RESEARCH PAPER

Can stochastic geographical evolution re-create macroecological richness–environment correlations?

J. Sebastián Tello[1,2]* and Richard D. Stevens[1]

[1]*Department of Biological Sciences, Louisiana State University, LA 70803, USA,* [2]*Museo de Zoología, Escuela de Ciencias Biológicas, Pontificia Universidad Católica del Ecuador, Avenue 12 de Octubre 1076 y Roca, Apdo. 17-01-2184, Quito, Ecuador*

ABSTRACT

Aim Richness gradients are frequently correlated with environmental characteristics at broad geographic scales. In particular, richness is often associated with energy and climate, while environmental heterogeneity is rarely its best correlate. These correlations have been interpreted as evidence in favour of environmental determinants of diversity gradients, particularly energy and climate. This interpretation assumes that the expected-by-random correlation between richness and environment is zero, and that this is equally true for all environmental characteristics. However, these expectations might be unrealistic. We investigated to what degree basic evolutionary/biogeographical processes occurring independently of environment could lead to richness gradients that correlate with environmental characteristics by chance alone.

Location Africa, Australia, Eurasia and the New World.

Methods We produced artificial richness gradients based on a stochastic simulation model of geographic diversification of clades. In these simulations, species speciate, go extinct and expand or shift their distributions independently of any environmental characteristic. One thousand two hundred repetitions of this model were run, and the resulting stochastic richness gradients were regressed against real-world environmental variables. Stochastic species–environment relationships were then compared among continents and among three environmental characteristics: energy, environmental heterogeneity and climate seasonality.

Results Simulations suggested that a significant degree of correlation between richness gradients and environment is expected even when clades diversify and species distribute stochastically. These correlations vary considerably in strength; but in the best cases, environment can spuriously account for almost 80% of variation in stochastic richness. Additionally, expected-by-chance relationships were different among continents and environmental characteristics, producing stronger spurious relationships with energy and climate than with heterogeneity.

Main conclusions We conclude that some features of empirical species–environment relationships can be reproduced just by chance when taking into account evolutionary/biogeographical processes underlying the construction of species richness gradients. Future tests of environmental effects on richness should consider structure in richness–environment correlations that can be produced by simple evolutionary null models. Research should move away from the naive non-biological null hypotheses that are implicit in traditional statistical tests.

Keywords

Diversification, diversity, environmental gradient, null model, richness, species–environment relationship.

*Correspondence: J. Sebastián Tello, Department of Biological Sciences, Louisiana State University, LA 70803, USA. E-mail: jtello1@lsu.edu

INTRODUCTION

Many hypotheses have been proposed to account for large-scale patterns of diversity and their most common representation: the latitudinal gradient of species richness (Rohde, 1992; Willig *et al.*, 2003). Among these proposed mechanisms, one idea that has received much attention and support is that diversity gradients are a consequence of underlying gradients in environmental characteristics. A major source of evidence in favour of this hypothesis is the frequent and strong correlations between species richness and environmental variables reported for many groups of organisms (Wright *et al.*, 1993; Hawkins *et al.*, 2003; Field *et al.*, 2009; Tello & Stevens, 2010).

It has also been recognized, however, that not all environmental characteristics have identical effects, and some might be more important than others (Currie, 1991; Field *et al.*, 2009). Indeed, various environmental characteristics have been shown to correlate differently with richness gradients. In the most recent review, Field *et al.* (2009) demonstrated that climate and energy are frequently the best predictors of richness gradients and produce stronger relationships than any other environmental or non-environmental hypothesis. In contrast, measures of environmental heterogeneity are rarely the best explanation for richness gradients. Field *et al.* (2009) and others (e.g. Currie, 1991; Wright *et al.*, 1993; Kalmar & Currie, 2007; Kreft & Jetz, 2007) have used these results to suggest that: (1) environmental gradients are important determinants of species richness, and (2) in particular it is climate and energy, not heterogeneity, that are more likely to drive diversity gradients.

Nevertheless, the exact mechanism behind such consistent relationships between richness and environment is not well understood. For example, many mechanisms have been proposed to explain how climatic and energetic variables determine numbers of species (e.g. Currie *et al.*, 2004, considered three explanations, while Evans *et al.*, 2005, considered nine). It is important to note that: (1) all of these proposed explanations assume a causal relationship between energy/climate and richness (namely diversity gradients result as a direct consequence of these environmental gradients), but (2) none has been strongly supported by evidence (see, for example, Currie *et al.*, 2004). Explanations for climate/energy–richness relationships have been more frequently discussed by macroecologists precisely because studies have indicated that these environmental characteristics are the best correlates of richness at broad scales, but similar conclusions can be reached regarding causes for relationships between environmental heterogeneity and richness.

Taking a step back in thinking about determinants of species richness at broad scales, it is clear that the most proximal mechanisms (i.e. those most closely linked to the production of richness gradients) are evolutionary/biogeographical processes associated with the diversification and distribution of clades. Richness in a particular region is determined by the processes of speciation, extinction and range dynamics, which move species distributions in and out of the region of interest (Bokma *et al.*, 2001; Wiens & Donoghue, 2004; Jablonski *et al.*, 2006; Mittelbach *et al.*, 2007; Arita & Vázquez-Domínguez, 2008; Gotelli

et al., 2009). Additionally, these processes typically occur within a bounded domain of distribution (e.g. continent, sea, large island, etc.; Colwell & Lees, 2000). Thus, mechanistic explanations of diversity gradients at broad scales should explicitly include these basic processes that are responsible for the construction of richness gradients.

These spatial processes and constraints associated with diversification and distribution of clades, by themselves, have the potential to create richness gradients that are spatially structured (Hennig, 1966; Bokma *et al.*, 2001; Stephens & Wiens, 2003; Stevens, 2006; Arita & Vázquez-Domínguez, 2008). This is true even if these processes occur independently of any environmental gradient. This is exemplified by simple simulation models that randomly position species distributions in a constrained domain. Typically, these models produce patterns of variation in richness that decrease monotonically from the centre to the edges of the domain (the mid-domain effect: Willig & Lyons, 1998; Colwell & Lees, 2000). Other recent models that incorporate additional processes can modify this simple pattern (e.g. Davies *et al.*, 2005; Colwell *et al.*, 2009), but still produce variation in richness that is spatially structured. Thus, spatially structured gradients of species richness can be produced as an emerging consequence of processes and constraints associated with the geographical diversification of clades (e.g. speciation, extinction, range shifts), and without a direct influence of the environment.

This spatial structure, which might be a necessary characteristic of richness gradients produced during geographic clade evolution, may have important consequences for understanding species–environment relationships. Primarily, an environmental gradient and an independently produced richness gradient distributed in the same domain might be predisposed to be correlated just by chance (throughout, we use 'chance' to mean noncausal coincidence of two independently generated gradients). However, little attention has been given to what richness–environment correlations should be expected in the absence of any effects of environment on their formation, but when richness gradients are produced during the stochastic diversification of clades and distribution of species. Here, we investigate to what degree features of empirical species–environment relationships can be re-created by stochastic simulation models based on the evolutionary/biogeographical first principles responsible for the construction of richness gradients, but where such first principles are completely independent of environmental variables. In particular, we address the following questions: (1) What is the expected strength of spurious correlation between richness and environmental gradients? (2) Can certain environmental characteristics produce higher correlations than others just by chance? and (3) Are random species–environment relationships different on different continents?

To answer these questions, we produced artificial richness gradients based on a stochastic simulation model of diversification and distribution of clades. In these simulations, species speciate, go extinct and expand or shift their distributions independently of the environment. Thus, the resulting artificial richness gradient is not directly affected by environmental gradients.

One thousand two hundred repetitions of this model were run across four different continents. The resulting stochastic richness gradients were regressed against real-world environmental variables that represent three environmental hypotheses. From these regressions, we estimated the strength of species–environment relationships that would be expected by chance alone (i.e. if environment has no influence on the geographic evolution of species richness). We conclude that some features of empirical species–environment relationships can be reproduced just by chance when taking into account the evolutionary/biogeographical processes underlying the construction of species richness gradients. We propose that future tests of environmental effects on richness should consider structure of richness–environmental correlations that can be produced by simple evolutionary models, and that macroecological research should move away from the use of non-biological non-mechanistic null hypotheses that are implicit in most traditional statistical tests.

METHODS

Simulations

We used a geographically explicit simulation model of diversification of clades to create stochastic species richness gradients. We call this simulation model the diversification and range dynamics model (DRD). An outcome of the DRD model is stochastic species distributions within a geographic domain. These distributions can then be transformed into richness gradients by counting the number of species with overlapping distributions within specified regions of the domain. For our simulations, we used as domains four continental masses: Africa, Australia, Eurasia and the New World. Each domain was divided into cells of 100 by 100 km.

The DRD model takes place in time steps. In the first time step, one cell from throughout the domain is randomly selected as the point of origin for the diversifying clade. The first species in the simulation colonizes this cell. Then, a target range size is selected at random from a pool of range sizes. Starting with the second time step, a number of events take place in the following sequence (more details are provided in Appendices S1 & S2 in Supporting Information).

1. Each species present in the domain can move its distribution. The probability of range movement is identical for all species and constant through time. If a species is selected to move its distribution, then a random direction is selected and the entire distribution moves one cell in that direction. Accordingly, species distributions follow random walks within the domain. A distribution can move partially outside of the domain. In such a case, that part of the distribution is lost and will need to be regained by spread of the range elsewhere within the domain (see below).

2. Each species in the domain that has not reached its target range size spreads its distribution by sending 'dispersers' from each occupied cell to surrounding cells. This spread is purely stochastic, and occurs at every time step until the target range size for that species has been reached.

3. Each species in the domain can give rise to a new species. The probability of speciation is identical for all species and constant through time. Speciation is modelled as a punctuated event. For each speciation event, a cell is selected at random from throughout the distribution of the parental species. This cell represents the point of origin for the distribution of the new species. This is equivalent to a population or individual from the parental distribution speciating and giving rise to a new species. Each time a new species appears, a target range size is randomly selected from a lognormally distributed pool of range sizes.

4. Each species in the continent can go extinct. All species present in the domain are evaluated for survival. If one goes extinct, then it disappears from the domain. Extinction probability is identical among all species, but it can either remain constant or change as a function of diversity. This produces two patterns of clade diversification: exponential and logistic.

After these events have taken place, the simulation moves to the next time step to start another cycle of stochastic range movements, range growths, speciations and extinctions. The simulation stops when the surviving number of species in the clade matches a predetermined number of species plus one. The time between the origin of the last required species and the additional species allows the last species to develop a distribution. The additional species is then eliminated from the output. At this point, a species richness gradient is produced by counting the number of range overlaps in each cell of the domain. Domains remained static throughout the simulation (no changes in shape or size).

The algorithm we have developed is similar to other models that simulate geographic diversification while leaving out environmental effects (Bokma *et al.*, 2001; Davies *et al.*, 2005; Rangel & Diniz-Filho, 2005; Arita & Vázquez-Domínguez, 2008; Connolly, 2009; reviewed by Gotelli *et al.*, 2009), but, it also differs in a number of details from previous algorithms (Appendix S1). More importantly, no previous study has used this type of model to systematically investigate the kind of richness–environment relationship that stochastic diversification can produce, and whether these expected-by-chance relationships change among continents or among environmental characteristics.

Our DRD model was repeated 300 times in each domain, leading to 1200 stochastic richness gradients. For each repetition, we modified simulation parameters (e.g. movement or speciation probabilities) to emulate variability observed in diversification and distribution of real clades. However, parameter variation was not based on real data. Consequently, these simulations produce entirely artificial richness gradients. Table 1 presents the varying parameters in the model and the parameter space covered by our simulations. We found that variation in parameter values usually did not have strong or consistent effects on simulation outcomes, and that most variation in the simulated species–environment relationships was produced by the stochasticity in the model (Appendix S1). Simulations were carried out in R (R Development Core Team,

Table 1 Parameters that varied among simulation runs in our analyses. Before a simulation run started, parameter values were randomly drawn from the range of possible values. For every parameter, all values had the same probability of being selected. Details of how parameters were varied can be found in Appendix S1. Additional analyses found that none of these parameters had a strong and consistent effect on the outcome of our simulations (Appendix S1).

Parameter	Values
Clade diversity	100 to 300 species
Mean proportional range size	5% to 70% of domain size
Place of clade origin	Any cell in domain
Range movement probability	0 to 1
Diversification type	Exponential or logistic
Speciation probability	0.0005 to 0.005
Extinction probability	0 to 75% of speciation probability

2008), using BioHPC of the Computational Biology Service Unit at Cornell University (http://cbsuapps.tc.cornell.edu). Code can be found in Appendix S2.

Environmental predictors

Stochastic species richness gradients were regressed against real-world environmental variables (see below). These variables were chosen to represent three environmental hypotheses frequently used to explain empirical richness at broad geographic extents: energy, environmental heterogeneity and climatic seasonality. Most environmental variables were obtained from WorldClim (Hijmans *et al.*, 2005) with a resolution of 30 arcsec (approximately 1 km^2). Net primary productivity (NPP) data were obtained from Imhoff *et al.* (2004) with a resolution of 0.25° squared. All environmental data had resolutions smaller than the cell size in the domains. This allowed calculation of statistics reflecting the central tendency (average) or spatial variability (standard deviation) of environmental variables within each cell (Beyer, 2004). The average number of raster pixels within a richness map cell was 10951.84 for all environmental predictors except NPP; for NPP, this number was 12.43. Energy was represented in our analyses by cell averages of: (1) mean annual temperature, (2) annual precipitation, and (3) annual NPP. These variables represent forms of or surrogates for both kinetic and chemical potential energy (productivity). Environmental heterogeneity was estimated by within-cell standard deviations of: (4) elevation, (5) mean annual temperature, (6) annual precipitation, and (7) annual NPP. Finally, seasonality was represented by cell averages of: (8) monthly coefficient of variation of precipitation, (9) standard deviation of month-to-month variation in temperature, and (10) monthly range of annual temperature (Appendix S1, Fig. S1).

Characterization of stochastic richness–environment relationships

The 1200 simulated richness gradients were regressed against three sets of predictors, each corresponding to one of the envi-

ronmental hypotheses considered. Ordinary least squares (OLS) were used for these multiple regressions. Adjusted R^2 values (Peres-Neto *et al.*, 2006) of these regressions were used to estimate the strength of species–environment relationships produced by our simulations. Additionally, to describe the direction of the stochastic species–environment relationships produced by our model, we investigated the frequency distribution of coefficients from univariate regressions between simulated species richness and each one of the environmental predictors. However, we focused our analyses on R^2 values because: (1) we were mainly interested in the explanatory power of environmental predictors and not on the direction of relationships, and (2) R^2 values have been the most frequently interpreted statistic in regressions between richness and environmental characteristics (e.g. Field *et al.*, 2009). For each regression, we excluded cells that: (1) did not have any species, (2) did not have information on one or more environmental predictors, (3) represented islands, (4) had less than 75% of their area over continental land, or (5) represented environmental outliers.

Although OLS might not be the most statistically appropriate analysis (Beale *et al.*, 2010), important ideas about richness–environment relationships have been produced by previous studies mostly using this type of regression. The objective of our study is to demonstrate the kinds of species–environment relationships that are expected by random diversification/distribution of clades and whether these relationships resemble those in the literature. Thus, we use the analyses that have been most commonly used in previous studies. However, in Appendix S1, we investigated whether the use of spatial regressions could modify our conclusions, and demonstrate that results are quantitatively different, but lead to the same conclusions reached using OLS analyses.

We described and analysed our stochastic species–environment relationships similarly to how Field *et al.* (2009) described empirical relationships in their recent meta-analysis. First, we calculated the primacy of each hypothesis as the proportion of cases (proportion of simulation runs) for which variables representing a particular hypothesis were the strongest correlates ('best predictors') of stochastic richness. Second, we characterized species–environment relationships for each hypothesis using primary adjusted-R^2 values. Primary adjusted-R^2 is the adjusted-R^2 of the hypothesis that correlated most strongly with richness in a particular simulation run.

Statistical analyses

We logit-transformed adjusted-R^2 values for all statistical tests where they were used as the dependent variable (Fox, 2009). However, for ease of interpretation, all plots have been produced showing untransformed adjusted-R^2s.

Question 1: What is the expected strength of correlation between richness and environmental gradients?

For each hypothesis on each continent, we constructed a frequency distribution of adjusted-R^2 values describing the central

tendency and variability of species–environment relationships produced by our simulations. We calculated 95% confidence intervals (CIs) using the 2.5% and 97.5% quantiles of this distribution as its limits (Chernick, 2007). CIs include the 95% most likely values of richness–environment correlations that could be expected by simple coincidence of environmental gradients and independently produced random richness gradients. Similarly, we constructed density distributions describing the variation in the univariate regression coefficients between richness and each environmental predictor.

Question 2: Can certain environmental characteristics exhibit higher correlations than others just by chance?

We determined whether primacy and primary R^2s were different among hypotheses (Field *et al.*, 2009). To compare primacy among hypotheses, we used a one-sample chi-square test, assuming that primacies should be identical for all hypotheses (33.33% for each). We also compared primary R^2 values using a two-way ANOVA, where primary R^2 was contrasted among hypotheses and continents. We followed this analysis with a post-hoc Tukey test. Significant 'hypothesis' main-effect or 'hypothesis-by-domain' interaction would indicate differences among hypotheses in terms of primary adjusted-R^2.

Questions 3: is strength of richness–environment relationships or the relative importance of environmental hypotheses different among continents?

To test whether hypothesis primacy was different among continents, we constructed a multiway contingency table where frequency of primacy for each hypothesis on each continent was recorded. This table was then analysed using log-linear models (Sokal & Rolf, 1994). We tested whether there was a significant 'hypothesis-by-domain' interaction by comparing a saturated model with a reduced model without this interaction. A significant difference indicates that the simpler model is a poorer fit than the complex model, and that the interaction is necessary to explain the data (Crawley, 2007).

We were also interested in whether primary adjusted-R^2 values were different among continents. We tested this using the same two-way ANOVA and post-hoc Tukey test described above, where primary adjusted-R^2 values were compared among hypotheses and continents.

RESULTS

When richness gradients are produced during the diversification and distribution of clades, moderate to strong species–environment relationships could be expected, even if richness is produced independently of environmental gradients (Figs 1 & 2). Histograms of stochastic species–environment correlations demonstrate that expected correlations are larger than zero, and 95% confidence intervals often span a broad range of R^2
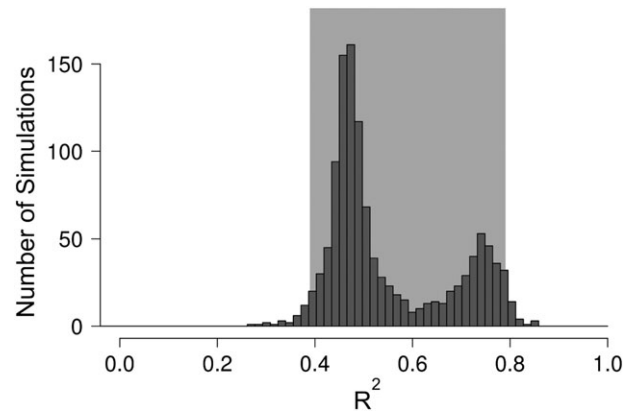

Figure 1 Histogram of species–environment relationships produced by a diversification and range dynamics (DRD) model when all environmental variables are used as predictors of stochastic richness gradients. Histograms are based on all adjusted-R^2s rather than only on the primary adjusted-R^2s. The grey box delimits the central 95% most common values (limits given by 2.5 and 97.5% quantiles of frequency distribution).

values (Figs 1 & 2). Additionally, the distribution of individual regression coefficients demonstrates that species–environment relationships are typically strongly biased away from zero (Fig. 3). The direction of bias and the range of variation, however, changes considerably among predictors and among continents.

Differences in primacy and primary adjusted-R^2s among environmental hypotheses

Stochastic diversification and distribution of clades can produce richness gradients that have patterns of correlation that differ among different environmental characteristics (Figs 2 & 4). Primacy varied significantly among environmental hypotheses ($\chi^2 = 978.8$, $P < 0.001$). Just by chance, variables associated with environmental heterogeneity were less frequently the strongest correlates of richness than variables representing energy or seasonality (Fig. 4a). Between energy and seasonality, climatic seasonality tended to be more frequently the best 'explanation' for stochastic richness gradients (Fig. 4a).

Also, expected-by-chance adjusted-R^2s of the primary predictor varied depending on the environmental hypothesis under consideration (hypothesis main effect: $F = 187.2$, $P < 0.001$; Fig. 4b; Appendix S1, Table S1). Heterogeneity typically could only 'account' for a relatively small fraction of variation in stochastic richness (less than energy: $P = 0.009$; less than seasonality: $P < 0.001$; Fig. 4b; Appendix S1, Table S1). Energy and seasonality, on the other hand, accounted for much larger proportions of variation, typically around 15 to 40%, but in the best cases almost as much as 80% (Figs 2 & 4b). Seasonality produced stronger primary adjusted-R^2s than energy ($P < 0.001$; Fig. 4b; Appendix S1, Table S1).
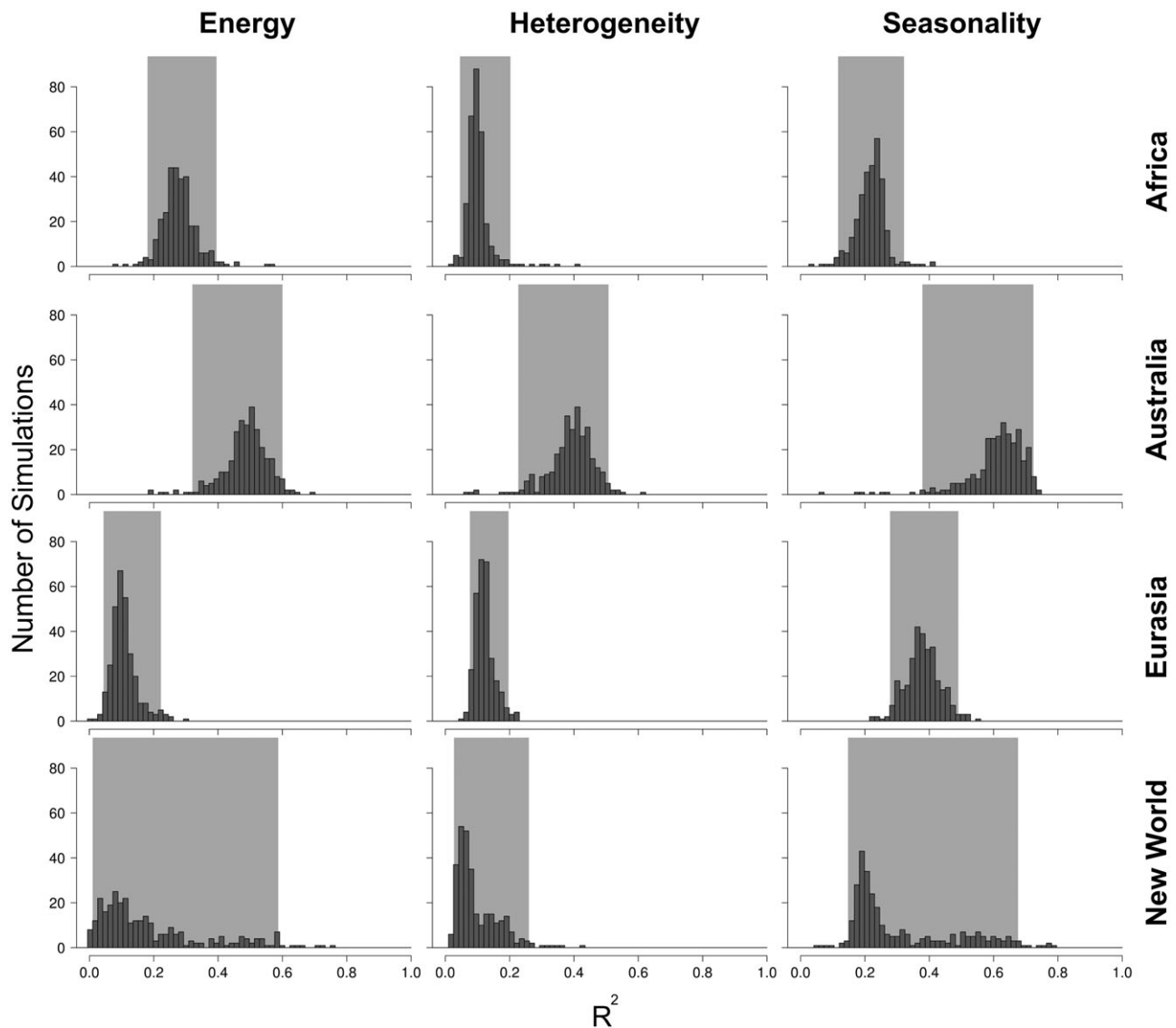
Figure 2 Histograms of species–environment relationships produced by a diversification and range dynamics (DRD) model by environmental hypothesis and continent. Histograms are based on all adjusted-R^2s rather than only on primary adjusted-R^2s. The grey boxes delimit the central 95% most common values (limits given by 2.5 and 97.5% quantiles of frequency distributions).

Differences in relative importance of environmental hypotheses among continents

The log-linear model for frequencies of primacy for each hypothesis demonstrated that no simpler model than the saturated model could successfully explain the data. Removing the hypothesis-by-continent interaction led to a model that was significantly different from the saturated model ($P < 0.001$). This suggests that primacy of different environmental hypotheses changes significantly across continents (Fig. 5). For example, energy has the highest primacy in Africa; but variables associated with seasonality more frequently accounted for the greatest proportion of variation in stochastic richness in all other continents (Fig. 5).

When directly analysing the strength of stochastic species–environment relationships (primary adjusted-R^2s), two-way ANOVA indicated a significant main effect of continent ($F = 458.8$, $P < 0.001$; Fig. 5; Appendix S1, Table S1); for example, for energy and seasonality, species–environment relationships tend to be stronger in Australia than in other continents. (Fig. 5; Appendix S1, Table S1). Also, there was a significant interaction between hypothesis and continent ($F = 7.5$, $P < 0.001$; Fig. 5; Appendix S1, Table S1); for example seasonality is not statistically different from energy in the New World, but has a higher adjusted-R^2 in Australia (Fig. 5; Appendix S1, Table S1). These results demonstrate a clear effect of continent and a potential change in relative importance of multiple hypotheses across domains.

Figure 3 Violin plots presenting probability densities for values of standardized regression coefficients of each environmental predictor on each continent. Regression coefficients are produced by univariate regressions of stochastic richness against each environmental predictor individually. Prec, precipitation; NPP, net primary productivity; Temp, temperature; Elev, elevation; Sea, seasonality.
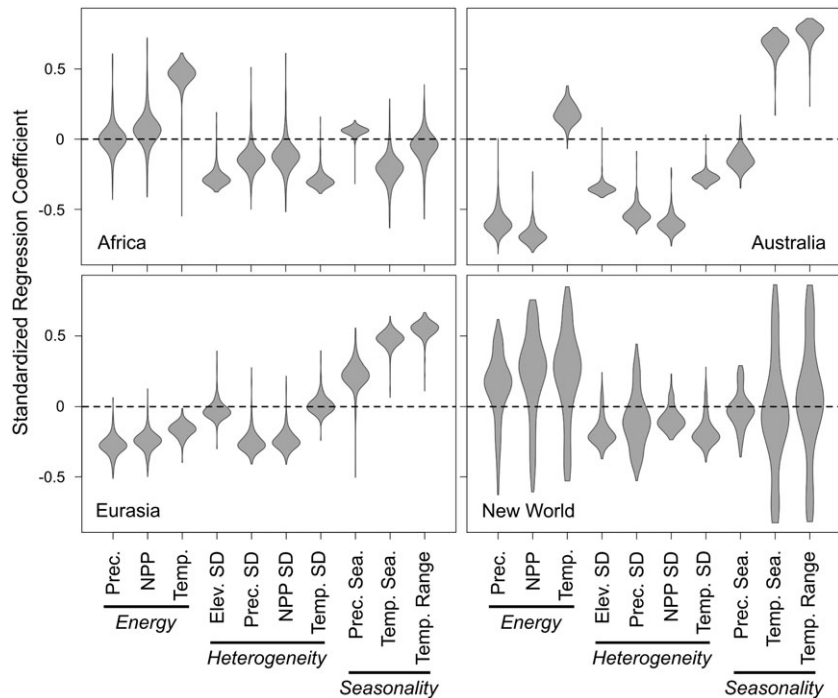

Figure 4 Simulated species–environment relationships produced by a diversification and range dynamics (DRD) model. Species–environment relationships are compared among three environmental hypotheses: energy, heterogeneity and seasonality. Comparisons are based on primacy (a) and primary adjusted-R^2s (b). Primacy is the proportion of times variables representing a particular hypothesis were the best correlates of richness compared with other hypotheses. Primary adjusted-R^2 is the proportion of variation 'explained' by the variables of the primary hypothesis in each simulation run.
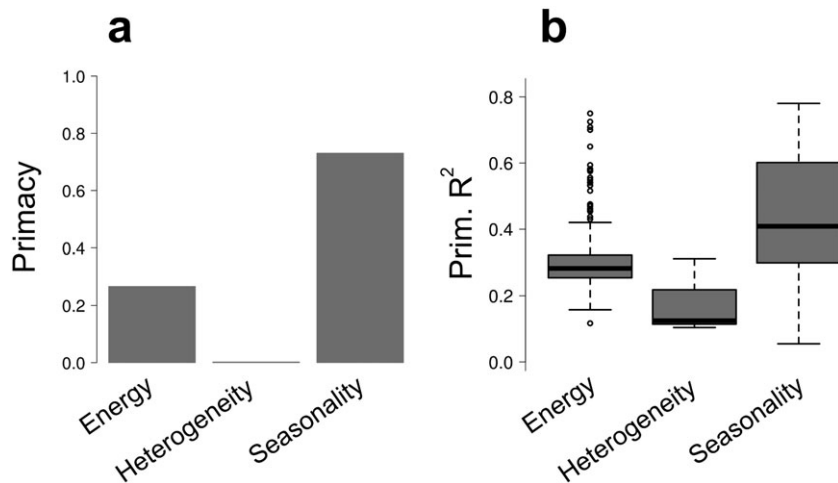
DISCUSSION

Stochastic processes as an explanation for richness–environment relationships

Many previous studies have demonstrated frequent and strong relationships between species richness and environmental gradients at broad geographic scales (Wright *et al.*, 1993; Hawkins *et al.*, 2003; Field *et al.*, 2009). All main explanations for these relationships assume a priori that richness–environment correlations reflect a causal relationship, where richness gradients are directly determined by environmental characteristics (Wright *et al.*, 1993; Currie *et al.*, 2004; Evans *et al.*, 2005). Nonetheless, stochastic processes could also lead to such correlations. In particular, stochastic models of distribution of species can produce

spatially structured species richness gradients (Colwell & Hurtt, 1994), and two spatially structured gradients distributed in the same domain are potentially correlated. Our simulations suggest that spurious correlations could explain, at least in part, the frequent species–environment relationships reported for many groups of organisms.

Simulation models used in this study are based on well-known evolutionary and biogeographical principles: species (1) originate from a spatially explicit process of speciation, (2) have limited geographic distributions, (3) shift their distributions through space, (4) go extinct, and (5) are distributed within constrained geographic domains (Gaston, 2003; Coyne & Orr, 2004; Arita & Vázquez-Domínguez, 2008). That these premises underlie the construction of richness gradients at broad scales is intuitive, and probably indisputable. An explanation for species
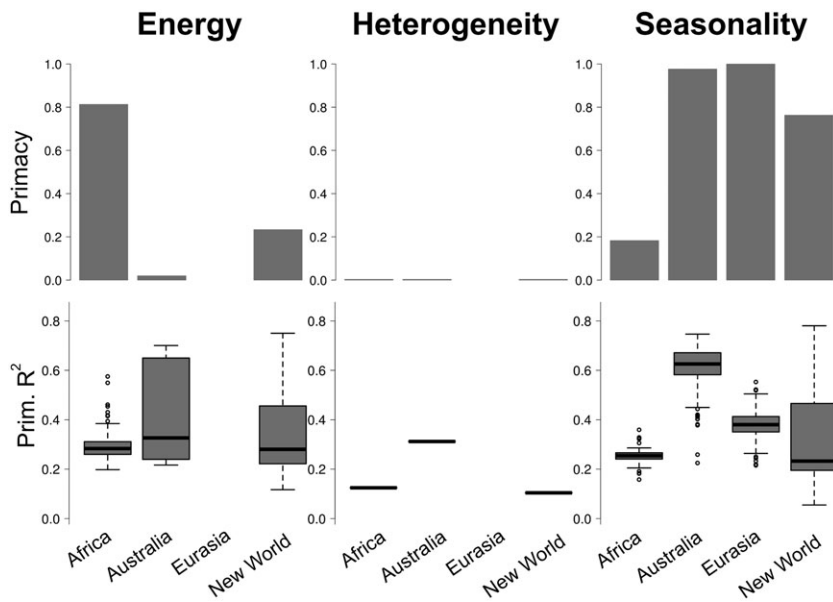
Figure 5 Simulated species–environment relationships produced by a diversification and range dynamics (DRD) model by continent. Species–environment relationships are compared among three environmental hypotheses (energy, heterogeneity and seasonality), in four different domains (Africa, Australia, Eurasia and the New World). Comparisons are based on primacy (a) and primary adjusted-R^2 (b).

richness gradients and for species–environment relationships that is based purely on these basic processes occurring stochastically across space and time is more parsimonious than an explanation involving influence of environmental factors. More traditional models (e.g. regressions of richness against temperature) do not include these processes explicitly, but assume that richness gradients are most proximally the result of these evolutionary/biogeographical first principles that determine the distribution of species. Thus, stochastic diversification and distribution of clades should not only be considered a possible explanation; it should also be the first to be scrutinized. We should consider more complex hypotheses involving additional processes only after concluding that purely stochastic evolution and distribution of species is not enough to explain species–environment correlations and spatial richness gradients.

We believe that research will probably demonstrate that simple coincidence of two independent gradients is not a complete explanation for many species–environment relationships; however, we consider that stochastic biogeographical and evolutionary processes have the potential to be important in some cases, or might interact with additional environmental and non-environmental mechanisms to produce richness gradients and richness–environment correlations. For example, Buckley *et al.* (2010) recently analysed the species–temperature relationships for a number of clades of mammals. They found that these relationships spanned a broad range of positive and negative values indicating that a single species–environment relationship is non-existent. We also found that there was not a single species–environment relationship expected by random diversification, but that there was considerable variation (Figs 2 & 3); some of this variation was associated with different domains and different environmental predictors. Buckley *et al.* (2010) suggested that the observed species–temperature relationships are probably the result of clade diversification plus phylogenetic niche conservatism, and not the result of the environment cre-

ating gradients in diversification rates or limits to species diversity. Nevertheless, how much of the pattern they document requires the role of niche conservatism, and how much could be accounted for by a purely stochastic geographic diversification model, where niche evolution is unconstrained, is unclear.

Implications for previous interpretations about the effects of different environmental characteristics

Our results have important implications for the interpretation of previously reported species–environment correlations. Many studies have compared the explanatory power of variables representing energy/climate versus environmental heterogeneity using some measure of strength of correlation (e.g. R^2 or F statistics; Currie, 1991; Tognelli & Kelt, 2004; Kreft & Jetz, 2007; Hortal *et al.*, 2008; Field *et al.*, 2009). Based on this research, many macroecologists have concluded that species richness is likely to be controlled by energetic and climatic determinants, while environmental heterogeneity is believed to be of little or no importance in most cases.

One of the most important conclusions we can reach from our results is that not all environmental characteristics have the same expected level of correlation with richness. As clades diversify and distribute stochastically, they tend to produce richness gradients that are likely to correlate significantly with variation in energy or climate; but these same clades produce richness gradients that are much more weakly correlated with variables that represent environmental heterogeneity. These results suggest that the reported predominance of energy/climate correlates might not necessarily represent evidence for a stronger effect of these environmental conditions on diversity gradients.

In our analyses, differences among environmental characteristics probably result from how environmental variables are distributed within domains. Typically, energetic and seasonality variables have relatively simple latitudinal gradients that are

partially a consequence of latitudinal variation in solar radiation and tilt of the earth with respect to the sun (Appendix S1, Fig. S1). These environmental gradients correlate relatively well with simple gradients produced by stochastic diversification and distribution of clades. In contrast, variables that represent environmental heterogeneity are typically strongly influenced by geological structures, such as major mountain chains. This makes the distribution of environmental heterogeneity more idiosyncratic (Appendix S1, Fig. S1), and consequently strong correlations with stochastic richness are less consistent.

Traditional comparisons between energy/climate and heterogeneity (and probably other hypotheses as well) might have been unfair or biased. The apparent importance of climate/energy over environmental heterogeneity might be, at least in part, due to their different probabilities of coincidental correlation with richness gradients. These results suggest reconsideration of previous evidence for the relative importance of different environmental hypotheses.

Previous research has also suggested that there are differences in how environmental variables correlate with species richness in different domains (e.g. Buckley & Jetz, 2007; Davies *et al.*, 2007). Our analyses indicate that domain differences could be expected simply from differences in: (1) how environmental variables are distributed within continents, and (2) how the geometry of a continent potentially affects richness gradients produced by stochastic diversification. However, our analyses are not exhaustive. Many studies have evaluated domains other than the ones we have used (e.g. Madagascar or Indo-Pacific oceans; Bellwood *et al.*, 2005; Lees & Colwell, 2007), or have divided continental masses into domains different from the ones we have defined (e.g. only South America or only sub-Saharan Africa; Jetz & Rahbek, 2002; Rahbek *et al.*, 2007). Despite the fact that not every domain has been evaluated in our analyses, we believe our results demonstrate the potential for spurious species–environment relationships to be possible under various domain configurations.

Stochastic simulations as null models for richness–environment relationships

Our results also suggest that we should reconsider the way we test for and compare the effects of multiple hypotheses. Species richness gradients are formed by overlap of species ranges, and current locations of these ranges are a consequence of the diversification and distribution of clades. Thus, a scenario where species speciate, go extinct, develop distributions and shift their geographic distributions randomly with respect to some particular mechanism of interest can be used as a null model to test the effects of such mechanism (Gotelli & Graves, 1996; Arita & Vázquez-Domínguez, 2008). Simulations like the ones used here can form the basis for more meaningful null models to test the effects of environmental characteristics on diversity gradients (see also Arita & Vázquez-Domínguez, 2008).

Although null models have an important history in ecology (Gotelli & Graves, 1996), much macroecological research has relied on simple OLS regression analyses. Null hypotheses implied by these regressions might be too null: they do not consider much of the relevant biology known about how species richness gradients are produced. Basic evolutionary/biogeographical processes occurring at random might lead to spatially structured richness gradients and consequently to spurious correlations with environmental variables.

Much has been discussed recently about appropriate statistical methods for studying the effects of predictors of species richness while accounting for spatial autocorrelation (Dormann *et al.*, 2007; Hawkins *et al.*, 2007; Bini *et al.*, 2009); as a consequence, many researchers have abandoned OLS and are using more complex spatial models. We believe this can be an important step forward, as these models can be a way to consider necessary corrections to regression coefficients, and might help alleviate some of the issues we have identified. Comparing spatial analyses with null model analyses requires further evaluation, but some preliminary analyses would suggest that spatial analyses are not enough to solve the problem we have identified with our simulations. In Appendix S1, we show that spatial models also produce spurious species–environment relationships, and that these relationships also differ among environmental characteristics and among continents (Appendix S1, Figs S3 & S4). Thus, these spatial methods might not be a complete solution for the problem identified in our study. Macroecologists might need to move away from the naive null hypothesis implied by most traditional statistical tests and instead use more appropriate null models (Gotelli & Graves, 1996; Arita & Vázquez-Domínguez, 2008).

Our results also suggest that R^2 values, F-values and regression coefficients calculated from regressions between richness and environmental predictors might be inflated or biased measures of effect size, making them inappropriate for comparing the effects of different predictors. When comparing the effects of multiple hypotheses, future research should consider their varying probabilities of correlation-by-chance with richness. One way is by estimating an effect size (ES) based on expected correlations produced by null models. The simplest option would be to use any measure of correlation between a predictor and richness to calculate a modified Hedges' d (Gotelli & Rohde, 2002; Hillebrand, 2008),

$$\text{ES} = \frac{C_{\text{obs}} - C_{\text{exp}}}{C_{\text{SD}}},$$

where C_{obs} is the observed correlation from the empirical richness–environment relationship, C_{exp} is the average correlation estimated from repeated null model runs and C_{SD} is the standard deviation of null species–environment correlations. Large positive or negative ES values would indicate that the observed effect is stronger than that expected under the null model. ESs like this could provide the basis for more appropriate comparisons of the relative importance of multiple predictors or hypotheses.

The distribution of regression coefficients also suggests that taking into account not only the strength, but also the directions of the species–environment relationships is fundamental. By

chance alone, the species–environment relationships are expected to change in direction among different environmental variables and among different continents. Empirical species–environment relationships need to be compared with these expectations produced by the random geographical diversification of clades.

Drawbacks of our simulation models

Our DRD model was conceived as an extension of simpler two-dimensional mid-domain models (e.g. Jetz & Rahbek, 2001). We tried to make the process by which richness gradients are constructed more realistic by including speciation, extinction and the dynamics of species distributions. This additional complexity, however, requires additional assumptions and many details of the model could have been specified differently (see, for example, Arita & Vázquez-Domínguez, 2008 or Connolly, 2009).

In our simulation model, there are multiple simplifications about processes underlying the construction of richness gradients. Three of the most important are the punctuated mode of speciation, the fixed nature of domains and the instantaneous extinction of species. The punctuated speciation that we have modelled in our simulations is possible, especially when speciation occurs by polyploidy (Otto & Whitton, 2000); but it is unlikely to be realistic for many other clades in which physical or ecological barriers are believed to have been important during speciation (Coyne & Orr, 2004; Phillimore et al., 2008). We have also assumed in our simulations that domains have not changed in size, shape or isolation from their present configuration. This is obviously not true; during the time that it has taken most real clades to diversify to their current stages, the geography of the earth has changed dramatically (Scotese, 2004). Finally, extinction, as incorporated into our model, does not take into account the process of range contraction that typically precedes extinction (Channell & Lomolino, 2000) and does not consider the varying probability of extinction as a function of species traits (like range size; e.g. Cardillo et al., 2008). Surely these and other assumptions made by our simulations have the potential to modify our results. However, we think it is unlikely that such additional complexities could lead to the destruction of spatially structured richness gradients produced by stochastic diversification and distribution of clades. Moreover, we believe that similar (if not more realistic) species–environment relationships could be expected in simulation models that consider additional complexities experienced by real-world species.

CONCLUSIONS

Our results suggest that even if diversification and distribution of clades occur independently of environmental gradients, some degree of correlation between richness and environment is expected simply by coincidence. Nevertheless, relationships between richness and environment have been mostly quantified and interpreted without considering this potential for spurious correlations. Thus, the frequency and strength of species–environment correlations could result, at least in part, from this coincidence of two independently produced gradients. Moreover, different environmental characteristics have different probabilities of spurious correlation with richness. By chance, energy and climate are more likely to correlate with richness than are measures of environmental heterogeneity. This bias could contribute to the reported predominance of energy and climate correlates of species richness. Our results suggest a re-evaluation of the frequency and strength of species–environment relationships using appropriate biogeographical/evolutionary null models. This might lead to an important reinterpretation of the determinants of diversity patterns at broad geographic extents.

REFERENCES

Arita, H.T. & Vázquez-Domínguez, E. (2008) The tropics: cradle, museum or casino? A dynamic null model for latitudinal gradients of species diversity. *Ecology Letters*, **11**, 653–663.

Beale, C.M., Lennon, J.J., Yearsley, J.M., Brewer, M.J. & Elston, D.A. (2010) Regression analysis of spatial data. *Ecology Letters*, **13**, 246–264.

Bellwood, D.R., Hughes, T.P., Connolly, S.R. & Tanner, J. (2005) Environmental and geometric constraints on Indo-Pacific coral reef biodiversity. *Ecology Letters*, **8**, 643–651.

Beyer, H.L. (2004) *Hawth's analysis tools for ArcGIS*. Available at: http://www.spatialecology.com/htools/ (accessed 15 May 2008).

Bini, L.M., Diniz-Filho, J.A.F., Rangel, T.F.L.V.B. *et al.* (2009) Coefficient shifts in geographical ecology: an empirical evaluation of spatial and non-spatial regression. *Ecography*, **32**, 193–204.

Bokma, F., Bokma, J. & Mönkkönen, M. (2001) Random processes and geographic species richness patterns: why so few species in the north? *Ecography*, **24**, 43–49.

Buckley, L.B. & Jetz, W. (2007) Environmental and historical constraints on global patterns of amphibian richness. *Proceedings of the Royal Society B: Biological Sciences*, **274**, 1167–1173.

Buckley, L.B., Davies, T.J., Ackerly, D.D., Kraft, N.J.B., Harrison, S.P., Anacker, B.L., Cornell, H.V., Damschen, E.I., Grytnes, J.-A., Hawkins, B.A., McCain, C.M., Stephens, P.R. & Wiens, J.J. (2010) Phylogeny, niche conservatism and the latitudinal diversity gradient in mammals. *Proceedings of the Royal Society B: Biological Sciences*, **277**, 2131–2138.

Cardillo, M., Mace, G.M., Gittleman, J.L., Jones, K.E., Bielby, J. & Purvis, A. (2008) The predictability of extinction: biological and external correlates of decline in mammals. *Proceedings of the Royal Society B: Biological Sciences*, **275**, 1441–1448.

Channell, R. & Lomolino, M.V. (2000) Trajectories to extinction: spatial dynamics of contraction of geographic ranges. *Journal of Biogeography*, **27**, 169–179.

Chernick, M.R. (2007) *Bootstrap methods: a guide for practitioners and researchers*, 2nd edn. Wiley-Interscience, Hoboken, NJ.

Colwell, R.K. & Hurtt, G.C. (1994) Nonbiological gradients in species richness and a spurious Rapoport effect. *The American Naturalist*, **144**, 570–595.

Colwell, R.K. & Lees, D.C. (2000) The mid-domain effect: geometric constraints on the geography of species richness. *Trends in Ecology and Evolution*, **15**, 70–76.

Colwell, R.K., Gotelli, N.J., Rahbek, C., Entsminger, G.L., Farrell, C. & Grave, G.R. (2009) Peaks, plateaus, canyons, and craters: the complex geometry of simple mid-domain effect models. *Evolutionary Ecology Research*, **11**, 355–370.

Connolly, S.R. (2009) Macroecological theory and the analysis of species richness gradients. *Marine macroecology* (ed. by J.D. Witman and K. Roy), pp. 279–309. University of Chicago Press, Chicago, IL.

Coyne, J.A. & Orr, H.A. (2004) *Speciation*. Sinauer Associates, Sunderland, MA.

Crawley, M.J. (2007) *The R book*. Wiley, Chichester.

Currie, D.J. (1991) Energy and large-scale patterns of animal-species and plant-species richness. *The American Naturalist*, **137**, 27–49.

Currie, D.J., Mittelbach, G.G., Cornell, H.V., Field, R., Guégan, J.F., Hawkins, B.A., Kaufman, D.M., Kerr, J.T., Oberdorff, T., O'Brien, E. & Turner, J. (2004) Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecology Letters*, **7**, 1121–1134.

Davies, T.J., Grenyer, R. & Gittleman, J.L. (2005) Phylogeny can make the mid-domain effect an inappropriate null model. *Biology Letters*, **1**, 143–146.

Davies, R.G., Orme, C.D.L., Storch, D., Olson, V.A., Thomas, G.H., Ross, S.G., Ding, T.S., Rasmussen, P.C., Bennett, P.M., Owens, I.P.F., Blackburn, T.M. & Gaston, K.J. (2007) Topography, energy and the global distribution of bird species richness. *Proceedings of the Royal Society B: Biological Sciences*, **274**, 1189–1197.

Dormann, C.F., McPherson, J.M., Araújo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Kissling, W.D., Kühn, I., Ohlemüller, R., Peres-Neto, P.R., Reineking, B., Schröder, B., Schurr, F.M. & Wilson, R. (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, **30**, 609–628.

Evans, K.L., Warren, P.H. & Gaston, K.J. (2005) Species–energy relationships at the macroecological scale: a review of the mechanisms. *Biological Reviews*, **80**, 1–25.

Field, R., Hawkins, B.A., Cornell, H.V., Currie, D.J., Diniz-Filho, J.A.F., Guégan, J.-F., Kaufman, D.M., Kerr, J.T., Mittelbach, G.G., Oberdorff, T., O'Brien, E.M. & Turner, J.R.G. (2009) Spatial species-richness gradients across scales: a meta-analysis. *Journal of Biogeography*, **36**, 132–147.

Fox, J. (2009) *car: companion to applied regression*. Available at: http://cran.r-project.org/web/packages/car/index.html (accessed 1 June 2010).

Gaston, K.J. (2003) *The structure and dynamics of geographic ranges*. Oxford University Press, Oxford.

Gotelli, N.J. & Graves, G.R. (1996) *Null models in ecology*. Smithsonian Institution Press, Washington, DC.

Gotelli, N.J. & Rohde, K. (2002) Co-occurrence of ectoparasites of marine fishes: a null model analysis. *Ecology Letters*, **5**, 86–94.

Gotelli, N.J., Anderson, M.J., Arita, H.T. *et al.* (2009) Patterns and causes of species richness: a general simulation model for macroecology. *Ecology Letters*, **12**, 873–886.

Hawkins, B.A., Field, R., Cornell, H.V., Currie, D.J., Guégan, J.-F., Kaufman, D.M., Kerr, J.T., Mittelbach, G.G., Oberdorff, T., O'Brien, E.M., Porter, E.E. & Turner, J.R.G. (2003) Energy, water, and broad-scale geographic patterns of species richness. *Ecology*, **84**, 3105–3117.

Hawkins, B.A., Diniz-Filho, J.A.F., Bini, L.M., De Marco, P. & Blackburn, T.M. (2007) Red herrings revisited: spatial autocorrelation and parameter estimation in geographical ecology. *Ecography*, **30**, 375–384.

Hennig, W. (1966) *Phylogenetic systematics*. University of Illinois Press, Urbana, IL.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hillebrand, H. (2008) Meta-analysis in ecology. pp. 3–4. *Encyclopedia of life sciences (ELS)*. Wiley, Chichester.

Hortal, J., Rodríguez, J., Nieto-Díaz, M. & Lobo, J.M. (2008) Regional and environmental effects on the species richness of mammal assemblages. *Journal of Biogeography*, **35**, 1202–1214.

Imhoff, M.L., Bounoua, L., Ricketts, T., Loucks, C., Harriss, R. & Lawrence, W.T. (2004) Global patterns in human consumption of net primary production. *Nature*, **429**, 870–873.

Jablonski, D., Roy, K. & Valentine, J.W. (2006) Out of the tropics: evolutionary dynamics of the latitudinal diversity gradient. *Science*, **314**, 102–106.

Jetz, W. & Rahbek, C. (2001) Geometric constraints explain much of the species richness pattern in African birds. *Proceedings of the National Academy of Sciences USA*, **98**, 5661–5666.

Jetz, W. & Rahbek, C. (2002) Geographic range size and determinants of avian species richness. *Science*, **297**, 1548–1551.

Kalmar, A. & Currie, D.J. (2007) A unified model of avian species richness on islands and continents. *Ecology*, **88**, 1309–1321.

Kreft, H. & Jetz, W. (2007) Global patterns and determinants of vascular plant diversity. *Proceedings of the National Academy of Sciences USA*, **104**, 5925–5930.

Lees, D.C. & Colwell, R.K. (2007) A strong Madagascan rainforest MDE and no equatorward increase in species richness: re-analysis of 'The missing Madagascan mid-domain effect', by Kerr J.T., Perring M. & Currie D.J. (*Ecology Letters* **9**:149–159, 2006). *Ecology Letters*, **10**, E4–E8.

Mittelbach, G.G., Schemske, D., Cornell, H.V. *et al.* (2007) Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecology Letters*, **10**, 315–331.

Otto, S.P. & Whitton, J. (2000) Polyploid incidence and evolution. *Annual Review of Genetics*, **34**, 401–437.

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

Phillimore, A.B., Orme, C.D., Thomas, G.H., Blackburn, T.M., Bennett, P.M., Gaston, K.J. & Owens, P.F. (2008) Sympatric speciation in birds is rare: insights from range data and simulations. *The American Naturalist*, **171**, 646–657.

R Development Core Team (2008) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna. Available at: http://www.R-project.org (accessed 1 June 2010).

Rahbek, C., Gotelli, N.J., Colwell, R.K., Entsmiger, G.L., Rangel, T. & Graves, G.R. (2007) Predicting continental-scale patterns of bird species richness with spatially explicit models. *Proceedings of the Royal Society B: Biological Sciences*, **274**, 165–174.

Rangel, T. & Diniz-Filho, J.A.F. (2005) Neutral community dynamics, the mid-domain effect and spatial patterns in species richness. *Ecology Letters*, **8**, 783–790.

Rohde, K. (1992) Latitudinal gradients in species-diversity – the search for the primary cause. *Oikos*, **65**, 514–527.

Scotese, C.R. (2004) Cenozoic and Mesozoic paleography: changing terrestrial biogeographic pathways. *Frontiers of biogeography: new directions in the geography of nature* (ed. by M.V. Lomolino and L.R. Heaney), pp. 9–26. Cambridge University Press, Cambridge.

Sokal, R.R. & Rolf, F.J. (1994) *Biometry. the principles and practices of statistics in biological research*, 3rd edn. W. H. Freeman, New York, NY.

Stephens, P.R. & Wiens, J.J. (2003) Explaining species richness from continents to communities: the time-for-speciation effect in Emydid turtles. *The American Naturalist*, **161**, 112–128.

Stevens, R.D. (2006) Historical processes enhance patterns of diversity along latitudinal gradients. *Proceedings of the Royal Society B: Biological Sciences*, **273**, 2283–2289.

Tello, J.S. & Stevens, R.D. (2010) Multiple environmental determinants of regional species richness and effects of geographic range size. *Ecography*, **33**, 796–808.

Tognelli, M.F. & Kelt, D.A. (2004) Analysis of determinants of mammalian species richness in South America using spatial autoregressive models. *Ecography*, **27**, 427–436.

Wiens, J.J. & Donoghue, M.J. (2004) Historical biogeography, ecology and species richness. *Evolution*, **19**, 639–644.

Willig, M.R. & Lyons, S.K. (1998) An analytical model of latitudinal gradients of species richness with an empirical test for marsupials and bats in the New World. *Oikos*, **81**, 93–98.

Willig, M.R., Kaufman, D.M. & Stevens, R.D. (2003) Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. *Annual Review of Ecology, Evolution, and Systematics*, **34**, 273–309.

Wright, D.H., Currie, D.J. & Maurer, B.A. (1993) Energy supply and patterns of species richness on local and regional scales. *Species diversity in ecological communities* (ed. by R.E. Ricklefs and D. Schluter), pp. 66–74. University of Chicago Press, Chicago, IL.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Additional figures, tables and analyses.
Appendix S2 Annotated R code.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

BIOSKETCHES

J. Sebastian Tello is interested in macroecology, biodiversity, metacommunity ecology and the use of biologically meaningful null models to test ecological and evolutionary hypotheses. He is currently a PhD candidate at Louisiana State University, and an associated researcher at the zoology museum at Pontificia Universidad Católica del Ecuador.

Richard D. Stevens is an associate professor in the Department of Biological Sciences at Louisiana State University. He is interested in understanding the processes that determine the community ecology, biogeography and macroecology of mammals.

Editor: Marcel Cardillo