

Mistaking geography for biology: inferring processes from species distributions

Dan. L. Warren¹, Marcel Cardillo², Dan F. Rosauer³, and Daniel I. Bolnick⁴

¹ Department of Biology, Macquarie University, North Ryde, NSW 2109, Australia

² Macroevolution and Macroecology Group, Research School of Biology, The Australian National University, Canberra, ACT 0200, Australia

³ Centre for Biodiversity Analysis, Research School of Biology, The Australian National University, Canberra, ACT 0200, Australia

⁴ Howard Hughes Medical Institute and Department of Integrative Biology, University of Texas at Austin, One University Station C0990 Austin, TX 78712, USA

Over the past few decades, there has been a rapid proliferation of statistical methods that infer evolutionary and ecological processes from data on species distributions. These methods have led to considerable new insights, but they often fail to account for the effects of historical biogeography on present-day species distributions. Because the geography of speciation can lead to patterns of spatial and temporal autocorrelation in the distributions of species within a clade, this can result in misleading inferences about the importance of deterministic processes in generating spatial patterns of biodiversity. In this opinion article, we discuss ways in which patterns of species distributions driven by historical biogeography are often interpreted as evidence of particular evolutionary or ecological processes. We focus on three areas that are especially prone to such misinterpretations: community phylogenetics, environmental niche modelling, and analyses of beta diversity (compositional turnover of biodiversity).

Biogeographic patterns in evolution and ecology

Some of the most profound and long-standing questions in ecology and evolution remain open. To what extent are ecological communities shaped by competitive exclusion? How do species' niches evolve? What are the relative roles of biotic and abiotic factors in dictating species' distributions and abundance? How often is speciation allopatric, parapatric, or sympatric? These questions continue to challenge biologists in part because they are inherently historical, concerning dynamics that play out over long time-scales that we cannot feasibly study directly. Consequently, biologists address such questions by searching for biogeographic fingerprints of past dynamics. For example, competitive exclusion is expected to reduce range overlap between close relatives, so early ecologists asked whether

the number of co-occurring species per genus is lower than expected [1]. The modern equivalent of this research program, community phylogenetics, quickly became a major theme in ecology over the past decade [2]. However, such biogeographic approaches to inferring past ecology have an unfortunate tendency to ignore evolutionary history. Conversely, biogeographic studies of the geography of speciation often ignore ecological processes affecting co-existence [3,4]. Here, we argue that both ecologists and evolutionary biologists need to better incorporate both perspectives into their analyses (and null models) of historical biogeography.

Glossary

Age range correlation (ARC): method of analysing geographic range overlap between related species as a function of time since divergence.

Compositional turnover or beta diversity: the difference in species composition between two areas, generally measured as a function of the number of species common to both areas, and the number found only at one area or the other.

Ecological sorting: the processes by which the set of species that co-occur in a given area is filtered from a broader regional pool of species, based on habitat suitability or biotic interactions, such as competitive exclusion. The term is usually applied as a distinction to evolutionary processes such as adaptation, with the implication that ecological sorting is a post-evolutionary mechanism operating on much shorter timescales.

Niche identity: the hypothesis that the ecological or environmental niches of two species are so similar as to be effectively interchangeable, so that data from the two species are effectively drawn from the same distribution; assumes that the rate of niche evolution is very slow.

Niche or distribution model: often called 'environmental niche models' (ENM) or 'species distribution models' (SDM), these are statistical models that infer environmental tolerances and potential spatial distributions of species by correlating the occurrences of those species with environmental predictor variables.

Niche similarity: the hypothesis that two species share some level of ecological or environmental niche similarity because of common ancestry; assumes that the rate of niche evolution is not so fast that it eliminates all signal of common ancestry in the ecological or environmental preferences of those species.

Phylogenetic beta diversity: the difference in composition between two areas measured as a function of the amount of branch length shared (and not shared) between biota of the two sites; indicates how closely related the biota of the two sites are.

Phylogenetic clustering: when a set of co-occurring species is more closely related to one another than an equal number of species drawn at random from a broader species pool.

Phylogenetic overdispersion: when a set of co-occurring species is less closely related to one another than an equal number of species drawn at random from a broader species pool.

Corresponding author: Warren, D.L. (dan.l.warren@gmail.com).

Keywords: biogeography; evolutionary ecology; ecological niche modelling; species distribution modelling; community phylogenetics; biodiversity; biodiversity turnover.

0169-5347/

Crown Copyright © 2014 Published by Elsevier Ltd. All rights reserved. <http://dx.doi.org/10.1016/j.tree.2014.08.003>

Analyses that infer biological processes from species distributions

Many popular statistical tests in ecological biogeography are predicated on the idea that ecological processes can be inferred from the distributions of species. In this opinion article, we focus on three analytical approaches: community phylogenetics, environmental niche models, and beta-diversity analyses. In community phylogenetics, co-occurrence of species within local communities is tested against a null distribution to determine whether close relatives co-occur less often than expected by chance, which is frequently interpreted as evidence of competitive exclusion between closely related species. In studies of niche evolution using environmental niche models (ENMs), researchers often test whether niche models of closely related species are more similar than expected compared with some null distribution. Models of beta diversity are used to study the ways in which environment and space explain compositional differences between communities.

These three analytical approaches all aim to identify ecological processes by comparing modern-day species distributions against null expectations that omit the focal ecological process. The null hypotheses are constructed by assuming that all species in the system have the potential to co-occur in communities or occupy similar environments. However, species distributions and co-occurrence are influenced by multifarious historical, ecological, and evolutionary factors. Communities are assembled through both colonisation by foreign species and speciation, both of which can lead to non-random distributions of species on a landscape. Certain areas on a landscape are more likely to receive and sustain invasive species [5,6]. New species are also less likely to co-occur with their closest relatives, either due to allopatric or parapatric speciation [7].

Although there is increasing evidence that speciation can happen in the presence of gene flow [7–9], it is still generally accepted that allopatric speciation is the predominant mode of speciation in many groups [10]. Following speciation, species' ranges change (translation, expansion, or contraction) due to both non-ecological processes (stochastic dispersal, evolution of dispersal behaviour [11]), and ecology (species interactions, abiotic conditions, stochastic processes). Therefore, species distributions are perpetually in flux for multiple reasons [12]. Consequently, even in the absence of any ecological interactions, species will be found non-randomly on a landscape. This non-random distribution is the background against which the dynamics that interest ecologists occur: niche evolution, range shifts due to climate change, and species interactions. Thus, modern-day biogeography represents the effects of the geography of speciation and subsequent range shifts (hereafter collectively referred to as 'historical processes'), as well as ecological processes.

We argue that these historical processes generate statistical results that are often interpreted as ecological phenomena. That is, allopatric speciation and colonisation dynamics mean that close relatives tend not to coexist in local communities, even in the absence of competition or niche evolution. Models of beta diversity can explicitly consider allopatric speciation (treating distance as a predictor of turnover), but their linear conception of distance

may be insufficient to capture the geography of speciation, which may also entail ecologically-driven reproductive isolation over fine spatial scales. In this opinion article, we examine the ways in which the geographic signal of evolutionary history in allopatrically (or parapatrically) diversifying clades can produce spurious inferences regarding species ecology from widely used statistical tests.

Community phylogenetics

The fundamental aim of community ecology is to understand the processes that produce two kinds of pattern: the structure of communities and the coexistence of species. Although closely linked, these two ecological patterns are usually addressed separately by two distinct approaches. Community structure is addressed using 'whole-assemblage' metrics that capture some aspect of the emergent structure of a community. In community phylogenetics, the most widely used whole-assemblage metrics are the Net Relatedness Index (NRI) and Nearest Taxon Index (NTI) [2,13], which allow a given community to be described as phylogenetically 'clustered' or 'overdispersed' compared with a regional species pool. By contrast, species coexistence is addressed using tests in which pairs of species are the units of analysis. The typical analysis in this approach is to calculate a metric quantifying the degree of co-occurrence of pairs of species across several sampling units, and plot this against the phylogenetic relatedness of each species pair (see Figure 1A in Box 1).

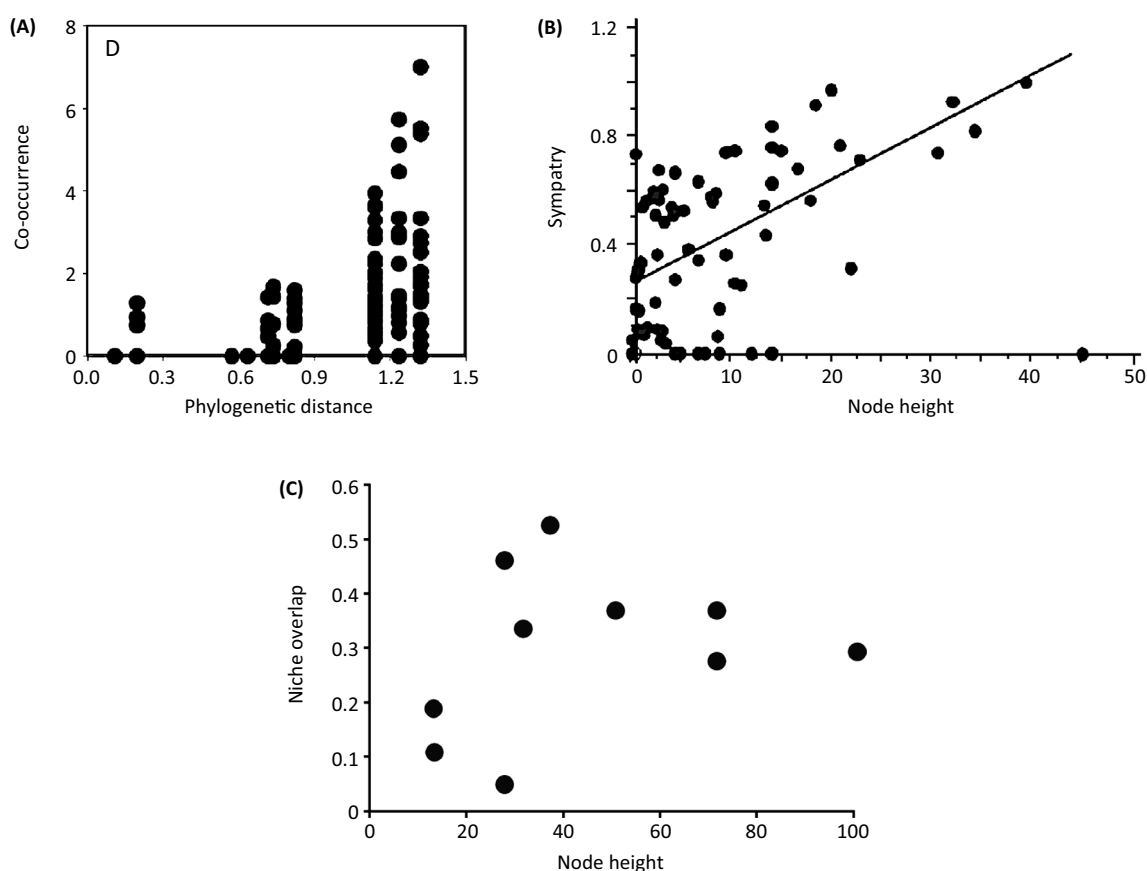
Overwhelmingly, community phylogenetics has continued the traditional emphasis of community ecology on patterns at spatially restricted ('local') scales, and on mechanisms based on ecological sorting, with interspecific competition as a process of particular interest. This emphasis began with the paper that sparked off the surge of interest in community phylogenetics [13], which introduced NRI and NTI to test two ecological-sorting hypotheses: that rainforest tree communities are structured primarily either by interspecific competition or by environmental filtering. Later papers used a pairwise approach to test the same two hypotheses (e.g., [14–18]). Assuming that phylogenetic relatedness reflects ecological similarity, and that all species in the study area have access to all sites, the expectation is that competition structuring will produce a positive association between phylogenetic distance and degree of co-occurrence across species pairs, whereas environmental filtering will produce a negative association (Figure 1A in [14]).

Age range correlation (ARC, Box 1) has developed over a longer period of time compared with community phylogenetics (the seminal paper was in 1989 [19]), but its impact has been far more subdued, with published papers numbering in the dozens rather than the hundreds. Although the methods have been refined over the years [3,20,21], ARC is essentially the same as the pairwise approach in community phylogenetics. Yet ARC has developed entirely separately from, and independently of, community phylogenetics, and it seeks to answer questions about the geography of speciation, not about ecological processes of competition or environmental filtering. Whereas community phylogenetics interprets a positive association between phylogenetic distance and local co-occurrence as

Box 1. ARC and ecological biogeography

ARC is an analytical approach that combines phylogenies and species distributions. The basic method of ARC is to plot the phylogenetic relatedness of pairs of taxa against the degree of overlap in their distributions [19,80]. ARC analyses demonstrate mixed results regarding the geography of speciation, but over 70% of sister pairs in studies reviewed by Fitzpatrick *et al.* [10] were entirely allopatric. However, the species considered in that review were primarily vertebrates; many groups of plants appear to have considerable range overlap even between relatively recently derived species pairs [8,9]. In ARC, a low intercept and positive slope on the relatedness \times overlap plot is interpreted not as the imprint of any specific ecological process, but as the historical signal of allopatric speciation followed by post-speciation range shifts that increase levels of sympatry.

Figure 1 provides examples of associations between species relatedness and relative distributions, interpreted in different ways. Figure 1A shows increasing degrees of local-scale co-occurrence of reticulate-sheathed *Schoenoid* sedges with increasing phylogenetic distance [18], interpreted as the possible influence of competition on community structure. Figure 1B shows increasing degrees of broad-scale geographic range overlap of primates with increasing phylogenetic distance, interpreted as the imprint of allopatric speciation followed by gradual range shifts [81]. Figure 1C shows increasing environmental overlap of Cuban anoles with increasing phylogenetic distance, interpreted as rapid niche evolution [39].



TRENDS in Ecology & Evolution

Figure 1. Examples of associations between species relatedness and relative distributions, interpreted in different ways. Reproduced, with permission, from [18] (A), [81] (B), and [39] (C).

the effect of competition (see Figure 1A in Box 1), ARC interprets the exact same association between phylogenetic distance and geographic overlap as the imprint of allopatric speciation (see Figure 1B in Box 1).

How has the situation arisen in which two independent subdisciplines of ecology and evolution are using essentially the same method to address two different questions? The answer is the issue of spatial scale. It is widely considered appropriate to address the two kinds of process, ecological sorting and speciation, using species distribution data at different spatial resolutions. To detect an effect of competition on species coexistence or community structure, it has long been considered necessary to adopt a spatially restricted circumscription of a community, under

the assumption that larger areas offer more potential for stable coexistence of potential competitors (e.g., [22]). By contrast, the arrangement of the geographic distributions of species on broad scales is considered more likely to carry the signal of macroevolutionary and biogeographic processes such as speciation and range shifts (e.g., [23]). However, are these assumptions really warranted? Perhaps not: it is fairly clear that, on the one hand, the legacy of speciation can impinge on the patterns of community phylogenetics, and on the other, that ecological sorting processes can influence species distributions at large geographic scales.

As an example of how speciation history can influence community patterns at local scales, consider two recent papers on the community phylogenetics of the plant genus

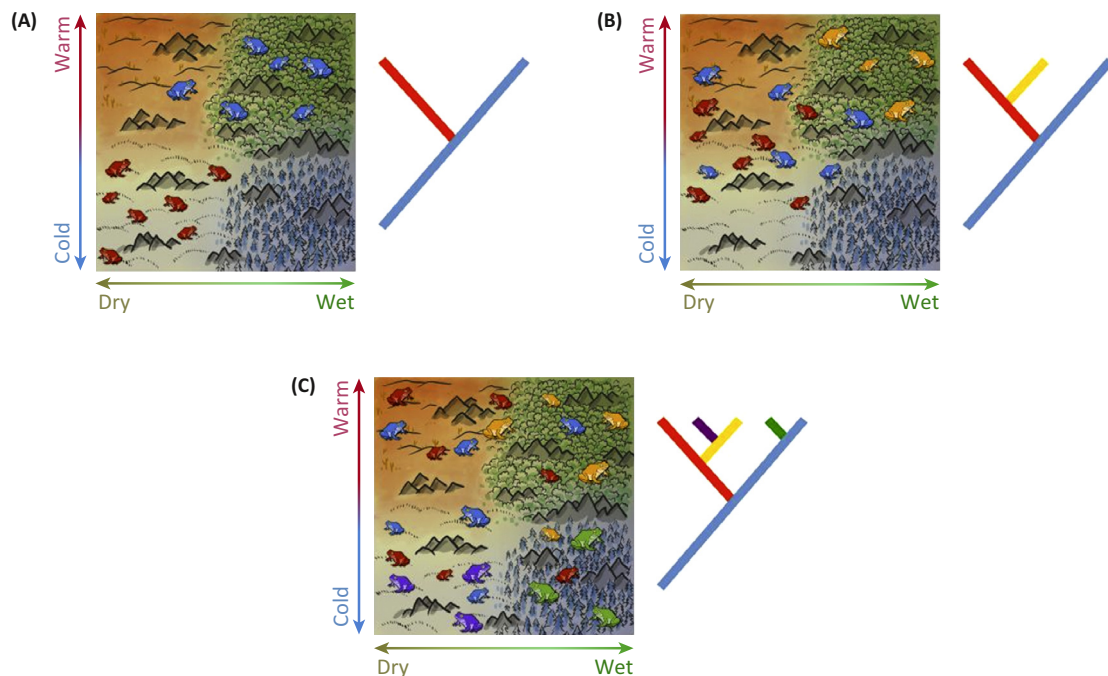
Box 2. Range and ENM overlap in a clade diversifying allopatrically

Figure 1 illustrates a process of diversification similar to that discussed in the main text. Diversification begins with allopatric speciation (Figure 1A). Subsequent speciation events (Figure 1B) result in allopatry between sister species (red and yellow frogs), whereas shifting ranges may lead to secondary range overlap between more distantly related species (red and blue), which also results in increased environmental similarity because of spatial autocorrelation present in the environment. A typical end result may be seen in Figure 1C: sister species exhibit little overlap in range or environmental variables (yellow and purple allopatric, blue and green allopatric), whereas more distantly related species pairs may overlap considerably both in geographic and environmental space (e.g., red and blue, red and yellow). An ARC analysis in this system would produce a plot similar to that seen in Figure 1B in Box 1.

Given this scenario, we can expect to obtain three results from commonly used methods of analysis using niche models (e.g., [17]). First, sister species will not be identical in their niche because of allopatry in a spatially autocorrelated environment. Second, sister species will be more similar than expected given the background, because the environmental background of all species contains some

patches of habitat that no species can tolerate (e.g., mountains in this example, but this is true if there is any environment outside of the union of the environmental tolerances of the species within the study area). Finally, a phylogenetic analysis of niche evolution will show greater divergence between sister species than it does between more distantly related pairs.

Unfortunately, we are unable to distinguish between two interpretations of these results that are almost diametrically opposed. The patterns seen in the spatial and environmental distributions of closely related species may reflect ecological phenomena; for instance, it is possible that competition between closely related species renders it impossible for them to coexist, leading them to occupy and eventually adapt to local climate regimes. Subsequent evolution along non-environmental niche axes could enable coexistence, allowing more distantly related species to occupy similar geographic (and hence environmental) distributions. However, these same statistical results would be expected based solely on the environmental signal of allopatric speciation and secondary range overlap, even if all species were effectively identical in their environmental niches and did not compete.



TRENDS in Ecology & Evolution

Figure 1. (A–C) Allopatric speciation and subsequent range overlap in a hypothetical clade. Illustration by Ainsley Seago.

Banksia in southwestern Australia. Merwin *et al.* [24] analysed the structure of *Banksia* communities across a large geographic region using whole-assemblage metrics, and found a prevalent pattern of phylogenetic clustering. That is, species within communities were more closely related to each other than expected, when compared with null communities constructed by shuffling species across the entire study region. This is the pattern expected if communities are assembled primarily by environmental filtering, but do the findings of Merwin *et al.* [24] allow us to conclude that environmental filtering is the primary mechanism of *Banksia* community assembly? Not necessarily, because many *Banksia* species in southwestern Australia are distributed allopatrically and have very restricted geographic distributions that are far smaller than the size

of Merwin *et al.*'s study region, which are most likely the legacy of rapid speciation and limited dispersal [25]. Therefore, a null model constructed from the entire region inevitably predisposes many communities to being phylogenetically clustered (to be fair to the authors, they were fully aware of this issue, and they acknowledged the likely contribution of speciation history to the prevalence of clustering in their results). The second paper analysed the community phylogenetics of southwest Australian *Banksia* using a pairwise approach [26]. In this study, the study sites were distributed across a far smaller region than those of Merwin *et al.* [24], and the author explicitly attempted to minimise the signal of speciation history by removing from the analysis pairs of species showing no overlap in their broad geographic distributions. This

analysis revealed no evidence for phylogenetic clustering; in fact, there was a significant pattern of phylogenetic overdispersion. The contrasting results of these two approaches to analysing community phylogenetics of the same set of species in the same region illustrates the potential for macroevolution and biogeography to impinge on the kinds of patterns that have, traditionally, been within the remit of community ecologists. Again, to be fair, many authors of community phylogenetics papers are aware of this, and attempt to deal explicitly with the issue by defining their study region to encompass a pool of species that are likely to have access to all sites included in the study (e.g., [14,17,27]). Of course, defining an appropriate source pool is itself fraught with assumptions, since species often differ in their abilities to disperse across a given region.

If macroevolution and biogeography can influence local community-level patterns, can ecological sorting influence species distributions (and, by extension, assemblage structure) at biogeographic scales? The answer is certainly yes: we need think only of clear examples in which the expansion of an invasive species has coincided with the contraction of the range of a native species, such as grey and red squirrels (*Sciurus carolinensis* and *S. vulgaris*) in Britain [28], or native and invasive bumblebees (*Bombus dahlbomii* and *B. terrestris*) in Patagonia [29]. Indeed, there is evidence for the imprint of competition on broad-scale assemblage structure using macroecological [30] and community-phylogenetic [31,32] methods.

Nonetheless, there still seems to be a prevailing mindset among researchers that dichotomizes the scale of species assemblages into 'local' and 'regional' [33,34]. This dichotomy is often accompanied by an expectation that ecological sorting processes are important at local scales, and macroevolutionary or biogeographic processes are important at large scales. However, this dichotomy of scale is probably false, since species disperse and interact across a wide variety of scales, and it seems more appropriate to consider that assemblages can be circumscribed across a continuum of scales, from 'local' to 'regional' or beyond [34]. We believe that the study of biodiversity would be well served by dispensing with a binary division of assemblages by spatial scale and the accompanying *a priori* expectations about processes.

Hypothesis testing using niche models

In comparative studies of niche models, locality data is used to construct estimates of ecological tolerances of species. Similarity between models for related species is then tested against null distributions to examine whether species ENMs are more or less similar than expected given some null hypothesis, or whether there is phylogenetic signal in the pattern of ENM overlap (e.g., [35–39]). There is still considerable disagreement about what aspects of the niche these models may estimate (e.g., [40–45]). However, for these models to be informative in studies of niche evolution, one must only assume that evolutionary processes affecting the niche are accurately represented in the differences between species ENMs, not that individual ENMs are necessarily accurate or complete niche estimates. The extent to which this assumption is valid is

usually unknown, and is likely to differ across study systems. In some cases, this assumption may be met with respect to the realised niche only, which is rarely the aspect of the niche that is of interest for evolutionary studies. In these cases, allopatric species analysed with ENM-based methods, such as those in Warren *et al.* [39], will frequently appear to be ecologically divergent even in the absence of evolution of environmental tolerances [39,46]. If spatial autocorrelation is sufficiently extreme, a phylogenetic test of ecological divergence using niche models will effectively be an ARC test filtered through a set of proxy variables, with a questionable ecological interpretation.

Given spatial autocorrelation in the environment, the extent to which the ranges of any two species overlap will place upper and lower bounds on the environmental similarity possible between them. Consider a clade in which speciation is allopatric, ranges drift over time, and the environment is spatially autocorrelated (Box 2). If niche evolution is investigated in this clade using ENMs (e.g., [37–39,47,48], many others reviewed in [49]), we may expect the three statistical patterns outlined in Box 2: sister species will not be identical in their environmental niches, but will typically be more similar than expected given the suites of environments available to them, and the shallowest clades in the phylogeny will often demonstrate the greatest levels of divergence. Of crucial importance is the fact that no assumptions need be made about niche evolution to obtain these results; they are expected strictly based on allopatric diversification and subsequent range shifts in a spatially autocorrelated environment.

Perhaps not coincidentally, these three patterns (niche similarity but not identity, rapid evolution in the shallowest clades on a phylogeny) are those most commonly seen in the empirical literature. In a recent review, Peterson ([49], supplementary information) surveyed studies using ENM-based methods to investigate niche evolution. Surveying 85 papers, Peterson found 249 out of 291 tests supporting niche similarity, with 22 more displaying mixed results. None of the six tests of niche identity included in that review supported identity. Although there have been relatively few phylogenetic studies using ENMs, the most common pattern seen is the one discussed above; as noted by Losos [50], it is common to find little evidence of phylogenetic signal in the environmental niche (e.g., [37,39,47,48,51]). Although it is possible that this indicates that the environmental niche is labile, it may simply be a secondary effect of the prevalence of allopatric speciation and the insufficiency of species distributions to fully characterise their environmental tolerances.

Therefore, we must ask how frequently we expect geography alone to produce these dominant empirical trends. Unfortunately the necessary conditions seem quite common: ARC analyses show that allopatric speciation and secondary range overlap are common in many groups (Box 1). Research comparing physiological limits to the limits of species geographic distributions shows that some species distributions may be primarily limited by physiological tolerances (e.g., [52]), but other processes, both ecological and historical, also have a role for many species [43]. Spatial autocorrelation is a ubiquitous feature of environmental variables [53,54]. Given that all of the necessary

conditions are frequently met, we are faced with the grim possibility that much of the research on niche evolution using ENMs may be primarily driven by the geography of speciation, rather than by the inferred ecological processes. This is not to say that the inferred patterns of niche evolution are necessarily wrong; rather, it suggests that these tests may be uninformative in some common situations. It is important to note that these concerns are not exclusive to analyses using ENMs; comparative studies using environmental data extracted from range maps or occurrence points [55,56] are subject to the same biases.

Newer analytical methods are available that correct for spatial autocorrelation in environmental predictors, which may lead to valuable new insights with less potential to be confounded by strictly spatial processes. Freckleton and Jetz [51] presented a method that disentangles spatial and phylogenetic autocorrelation in species environmental niches, and found that patterns in the environmental niche tend to be more driven by spatial autocorrelation in species ranges than by evolutionary history. Broennimann *et al.* [35] demonstrated a set of methods that extend those in Warren *et al.* [39], allowing hypothesis tests that correct for the expected similarity based on spatial autocorrelation in the environment. However, these methods remain underutilised compared with more accessible methods that do not correct for spatial autocorrelation (e.g., [57]).

Much attention has been devoted to processes that limit our ability to use ENMs; processes such as biotic interactions and dispersal limitation may limit species distributions, and the effects of these non-target phenomena are included in models through spurious spatial correlations with environmental predictors. The geographic pattern of diversification within a clade may have similar effects on inferences of the evolutionary history of the niche, because the pattern of species ranges and environmental distributions may be spatially and temporally autocorrelated for reasons having little to do with the ecological processes we are attempting to infer.

Modelling the drivers of compositional turnover

The spatial structure of biodiversity can also be described and explained through the lens of compositional turnover or beta diversity: the difference between the biotic assemblages at different locations. The proportion of species, alleles, or other evolutionary units shared between sites tends to decrease with the separation in geographic or environmental space [58]. Numerous studies have investigated the drivers of beta diversity to identify the processes that determine the spatial structure of compositional turnover between locations [27], and to develop predictive models of turnover for application in biodiversity conservation and macroecology [59]. The term ‘beta diversity’ is used in various ways [60], but in this article, we refer to the dissimilarity between two assemblages at different places or times. Turnover in community composition can be seen as an aggregate property of the phenomena discussed above: community level interactions, ecological niches, and the spatial signal of evolutionary history.

Compositional turnover can be statistically partitioned into components explained by environment, by location, or by both together [61], quantifying the relation between

matrices of compositional, geographic, and environmental distance via techniques including canonical partitioning [61], Mantel tests, null models of expected turnover [62], and generalised dissimilarity models (GDM) [59,63]. The component of turnover associated with environment has a straightforward explanation. Species are replaced as one moves along an environmental gradient, entering or leaving the range of suitable conditions for each species at points along the gradient. If ecological niche were the sole determinant of species distributions, with each species occupying all areas of suitable environment, we could expect an almost perfect relation between environmental and compositional turnover. Explaining turnover attributed to distance is less straightforward. Under a neutral model of biodiversity [64] the difference in species composition between sites should be stochastic, decaying logarithmically with increasing geographic distance until the probability of sites sharing any species is close to zero [65,66]. However, the geographic signal of historical diversification can result in turnover, independent of environment, due to a range of spatially structured nonstochastic phenomena, including isolation by past and current barriers to dispersal, which do not always correlate to linear distance.

The conflation of environment and geography is as much of an issue when identifying drivers of beta diversity as it is for ENMs, but it has arguably been handled somewhat better for beta diversity. Methods for examining beta diversity often explicitly account for both environment and distance as drivers of turnover, and also identify a component explained jointly by location and environment [61,66,67]. Let us consider again the case of allopatric sister species each occupying just part of their fundamental niche (e.g., Box 2). For an ENM, the difference in locations of the species may be misattributed to niche evolution, but for a species beta model, this turnover of composition with little environmental difference could be attributed to distance, or simply as unexplained ‘noise’ in the model. However, if one explicitly incorporates evolutionary relationships, modelling not species but phylogenetic beta diversity [59,68], then the occurrence of sister species in identical or similar environments (a conjunction of small phylogenetic and small environmental turnover) could be entirely captured in the model. A good model would attribute the shared evolutionary history of these sister species to the similar environment, and their limited phylogenetic dissimilarity at the tips to isolation.

We highlight three ways in which erroneous conclusions about the drivers of compositional turnover could arise. First, in an analysis of *species* turnover, if the correlation with environment occurs above or below species level, then turnover resulting from environmental gradients may be unexplained or mistakenly attributed to distance. Swenson [69] suggested and Rosauer *et al.* [70] found that phylogenetic beta diversity was more strongly related to environment in a GDM than was species beta diversity, because species are more likely to be replaced by close than distant relatives along an environmental gradient. However, the relationship of turnover to environment peaked at a particular tree depth in each case, with deeper branches confounding the relationship. Just as longer

diverged species are more often found in sympatry, (see Figure 1B in Box 1) deeper phylogenetic clades are more likely to overlap in geographic or environmental space, weakening the relationship of environment to phylogenetic beta diversity. Although it is possible that this phenomenon is a meaningful signal of niche evolution, it may simply be due to the fact that older clades generally have more extant descendant lineages and more time in which allopatrically or parapatrically diversifying species can develop secondary range overlap.

Second, the rate of turnover is rarely constant along an environmental gradient, with some parts of the gradient associated with far more compositional turnover than others [71]. For example, Fitzpatrick *et al.* [71] found far more compositional turnover in southwest Australian plants associated with differences in winter precipitation below 200 mm, than with a difference of the same magnitude at a wetter part of the gradient. As a result, models that assume a constant rate of decay in compositional similarity with environmental distance, rather than capturing the variable rate (non-stationarity) in environment-driven turnover [59,72], may indicate that environment explains less of the turnover than it does, and favour explanations based on isolation or history.

Third, although geographic distance is usually a strong predictor of beta diversity, it is merely a surrogate for the isolation that can generate and maintain distinct evolutionary histories. Isolation is important to turnover specifically because it supports the development and maintenance of separate evolutionary histories at all levels, including locally divergent lineages, allopatric speciation, and endemic radiations. This isolation varies with location and direction [73] and depends on the dispersal capabilities of the organisms in question and the suitability of the intervening environment. It also depends, of course, on the legacy of historical environments and barriers, and their strength and duration over evolutionary timescales, none of which are correlated in a simple way to distance.

A challenge here is not only to quantify the amount of turnover explained by isolation, but also to develop spatially explicit models that define the locations and strength of isolation for a given group of taxa. Although we can quantify for a known barrier, such as Wallace's Line or the Andes [74], the amount of turnover that can be attributed to crossing the barrier itself rather than to spatial or environmental separation, it is harder to model for more diffuse barriers in heterogeneous landscapes. The solution is likely to involve not only compositional difference between regions as a surrogate for geographically distinct evolutionary histories [70], but also approaches such as least-cost paths [75] or resistance surfaces [76] that represent the 'cost' of connectivity between areas as a function of their current habitat suitability. Such approaches have been applied in models of beta diversity at landscape and regional scales [77,78], including in GDM [79]. Moving beyond cost surfaces based on current habitat suitability, we can imagine a framework for spatially continuous models of isolation, in effect hypotheses about the location and relative degree of 'barrierness' or long-term isolation contributed by locations across a study region. Such an

approach may help to better distinguish components of turnover due to niche differentiation from the geographic signal of historical diversification, for both quantitative explanations and predictive models of turnover.

Concluding remarks

The past decade has seen a rapid proliferation of methods and data sources that allow unprecedented insight into questions at the interface of evolution and ecology. However, many of these analytical methods are regularly applied to processes that are both spatially and temporally autocorrelated, without sufficient consideration being given to those sources of autocorrelation. In some cases, we may be inferring nonexistent ecological processes through spurious correlations between the spatial pattern of diversification and the spatial autocorrelation present in the environment. Although new methods have been developed that address some of these issues, adoption of these approaches has been slow in some cases. Until ecological biogeography is founded on analytical methods that can tease apart the geographic signals of evolutionary and ecological processes, both ecologists and evolutionary biologists should be careful when inferring ecological drivers of biogeographic patterns. At a minimum, any interpretation of analytical results must take into account the possibility that strictly geographic historical processes could have produced the patterns being inferred.

In this opinion article, we have examined some of the most important ways in which geographic and ecological processes can be conflated with commonly used tests in the current literature, and the ways in which investigators have begun to deal with these issues. For community phylogenetics, macroecology, and biogeography, researchers should explicitly include both kinds of process (speciation history and ecological sorting) as potential explanatory variables in analyses of species coexistence and assemblage structure.

References

- 1 Elton, C. (1946) Competition and the structure of ecological communities. *J. Anim. Ecol.* 15, 54–68
- 2 Webb, C.O. *et al.* (2002) Phylogenies and community ecology. *Annu. Rev. Ecol. Syst.* 33, 475–505
- 3 Barraclough, T.G. and Vogler, A.P. (2000) Detecting the geographical pattern of speciation from species-level phylogenies. *Am. Nat.* 155, 419–434
- 4 Berlocher, S.H. (1998) Can sympatric speciation via host or habitat shift be proven from phylogenetic and biogeographic evidence. In *Endless Forms: Species and Speciation*, pp. 99–113, Oxford University Press
- 5 Petren, K. and Case, T.J. (1998) Habitat structure determines competition intensity and invasion success in gecko lizards. *Proc. Natl. Acad. Sci. U.S.A.* 95, 11739–11744
- 6 Theoharides, K.A. and Dukes, J.S. (2007) Plant invasion across space and time: factors affecting nonindigenous species success during four stages of invasion. *New Phytol.* 176, 256–273
- 7 Bolnick, D.I. and Fitzpatrick, B.M. (2007) Sympatric speciation: models and empirical evidence. *Annu. Rev. Ecol. Syst.* 38, 459–487
- 8 Anacker, B.L. and Strauss, S.Y. (2014) The geography and ecology of plant speciation: range overlap and niche divergence in sister species. *Proc. R. Soc. B: Biol. Sci.* 281, 20132980
- 9 Grossenbacher, D.L. *et al.* (2014) Niche and range size patterns suggest that speciation begins in small, ecologically diverged populations in North American monkeyflowers (*Mimulus* spp.). *Evolution* 68, 1270–1280

- 10 Fitzpatrick, B. *et al.* (2009) Pattern, process and geographic modes of speciation. *J. Evol. Biol.* 22, 2342–2347
- 11 Phillips, B.L. *et al.* (2006) Invasion and the evolution of speed in toads. *Nature* 439, 803
- 12 Hewitt, G. (2000) The genetic legacy of the Quaternary ice ages. *Nature* 405, 907–913
- 13 Webb, C.O. (2000) Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *Am. Nat.* 156, 145–155
- 14 Cavender-Bares, J. *et al.* (2006) Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. *Ecology* 87, S109–S122
- 15 Cavender-Bares, J. *et al.* (2004) Phylogenetic overdispersion in Floridian oak communities. *Am. Nat.* 163, 823–843
- 16 Kembel, S.W. and Hubbell, S.P. (2006) The phylogenetic structure of a neotropical forest tree community. *Ecology* 87, S86–S99
- 17 Lovette, I.J. and Hochachka, W.M. (2006) Simultaneous effects of phylogenetic niche conservatism and competition on avian community structure. *Ecology* 87, S14–S28
- 18 Slingsby, J.A. and Verboom, G.A. (2006) Phylogenetic relatedness limits co-occurrence at fine spatial scales: evidence from the *Schoenoid* sedges (*Cyperaceae: Schoeneae*) of the Cape Floristic Region, South Africa. *Am. Nat.* 168, 14–27
- 19 Lynch, J.D. (1989) The gauge of speciation: on the frequencies of modes of speciation. In *Speciation and its Consequences* (Otte, D. and Endler, J.A., eds), pp. 527–556. Sinauer Associates
- 20 Fitzpatrick, B.M. and Turelli, M. (2006) The geography of mammalian speciation: mixed signals from phylogenies and range maps. *Evolution* 60, 601–615
- 21 Phillimore, A.B. *et al.* (2008) Sympatric speciation in birds is rare: insights from range data and simulations. *Am. Nat.* 171, 646–657
- 22 Shmida, A. and Ellner, S. (1985) Coexistence of plant species with similar niches. In *Plant Community Ecology: Papers in Honor of Robert H. Whittaker* (Peet, R.K., ed.), pp. 275–301. Springer
- 23 Brown, J.H. (1995) *Macroecology*, University of Chicago Press
- 24 Merwin, L. *et al.* (2012) Phylogenetic and phenotypic structure among Banksia communities in south-western Australia. *J. Biogeogr.* 39, 397–407
- 25 Cardillo, M. and Pratt, R. (2013) Evolution of a hotspot genus: geographic variation in speciation and extinction rates in *Banksia* (*Proteaceae*). *BMC Evol. Biol.* 13, 155
- 26 Cardillo, M. (2012) The phylogenetic signal of species co-occurrence in high-diversity shrublands: different patterns for fire-killed and fire-resistant species. *BMC Ecol.* 12, 21
- 27 Anderson, M.J. *et al.* (2011) Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. *Ecol. Lett.* 14, 19–28
- 28 Sandro, B. (2008) Introduction of the American grey squirrel (*Sciurus carolinensis*) in Europe: a case study in biological invasion. *Curr. Sci.* (00113891) 95, 903–906
- 29 Morales, C.L. *et al.* (2013) Rapid ecological replacement of a native bumble bee by invasive species. *Front. Ecol. Environ.* 11, 529–534
- 30 Gotelli, N.J. *et al.* (2010) Macroecological signals of species interactions in the Danish avifauna. *Proc. Natl. Acad. Sci. U.S.A.* 107, 5030–5035
- 31 Cardillo, M. (2011) Phylogenetic structure of mammal assemblages at large geographical scales: linking phylogenetic community ecology with macroecology. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 366, 2545–2553
- 32 Cardillo, M. *et al.* (2008) Global patterns in the phylogenetic structure of island mammal assemblages. *Proc. R. Soc. B: Biol. Sci.* 275, 1549–1556
- 33 Harrison, S. and Cornell, H. (2008) Toward a better understanding of the regional causes of local community richness. *Ecol. Lett.* 11, 969–979
- 34 Ricklefs, R.E. (2008) Disintegration of the ecological community. *Am. Nat.* 172, 741–750
- 35 Broennimann, O. *et al.* (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecol. Biogeogr.* 21, 481–497
- 36 Glor, R.E. and Warren, D. (2011) Testing ecological explanations for biogeographic boundaries. *Evolution* 65, 673–683
- 37 Graham, C.H. *et al.* (2004) Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. *Evolution* 58, 1781–1793
- 38 Peterson, A. *et al.* (1999) Conservatism of ecological niches in evolutionary time. *Science* 285, 1265–1267
- 39 Warren, D.L. *et al.* (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution* 62, 2868–2883
- 40 Godsoe, W. (2010) I can't define the niche but I know it when I see it: a formal link between statistical theory and the ecological niche. *Oikos* 119, 53–60
- 41 McNerny, G.S. and Etienne, R.S. (2013) 'Niche' or 'distribution' modelling? A response to Warren. *Trends Ecol. Evol.* 28, 13–14
- 42 Peterson, A.T. *et al.* (2011) *Ecological Niches and Geographic Distributions (MPB-49)*, Princeton University Press
- 43 Soberon, J. and Peterson, A.T. (2005) Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodivers. Inform.* 2, 1–10
- 44 Warren, D.L. (2012) In defense of 'niche modeling'. *Trends Ecol. Evol.* 27, 497–500
- 45 Warren, D.L. (2013) "Niche modeling": that unpleasant sensation means it's working. *Trends Ecol. Evol.* 28, 193–194
- 46 Godsoe, W. (2010) Regional variation exaggerates ecological divergence in niche models. *Syst. Biol.* 59, 298–306
- 47 Knouft, J.H. *et al.* (2006) Phylogenetic analysis of the evolution of the niche in lizards of the *Anolis sagrei* group. *Ecology* 87, S29–S38
- 48 Rice, N.H. *et al.* (2003) Ecological niche differentiation in the *Aphelocoma* jays: a phylogenetic perspective. *Biol. J. Linn. Soc.* 80, 369–383
- 49 Peterson, A.T. (2011) Ecological niche conservatism: a time-structured review of evidence. *J. Biogeogr.* 38, 817–827
- 50 Losos, J.B. (2011) Seeing the forest for the trees: the limitations of phylogenies in comparative biology. *Am. Nat.* 177, 709–727
- 51 Freckleton, R.P. and Jetz, W. (2009) Space versus phylogeny: disentangling phylogenetic and spatial signals in comparative data. *Proc. R. Soc. B: Biol. Sci.* 276, 21–30
- 52 Kellermann, V. *et al.* (2009) Fundamental evolutionary limits in ecological traits drive *Drosophila* species distributions. *Science* 325, 1244–1246
- 53 Diniz-Filho, J.A.F. *et al.* (2003) Spatial autocorrelation and red herrings in geographical ecology. *Global Ecol. Biogeogr.* 12, 53–64
- 54 Legendre, P. (1993) Spatial autocorrelation: trouble or new paradigm? *Ecology* 74, 1659–1673
- 55 Hua, X. and Wiens, J.J. (2013) How does climate influence speciation. *Am. Nat.* 182, 1–12
- 56 Kozak, K.H. and Wiens, J.J. (2010) Accelerated rates of climatic-niche evolution underlie rapid species diversification. *Ecol. Lett.* 13, 1378–1389
- 57 Warren, D.L. *et al.* (2010) ENMTTools: a toolbox for comparative studies of environmental niche models. *Ecography* 33, 607–611
- 58 Nekola, J.C. and White, P.S. (1999) The distance decay of similarity in biogeography and ecology. *J. Biogeogr.* 26, 867–878
- 59 Ferrier, S. *et al.* (2007) Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Divers. Distrib.* 13, 252–264
- 60 Tuomisto, H. (2010) A diversity of beta diversities: straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography* 33, 2–22
- 61 Legendre, P. *et al.* (2005) Analyzing beta diversity: partitioning the spatial variation of community composition data. *Ecol. Monogr.* 75, 435–450
- 62 Kraft, N.J. *et al.* (2011) Disentangling the drivers of β diversity along latitudinal and elevational gradients. *Science* 333, 1755–1758
- 63 Bell, K.L. *et al.* (2013) The role of geography and environment in species turnover: phytophagous arthropods on a Neotropical legume. *J. Biogeogr.* 40, 1755–1766
- 64 Hubbell, S.P. (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*, Princeton University Press
- 65 Soinenen, J. *et al.* (2007) The distance decay of similarity in ecological communities. *Ecography* 30, 3–12
- 66 Zhang, J.L. *et al.* (2013) Phylogenetic beta diversity in tropical forests: implications for the roles of geographical and environmental distance. *J. Syst. Evol.* 51, 71–85
- 67 Stegen, J.C. and Hurlbert, A.H. (2011) Inferring ecological processes from taxonomic, phylogenetic and functional trait β -diversity. *PLoS ONE* 6, e20906
- 68 Graham, C.H. and Fine, P.V. (2008) Phylogenetic beta diversity: linking ecological and evolutionary processes across space in time. *Ecol. Lett.* 11, 1265–1277

- 69 Swenson, N.G. (2011) Phylogenetic beta diversity metrics, trait evolution and inferring the functional beta diversity of communities. *PLoS ONE* 6, e21264
- 70 Rosauer, D.F. *et al.* (2014) Phylogenetic generalised dissimilarity modelling: a new approach to analysing and predicting spatial turnover in the phylogenetic composition of communities. *Ecography* 37, 21–32
- 71 Fitzpatrick, M.C. *et al.* (2013) Environmental and historical imprints on beta diversity: insights from variation in rates of species turnover along gradients. *Proc. R. Soc. B: Biol. Sci.* 280, 20131201
- 72 Ellis, N. *et al.* (2011) Gradient forests: calculating importance gradients on physical predictors. *Ecology* 93, 156–168
- 73 Burley, H.M. *et al.* (2012) Spatial non-stationarity and anisotropy of compositional turnover in eastern Australian Myrtaceae species. *Int. J. Geogr. Inf. Sci.* 26, 2065–2081
- 74 Graham, C.H. *et al.* (2009) Phylogenetic structure in tropical hummingbird communities. *Proc. Natl. Acad. Sci. U.S.A.* 106, 19673–19678
- 75 Ray, N. (2005) Pathmatrix: a geographical information system tool to compute effective distances among samples. *Mol. Ecol. Notes* 5, 177–180
- 76 McRae, B.H. *et al.* (2008) Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology* 89, 2712–2724
- 77 Jiménez-Valverde, A. *et al.* (2010) Climate and regional beta-diversity gradients in spiders: dispersal capacity has nothing to say? *Insect Conserv. Divers.* 3, 51–60
- 78 Weinstein, B.G. *et al.* (2014) Taxonomic, phylogenetic, and trait beta diversity in South American hummingbirds. *Am. Nat.* 184, 211–224
- 79 Thomassen, H.A. *et al.* (2010) Modeling environmentally associated morphological and genetic variation in a rainforest bird, and its application to conservation prioritization. *Evol. Appl.* 3, 1–16
- 80 Chesser, R.T. and Zink, R.M. (1994) Modes of speciation in birds: a test of Lynch's method. *Evolution* 48, 490–497
- 81 Barraclough, T.G. *et al.* (1998) Revealing the factors that promote speciation. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 353, 241–249