Environmental and historical imprints on beta diversity: insights from variation in rates of species turnover along gradients

Matthew C. Fitzpatrick¹, Nathan J. Sanders², Signe Normand³, Jens-Christian Svenning⁴, Simon Ferrier⁵, Aaron D. Gove⁶ and Robert R. Dunn⁷.⁸

¹Appalachian Laboratory, University of Maryland Centre for Environmental Science, Frostburg, MD 21532, USA
²Department of Ecology and Evolutionary Biology, University of Tennessee, 569 Dabney Hall, Knoxville, TN 37996-1610, USA
³Landscape Dynamics, Swiss Federal Research Institute WSL, Birmensdorf, Switzerland
⁴Ecoinformatics and Biodiversity Group, Department of Bioscience, Aarhus University, Aarhus, Denmark
⁵CSIRO Ecosystem Sciences, Canberra, Australian Capital Territory, Australia
⁶Department of Environment and Agriculture, Curtin University, Perth, Western Australia, Australia
⁷Department of Biology, and ⁸Keck Center for Behavioral Biology, North Carolina State University, Raleigh, NC 27695-7617, USA

A common approach for analysing geographical variation in biodiversity involves using linear models to determine the rate at which species similarity declines with geographical or environmental distance and comparing this rate among regions, taxa or communities. Implicit in this approach are weakly justified assumptions that the rate of species turnover remains constant along gradients and that this rate can therefore serve as a means to compare ecological systems. We use generalized dissimilarity modelling, a novel method that accommodates variation in rates of species turnover along gradients and between different gradients, to compare environmental and spatial controls on the floras of two regions with contrasting evolutionary and climatic histories: southwest Australia and northern Europe. We find stronger signals of climate history in the northern European flora and demonstrate that variation in rates of species turnover is persistent across regions, taxa and different gradients. Such variation may represent an important but often overlooked component of biodiversity that complicates comparisons of distance-decay relationships and underscores the importance of using methods that accommodate the curvilinear relationships expected when modelling beta diversity. Determining how rates of species turnover vary along and between gradients is relevant to understanding the sensitivity of ecological systems to environmental change.

1. Introduction

Understanding the role of environmental and spatial processes in determining geographical variation in biodiversity remains a fundamental pursuit of macroecology [1]. While many studies have considered why some places have more species than others, the past decade has witnessed a renewed emphasis on understanding spatial variation in species composition, or beta diversity [2–6]. A better understanding of the environmental and geographical factors that underlie patterns of beta diversity and how these relationships vary between regions with contrasting ecological and evolutionary histories may provide insights into the processes structuring ecological communities [7].

A common approach for comparing patterns of beta diversity for different regions or taxa is to compare the rates at which similarity in species composition decreases (or conversely, the rates at which dissimilarity increases) as a function of geographical and/or environmental distance [4], with greater rates indicating more rapid turnover of species in space [6]. Distance-decay relationships arise from multiple causes, but studies typically emphasize two main drivers [8–10]:

(i) the decrease in environmental similarity with distance (niche-based processes),
species turnover along environmental gradients, and differ-
we use GDM to examine variation in the magnitude and rate of
during the Pleistocene [22,23]. In the light of these differences,
dispersal-limited species that persisted in numerous refugia
southwest Australia has remained unglaciated for millions of
long-distance seed dispersal [20]. By contrast, megadiverse
patterns of beta diversity between the floras of southwest
[16]), a novel, nonlinear statistical approach to compare
will depend on where along the gradient it is measured
measured on essentially arbitrary scales from a biological
between 0 and 1, and because environmental variables are
log-transformations) of similarity versus environmental and/
constant and that beta diversity can therefore be quantified
the slope of a linear regression (typically incorporating log-transformations) of similarity versus environmental and/or geographical distance [2,3,8]. Linear approaches present two problems. First, linear models cannot accommodate the pronounced curvilinear relationships expected because most measures of compositional dissimilarity are constrained between 0 and 1, and because environmental variables are measured on essentially arbitrary scales from a biological perspective (e.g. a difference of 100 mm in annual precipitation is likely to have a much bigger effect on biological composition in a desert than in a rainforest). Second, if compositional similarity is log-transformed, then parameter estimates can be highly sensitive to assemblages that share no species in common because the log of zero is undefined [13].

The consequences of assuming constant rates of turnover remain poorly considered in the literature. However, failure to account for non-stationarity of distance–decay rates may artificially create the appearance that environment explains less of beta diversity than it actually does—possibly in favour of dispersal- or history-related explanations [14]. A second and possibly less well-appreciated consequence is that if species composition changes at different rates along different portions of environmental gradients, no single rate of turnover exists. Instead, the rate of compositional turnover will depend on where along the gradient it is measured [14,15], thereby rendering it of little use as a means to quantify and compare patterns of beta diversity as is often performed.

Here, we use generalized dissimilarity modelling (GDM; [16]), a novel, nonlinear statistical approach to compare patterns of beta diversity between the floras of southwest Australia and northern Europe, with a focus on plants that differ in seed dispersal adaptations. Northern Europe experienced multiple episodes of cooling and glaciation during previous ice ages, the effects of which remain evident both as post-glacial migration lags in plant distributions [17–19] and a greater representation of plants with adaptations for long-distance seed dispersal [20]. By contrast, megadiverse southwest Australia has remained unglaciated for millions of years [21] and contains a high proportion of endemics and dispersal-limited species that persisted in numerous refugia through changes mainly in aridity rather than temperature during the Pleistocene [22,23]. In the light of these differences, we use GDM to examine variation in the magnitude and rate of species turnover along environmental gradients, and differences in the extent to which environment versus space uniquely explains patterns of beta diversity. Specifically, we use GDM to test the following predictions: (i) the diverse, narrowly distributed flora of southwest Australia will exhibit greater turnover than that of northern Europe, though the magnitude and rate of turnover will vary as a function of individual gradients and with position along each gradient, respectively; (ii) precipitation gradients will have greater influence on beta diversity in arid southwest Australia, whereas temperature gradients will have greater influence in northern Europe; (iii) between regions, geographical distance will explain a greater proportion of spatial variation in species composition in northern Europe, where post-glacial migration lags remain, than in more climatically stable southwest Australia; and (iv) within regions, geographical distance will explain a greater proportion of spatial variation in species composition for plants lacking adaptations for long-distance seed dispersal when compared with plants with higher potential for long-distance dispersal.

Lastly, we wished to ascertain whether accommodating nonlinearity improves explanation of observed patterns. Thus, for comparative purposes, we also ask whether we obtain similar inferences using linear models as we do from GDM.

2. Material and methods

(a) Floristic datasets
Northern Europe and southwest Australia exhibit little overlap in floristic composition. In northern Europe, deciduous trees, conifers, herbs and grasses dominate the temperate flora, whereas sclerophyllous scrub vegetation adapted to a Mediterranean climate and low-nutrient soils characterize the flora of southwest Australia [24]. For each region, we compiled datasets on the distribution of plant species and their seed morphology. While not a perfect proxy for dispersal ability [25], morphological adaptations to different primary dispersal vectors have been shown to influence macroecology of plants in ways consistent with perceived dispersal ability [4,8,12,19]. We considered species that could be assigned to one of four dispersal modes: ant, passive, vertebrate or wind. For northern Europe, we used Atlas Florae Europaeae (AFE), which covers approximately 20% of the European flora on a 50 × 50 km grid (AFE cells; [26]). Given data constraints, we focus on Europe within 48° N–71° N, 11° W–32° W. In northern Europe, we obtained data for 785 species, 39 (5%) of which were considered ant-dispersed, 35 (4%) passively dispersed, 279 (36%) vertebrate-dispersed and 432 (55%) wind-dispersed using dispersal trait data obtained from [27] and FloraWeb (http://www.floraweb.de). In southwest Australia, we used seed morphology and publications on seed dispersal to assign dispersal mode at the level of plant genera, except for Acacia, which was assigned as species to accommodate within-genus variation. We obtained data for 2936 species (approx. 50% of the flora) from the Western Australia Herbarium (PERTH; data provided May 2005), of which 973 (33%) were considered ant-dispersed, 1496 (51%) passively dispersed, 128 (4%) vertebrate-dispersed and 339 (12%) wind-dispersed.

We controlled for possible confounding effects of regional differences in scale and extent on patterns of beta diversity [9,10,28] in two ways. First, for southwest Australia, occurrence records were assigned to the centroids of 50 × 50 km cells to match the resolution of the AFE data (figure 1a). Second, to account for the greater extent of northern Europe and the fact that the North and Baltic Seas probably act as dispersal barriers, we performed primary analyses on all of northern Europe (figure 2a) and secondary analyses on two subregions with the same maximum spatial extent as southwest Australia, including a southern subregion not dissected by the sea (see the electronic supplementary material, figures S1a and S2a).
(b) Environmental datasets
We assembled climate and soil datasets to describe environmental gradients. For climatic gradients, we used an uncorrelated subset \((r < 0.75)\) of the 19 bioclimatic variables from WorldClim (www.worldclim.org; [29]). After assessing correlations for each region separately, the same five climate variables were selected for southwest Australia and northern Europe, including minimum temperature, maximum temperature, precipitation seasonality, and summer and winter precipitation. Ideally, datasets of soil properties would be congruent among regions; to date, however, they are not. Moreover, given the coarse resolution of global soil datasets and a high degree of spatial variability in southwest Australia, we used soil layers specific to each region. In northern Europe, uncorrelated soil variables included pH, sand content and calcium carbonate concentration [30]. For southwest Australia, we retained five uncorrelated soil variables from the Australian Natural Resources Data Library (http://data.brs.gov.au/; accessed September 2006), including sand content, plant-available phosphorus, saturated hydraulic conductivity, plant-available water capacity, and soil depth. To match the AFE data, the resolution of the environmental data for southwest Australia was decreased to \(50 \times 50\) km by calculating the mean of surrounding cells using the raster library [31] in R v. 2.12.2 [32].

(c) Statistical modelling
We used GDM and—in supplement—multiple linear regression to compare patterns of beta diversity between the floras of
southwest Australia and northern Europe to evaluate the contribution of environment and space in explaining these patterns. GDM is a nonlinear matrix regression technique for analysing spatial patterns in the compositional dissimilarity (quantified with the Sørensen measure) between pairs of locations as a function of environmental dissimilarity and geographical distance [16]. Unlike classical linear matrix regression, GDM accommodates (i) variation in the rate of compositional turnover (non-stationarity) at different positions along a given gradient, and (ii) the curvilinear relationship between compositional dissimilarity and increasing environmental/geographical distance between sites. To address non-stationarity, GDM first uses maximum-likelihood estimation and flexible I-splines to transform each of the predictor variables and provide the best supported relationship between intersite environmental/geographical separation and compositional dissimilarity [16]. Second, this combined scale of inter-site distances is transformed via a link function to accommodate the curvilinear relationship between compositional dissimilarity (constrained between 0 and 1) and environmental/geographical separation (see [16] for details). We used the default of three I-spline basis functions per predictor. When plotted, the maximum height of each I-spline represents the total amount of compositional turnover associated with that variable, holding all other variables constant. As such, the I-splines are partial regression fits that serve as an indication of the importance of each variable in determining patterns of beta diversity. Second, the slope of the I-spline indicates the rate of species turnover and, importantly, how this rate varies at any point along the gradient concerned (holding all other variables constant). Lastly, the difference in height between any two sites along the I-spline corresponds to the modelled contribution of that predictor variable to the total ecological distance between those sites.

To fit GDMs, we constructed site-by-species and site-by-environment matrices for each flora and dispersal mode, where sites are 50 x 50 km cells. In addition to climate and soil predictors, the environment matrices included geographical coordinates of cell centroids based on equidistant conic projections. It is from these predictors that GDM derives sets of I-splines and calculates distances between all possible pairs of sites. We tested variable significance using Monte Carlo permutation (see [16,33]) and retained only significant variables in final models. The results included, for each flora and dispersal mode therein: (i) a set of significant predictor variables, (ii) a unique fitted I-spline for each significant predictor variable describing the relationship between beta diversity and that gradient, and (iii) per cent deviance explained by the model, the metric used by GDM to assess model fit. We plotted the I-splines to assess how magnitudes and rates of species turnover varied along and between gradients and how these patterns differed between regions and dispersal modes. To quantify the magnitude of turnover along each gradient and the relative importance of that gradient in driving species turnover, we summed the coefficients of the I-splines (each spline has three coefficients), which is equivalent to the maximum height obtained by the curve (see [16]). To evaluate the unique contributions of environment and space in explaining species turnover of each flora and dispersal mode, we partitioned the deviance resulting from sets of three GDMs that used either environmental variables, geographical distance or both as predictor variables [34]. All
analyses were performed in R v. 2.12.2 [32] using GDM functions available from http://www.biomaps.net.au/gdm/.

To evaluate how much improvement in explanation is being provided by GDM relative to a linear model, we also applied the more standard approach of using multiple linear regression to relate the natural logarithm of compositional similarity ($s$) to environmental and geographical distance [35], emphasizing analyses of the entire floras of southwest Australia and all of northern Europe. We used the log($s + 0.01$) transformation to accommodate site pairs that share no species in common ($s = 0$, the log of which is undefined). To assess variable contributions, we used the hier.part R library [36]. Data and R scripts are available from Dryad (http://dx.doi.org/10.5061/dryad.b1P60).

### 3. Results

(a) Patterns of species turnover

Southwest Australia exhibited greater floristic turnover than northern Europe (figures 1g and 2f). However, patterns of species turnover in all regions varied by environmental gradient and geographical distance and by seed dispersal mode. In southwest Australia, winter precipitation was the most important gradient for determining turnover for the entire flora and for ant- and passive-dispersed plants, whereas maximum temperature and geographical distance were most important for vertebrate- and wind-dispersed plants, respectively (bold text in table 1). These three predictors along with total phosphorus consistently were found by GDM to be significant predictors for all groups of flora. Notably, winter precipitation, the most biologically important climatic gradient in southwest Australia, both at present and historically [23,24], was consistently one of the top two most important predictors of beta diversity. In nearly all instances, the fitted functions describing the rates and magnitude of turnover along each gradient were nonlinear, with rates of turnover varying with position along gradients and being greatest at low levels of winter precipitation and soil phosphorus (figure 1b–e). Predictors identified as most important by linear models differed from those of GDM and included geographical distance, sand content and soil depth, while largely excluding precipitation variables (see the electronic supplementary material, figure S3a).

In contrast to the importance of precipitation gradients in southwest Australia, GDM consistently identified geographical distance and temperature gradients as the most important drivers of beta diversity in northern Europe (bold text in table 2), though their importance varied by the region considered and by dispersal mode. For all of northern Europe (table 2) and the southern subregion (see the electronic supplementary material, table S2), geographical distance was the most important driver of floristic turnover, whereas temperature was most important in the circular subregion (see the electronic supplementary material, table S1). Soil pH was also commonly a significant predictor of turnover, though mainly for the subregions. The fitted functions for northern European plants also were often nonlinear (figures 2b–d and the electronic supplementary material, S1b–e and S2b–c), though less so than those for southwest Australia. As with GDM, linear models also identified geographical distance and temperature as the most important predictors of turnover in northern Europe (see the electronic supplementary material, figure S3b).

(b) Contributions of environment and space to beta diversity

GDMs including both environmental and geographical distance outperformed models with these predictors in isolation
and explained more than 80% of the deviance in observed floristic dissimilarities (southwest Australia 81.1%; northern Europe 82.7%). Except for ant-dispersed plants in northern Europe, GDM consistently explained more variation in species turnover than did linear models (electronic supplementary material, figure S4). The flora of southwest Australia exhibited the greatest differences between deviance explained by GDM and linear models, except for passive-dispersed plants.

When the deviance explained from GDM was partitioned into unique and shared components of environment and space, the unique contribution of space was, as predicted, always greater for the entire flora of northern Europe (figure 2c), regardless of the region considered (see the electronic supplementary material, figures S1f and S2f), than for southwest Australia (figure 1f). By contrast, linear models increased the unique contribution of space in southwest Australia and decreased it in all regions of northern Europe relative to GDM. As a result, the unique contribution of space from linear models was less in all of northern Europe than in southwest Australia (7.2% versus 8.8%, respectively), but was greater than southwest Australia in the circular (11.7%) and southern (22.9%) subregions of northern Europe.

For dispersal modes individually, the unique contribution of space from GDM was greater in northern Europe than in southwest Australia, with the exception of vertebrate-dispersed plants in southwest Australia. For the entire floras, the unique contribution of environment from GDM was always greater in southwest Australia than any region of northern Europe, albeit weakly when compared with the circular subregion (see the electronic supplementary material, figure S1). In contrast to our prediction, and for both GDM and linear models, space did not uniquely explain a greater proportion of total deviance for plants lacking adaptations for long-distance dispersal (i.e. ant- and passively dispersed) than for wind- and vertebrate-dispersed plants, in either southwest Australia or for any region of northern Europe.

4. Discussion

We examined relationships among species traits and suites of environmental predictors to compare patterns of beta diversity in two distinct floras with contrasting climatic histories. Consistent with northern Europe’s glacial history, we found using GDM that space had a greater contribution in explaining beta diversity, where the proportion of total explained deviance attributable purely to space was between 44% and 269% greater than in southwest Australia (depending on the region of northern Europe considered). In addition, we found that geographical distance and temperature gradients were most important for driving species turnover in northern Europe. By contrast, the contribution of space tended to be less, and the contribution of environment greater, in explaining beta diversity in the comparatively stable and semi-arid climate of southwest Australia, where the gradient of winter precipitation was among the most important drivers of species turnover. While we consider differences in the nature of climatic change during Quaternary glacial periods the most likely explanation for the differences in the role of space in explaining beta diversity, we cannot rule out other possibilities, for example, failure to include important environmental variables in one or both regions. The dispersal barriers posed by the North and Baltic seas and the absence of similar barriers in southwest Australia cannot explain these regional differences given that space tended to have the greatest contribution in the subregion of northern Europe that did not include these barriers (see the electronic supplementary material, figure S2f) and least in the regions that did (figure 2c and the electronic supplementary material, S1f).

Our results suggesting historic effects are an important factor determining macroecological patterns in northern Europe is consistent with numerous other studies [18,19,37]. However, our study is unique in comparing the magnitude of these effects with other regions. Although both regions have been impacted by previous glacial cycles to some extent, we expected spatial signals to be stronger in northern Europe, where the rate, magnitude and extent of climate change are considered to be greater [20,24,38], and where the strong north to south temperature gradient would tend to structure diversity patterns most prominently with latitude. By contrast, changes in aridity in southwest Australia allowed persistence of species in multiple refugia with suitable microclimates [22]. Given the triangular geometry of southwest Australia and the rapid decrease in precipitation moving inland, these refugia would have been most common along both coastlines, thereby allowing recolonization from multiple fronts and resulting in less spatial structuring of diversity patterns beyond that attributable to environmental gradients alone. Thus, all else being equal, in northern Europe species would have had to migrate greater distances, requiring more time, to attain equilibrium with climate.

In contrast to our prediction, we found little evidence that seed dispersal morphology exhibited a consistent relationship with the extent to which environment versus space explained patterns of beta diversity in either region. The simplest explanation is that seed morphology is a poor predictor of dispersal ability, perhaps because long-distance movements by non-standard means might yield dispersal events exceeding those implied by morphology [25]. For example, emu (Dromaius novaehollandiae) and kangaroo (Macropus fuliginosus) have been shown to disperse ant-dispersed plants in southwest Australia [39]. Although this explanation is satisfying for dispersal-limited species, it is not clear why non-standard mechanisms would trump those of species with morphological adaptations for long-distance movements via their standard means of dispersal.

Using GDM, which accommodates the curvilinear relationships expected when modeling beta diversity, we demonstrated that rates of species turnover vary substantially as a function of the environmental gradient considered and with position along gradients (e.g. the dry versus the wet end of a precipitation gradient). While our study is not the first to document that rates of turnover can vary along environmental gradients [9,15,33,40], it does show that such non-stationarity can be pervasive across regions, taxa and environmental gradients. Linear models, which cannot accommodate non-stationarity, explained less variation in species turnover than GDM, especially in southwest Australia where species turnover is rapid and non-stationarity is pervasive. Also important is that turnover in northern Europe is not sufficiently high to result in dissimilarities of one (complete turnover between sites), whereas in southwest Australia this is common (compare figures 1g and 2f). GDM’s nonlinear link function provides a means to fit this pattern and, unlike a linear model, can do so without the complication of log-transforming sites that share no species in common [13].
These issues are less important in northern Europe, where a straight line nearly suffices. As a result, GDM and linear models identified similar sets of important environmental variables in northern Europe, but different sets in southwest Australia. It is important to note that the variables identified as most important by GDM (e.g. winter precipitation, soil nutrients) in southwest Australia are more consistent with the ecology and biogeography of the region than those identified by linear models [23,24]. Finally, when the deviance explained using linear models was partitioned, the unique contribution of space became greater in southwest Australia and lower in northern Europe than the values obtained from GDM. For all of northern Europe, this change effectively reversed the inference regarding the relative effects of history on the floras of southwest Australia and northern Europe. While the linear assumption may be more tenable for regions such as northern Europe that exhibit lower levels of beta diversity, and where most sites share some species in common, our results confirm the importance of using a method that can accommodate curvilinear relationships expected when modeling compositional dissimilarity. Taken together, our findings suggest that: (i) variation in the rate of species turnover could represent an important and apparently underappreciated component of beta diversity, and that (ii) comparisons of patterns of beta diversity from different ecological systems that assume constant distance-decay rates should be interpreted with caution (or avoided entirely) as conclusions could depend on the gradient considered and its extent (see also [28]).

The curvilinear functions produced by GDM depict how patterns of beta diversity vary by environmental and geographical predictors simultaneously, while holding all other variables constant, and provide a measure of the importance of each variable in driving species turnover. Instead of relying solely on distance-decay rates as a means to quantify and compare beta diversity, these functions offer a complementary or alternate means to quantify geographical variation in species composition, while providing insights not readily apparent or obtainable from distance-decay relationships. Beyond indicating greater overall turnover in the flora of southwest Australia than in northern Europe, the functions from GDM suggested that the environmental gradients most strongly associated with beta diversity are also those likely to have influenced vegetation during previous climatic changes (i.e. precipitation in southwest Australia and temperature in northern Europe). By extension, changes in these gradients in the future might be expected to result in the most dramatic biotic responses and particularly if climatic changes were to occur along those portions of environmental gradients where turnover is most rapid. In this sense, determining which environmental gradients are most closely associated with patterns of beta diversity and, perhaps more importantly, how rates of turnover vary along these gradients, is not only of theoretical interest, it could also serve as a means to quantify the sensitivity of different ecological systems to future climatic change [41].

Funding statement. M.C.F. was supported by US NSF award DEB-1257164, NOAA MDSG NA10OAR4170072 SA7528114-DDD, and by funding from UMCES. J.C.S. was supported by the European Research Council (ERC-2012-StG-310886-HISTFUNC), the Danish Council for Independent Research Natural Sciences (12-125079) and Aarhus University and Aarhus University Research Foundation under the AU IDEAS programme (via the Center for Informatics Research on Complexity in Ecology). S.N. was supported by the Danish Council for Independent Research Natural Sciences (10-080506). N.J.S. and R.R.D. were supported by DOE-FER DE-FG02-08ER64510. R.R.D. was supported by NASA award NNX09AK22G and NSF-CAREER 0953390.

References


