protocols should also prove useful while comparing and contrasting the performance of different frameworks in future developments. When analyzing the effectiveness of different alternative methods there is always the question of whether simulated data, rather than real data, represent plausible ecological scenarios and how applicable the conclusions are. The use of simulated data is preferable because their characteristics are known and a systematic change in conditions lead to a better understanding of the main features being evaluated. However, even if the differences are not in statistical performance per se, simulation studies can help resolve many issues, including those related to analytical goals.

Acknowledgments

P. R. Peres-Neto was supported by the Canada Research Chair program (Chair in Spatial Modelling and Biodiversity), NSERC Discovery grants, and the Canadian Foundation for

Innovation. M. A. Leibold was supported by NSF-DEB 0717370. We also thank Jeannine Cavender-Bares, Jonathan Davies, Alexandre Diniz-Filho, Matt Helmus, Liam Revell, and Sharon Strauss for stimulating some of the ideas discussed here. Two anonymous reviewers provided extensive comments on the manuscript. Anthony Ives kindly shared his MATLAB codes that were used in our second set of simulations.

Literature Cited

Ackerly, D. D., and W. K. Cornwell. 2007. A trait-based approach to community assembly: partitioning of species trait values into within- and among-community components. Ecology Letters 10:135–145.

Aitkin, M. 1987. Modelling variance heterogeneity in normal regression using GLIM. Applied Statistics 36:332–339.

Barnett, A., and B. E. Beisner. 2007. Biodiversity and lake trophic state: Explanations invoking resource abundance and distribution. Ecology 88:1675–1686.

Beisner, B. E., P. R. Peres-Neto, E. S. Lindstrom, A. Barnett, and M. L. Longhi. 2006. The role of environmental and spatial processes in structuring lake communities from bacteria to fish. Ecology 87:2985–2991.

Borcard, D., and P. Legendre. 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. Ecological Modelling 153:51–68.

Borcard, D., P. Legendre, C. Avois-Jacquet, and H. Tuomisto. 2004. Dissecting the spatial structure of ecological data at multiple scales. Ecology 85:1826–1832.

Brown, J. H., B. J. Fox, and D. A. Kelt. 2000. Assembly rules: desert rodent communities are structured at scales from local to continental. American Naturalist 156:314–321.

Cadotte, M. W., J. Cavender-Bares, D. Tilman, and T. H. Oakley. 2009. Using phylogenetic, functional and trait diversity to understand patterns of plant community productivity. PLoS ONE 4:e5695.

Cadotte, M. W., and T. Fukami. 2005. Dispersal, spatial scale and species diversity in a hierarchically structured experimental landscape. Ecology Letters 8:548–557.

Cavender-Bares, J., A. Keen, and B. Miles. 2006. Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. Ecology 87(Supplement):S109–S122.

Cavender-Bares, J., K. H. Kozak, P. V. A. Fine, and S. W. Kembel. 2009. The merging of community ecology and phylogenetic biology. Ecology Letters 12:693–715.

Cottenie, K. 2005. Integrating environmental and spatial processes in ecological community dynamics. Ecology Letters 8:1175–1182.

Dale, M. R. T., and M.-J. Fortin. 2009. Spatial autocorrelation and statistical tests: some solutions. Journal of Agricultural, Biological, and Environmental Statistics 14:188–206.

Diamond, J. M. 1975. Assembly of species communities. Pages 342–444 *in* M. L. Cody and J. M. Diamond, editors. Ecology and evolution of communities. Belknap Press, Cambridge, Massachusetts, USA.

Diniz-Filho, J. A. F., C. E. R. Sant'ana, and L. M. Bini. 1998. An eigenvector method for estimating phylogenetic inertia. Evolution 52:1247–1262.

Dray, S., and P. Legendre. 2008. Testing the species traits-environment relationships: the fourth-corner problem revisited. Ecology 89:3400–3412.

Dray, S., P. Legendre, and P. R. Peres-Neto. 2006. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbor matrices (PCNM). Ecological Modelling 196:483–493.

Emerson, B. C., and R. G. Gillespie. 2008. Phylogenetic analysis of community assembly and structure over space and time. Trends in Ecology and Evolution 23:619–630.

Garland, T., Jr., A. W. Dickerman, C. M. Janis, and J. A. Jones. 1993. Phylogenetic analysis of covariance by computer simulation. Systematic Biology 42:265–292.

Gilbert, B., and J. R. Bennett. 2010. Partitioning variation in ecological communities: do the numbers add up? Journal of Applied Ecology 47:1071–1082.

Gonzalez, A. 2000. Community relaxation in fragmented landscapes: the relation between species, area and age. Ecology Letters 3:441–448.

Graham, C. H., J. L. Parra, C. Rahbek, and J. A. McGuire. 2009. Phylogenetic structure in tropical hummingbird communities. Proceedings of the National Academy of Sciences USA 106:19673–19678.

Griffith, D. A., and P. R. Peres-Neto. 2006. Spatial modeling in ecology: the flexibility of eigenfunction spatial analyses. Ecology 87:2603–2613.

Hansen, T. F. 1997. Stabilizing selection and the comparative analysis of adaptation. Evolution 51:1341–1351.

Helmus, M. R., and A. R. Ives. 2012. Phylogenetic diversity–area curves. Ecology 93(Supplement):S31–S43.

Ives, A. R., and M. R. Helmus. 2011. Generalized linear mixed models for phylogenetic analyses of community structure. Ecological Monographs 81:511–525.

Jacobson, B., and P. R. Peres-Neto. 2010. Quantifying and disentangling dispersal in metacommunities: How close have we come? How far is there to go? Landscape Ecology 25:494–507.

Jombart, T., S. Dray, and A. B. Dufour. 2009. Finding essential scales of spatial variation in ecological data: a multivariate approach. Ecography 32:161–168.

Keddy, P. A. 1992. Assembly and response rules: two goals for predictive community ecology. Journal of Vegetation Science 3:157–164.

Kembel, S. W. 2009. Disentangling niche and neutral influences on community assembly: assessing the performance of community phylogenetic structure tests. Ecology Letters 12:949–960.

Kühn, I., M. P. Nobis, and W. Durka. 2009. Combining spatial and phylogenetic filtering in a trait analysis of plant flowering phenology in Switzerland. Global Ecology and Biogeography 18:745–758.

Lavorel, S., and E. Garnier. 2002. Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the Holy Grail. Functional Ecology 16:545–556.

Lavorel, S., K. Grigulis, P. Lamarque, M. P. Colace, G. Garden, J. Girel, G. Pellet, and R. Douzet. 2011. Using plant functional traits to understand the landscape distribution of multiple ecosystem services. Journal of Ecology 99:135–147.

Legendre, P., D. Borcard, and P. R. Peres-Neto. 2005. Analyzing beta diversity: partitioning the spatial variation of community composition data. Ecological Monographs 75:435–450.

Leibold, M. A., E. P. Economo, and P. R. Peres-Neto. 2010. Metacommunity phylogenetics: separating the roles of environmental filters and historical biogeography. Ecology Letters 13:1290–1299.

Leibold, M. A., M. Holyoak, N. Mouquet, P. Amarasekare, J. M. Chase, M. F. Hoopes, R. D. Holt, J. B. Shurin, R. Law, D. Tilman, M. Loreau, and A. Gonzalez. 2004. The metacommunity concept: a framework for multi-scale community ecology. Ecology Letters 7:601–613.

Leibold, M. A., and G. M. Mikkelson. 2002. Coherence, species turnover, and boundary clumping: elements of meta-community structure. Oikos 97:237–250.

Lessard, J.-P., M. K. Borregaard, J. A. Fordyce, C. Rahbek, M. D. Weiser, R. R. Dunn, and N. J. Sanders. 2011. Strong influence of regional species pools on continent-wide structuring of local communities. Proceedings of the Royal Society B 279:266–274.

Levin, S. 1992. The problem of pattern and scale in ecology. Ecology 73:1943–1967.

Lima, S. L., and P. A. Zollner. 1996. Towards a behavioral ecology of ecological landscapes. Trends in Ecology and Evolution 11:131–135.

Machac, A., M. Janda, R. R. Dunn, and N. J. Sanders. 2011. Elevational gradients in phylogenetic structure of ant communities reveal the interplay of biotic and abiotic constraints on diversity. Ecography 34:364–371.

Martins, E. P. 1995. Phylogenies and comparative data, a microevolutionary perspective. Philosophical Transactions of the Royal Society of London B 349:85–91.

McIntire, E. J. B., and A. Fajardo. 2009. Beyond description: the active and effective way to infer processes from spatial patterns. Ecology 90:46–56.

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

Menge, B. A., and A. M. Olson. 1990. Role of scale and environmental factors in regulation of community structure. Trends in Ecology and Evolution 5:52–57.

Paradis, E., J. Claude, and K. Strimmer. 2004. APE: Analyses of phylogenetics and evolution in R language. Bioinformatics 20:289–290.

Parent, C. E., and B. J. Crespi. 2006. Sequential colonization and diversification of Galapagos endemic land snail genus *Bulimulus* (Gastropoda, Stylommatophora). Evolution 60:2311–2328.

Pavoine, S., E. Vela, S. Gachet, G. de Bélair, and M. B. Bonsall. 2011. Linking patterns in phylogeny, traits, abiotic variables and space: a novel approach to linking environmental filtering and plant community assembly. Journal of Ecology 99:165–175.

Pedruski, M. T., and S. E. Arnott. 2011. The effects of habitat connectivity and regional heterogeneity on artificial pond metacommunities. Oecologia 166:221–228.

Peres-Neto, P. R., and P. Legendre. 2010. Estimating and controlling for spatial autocorrelation in the study of ecological communities. Global Ecology and Biogeography 19:174–184.

Peres-Neto, P. R., P. Legendre, S. Dray, and D. Borcard. 2006. Variation partitioning of species data matrices: estimation and comparison of fractions. Ecology 87:2614–2625.

Pillar, V. D., and L. D. S. Duarte. 2010. A framework for metacommunity analysis of phylogenetic structure. Ecology Letters 13:587–596.

Presley, S. J., C. L. Higgins, and M. R. Willig. 2010. A comprehensive framework for the evaluation of metacommunity structure. Oikos 119:908–917.

Revell, L. J. 2010. Phylogenetic signal and linear regression on species data. Methods in Ecology and Evolution 1:319–329.

Ricklefs, R. E. 1987. Community diversity: relative roles of local and regional processes. Science 235:167–171.

Ricotta, C., and M. Moretti. 2011. CWM and Rao's quadratic diversity: a unified framework for functional ecology. Oecologia 167:181–188.

Savage, J. A., and J. M. Cavender-Bares. 2012. Habitat specialization and the role of trait lability in structuring diverse willow (genus: *Salix*) communities. Ecology 93(Supplement):S138–S150.

Seidler, T. G., and J. B. Plotkin. 2006. Seed dispersal and spatial pattern in tropical trees. PLoS ONE 4:e344.

Sharma, S., P. Legendre, M. De Caceres, and D. Boisclair. 2011. The role of environmental and spatial processes in structuring native and non-native fish communities across thousands of lakes. Ecography 34:762–771.

Simberloff, D. 2004. Community ecology: is it time to move on? American Naturalist 163:787–799.

Strauss, S. Y., C. O. Webb, and N. Salamin. 2006. Exotic taxa less related to native species are more invasive. Proceedings of the National Academy of Sciences USA 103:5841–5845.

Swenson, N. G., B. J. Enquist, J. Thompson, and J. K. Zimmerman. 2007. The influence of spatial and size scales on phylogenetic relatedness in tropical forest communities. Ecology 88:1770–1780.

Vellend, M. 2010. Conceptual synthesis in community ecology. Quarterly Review of Biology 85:183–206.

Webb, C. O., D. D. Ackerly, M. A. McPeek, and M. H. Donoghy. 2002. Phylogenies and community ecology. Annual Review Ecology and Systematics 33:475–505.

Webb, C. O., C. H. Cannon, and S. J. Davies. 2008. Ecological organization, biogeography, and the phylogenetic structure of tropical forest tree communities. Pages 79–97 *in* W. P. Carson and S. A. Schnitzer, editors. Tropical forest community ecology. Wiley-Blackwell, Oxford, UK.

Wiens, J. A. 1989. Spatial scaling in ecology. Functional Ecology 3:385–397.

Willis, C. G., M. Halina, C. Lehman, P. B. Reich, A. Keen, S. McCarthy, and J. Cavender-Bares. 2010. Community phylogenetic structure in Minnesota oak savanna is influenced by spatial extent and environmental variations. Ecography 33:565–577.

Supplemental Material

Appendix A

Computational details of our metacommunity phylogenetics framework (*Ecological Archives* E093-175-A1).

Appendix B

Analytical details on representing phylogenetic variation (*Ecological Archives* E093-175-A2).

Appendix C

Simulations regarding the phylogenetic gradient analyses (*Ecological Archives* E093-175-A3).

Supplement

MATLAB functions for the analytical framework for metacommunity trait/phylogenetics (*Ecological Archives* E093-175-S1).