Evolutionary Ecology Research, 2009, **11**: 355–370

Peaks, plateaus, canyons, and craters: the complex geometry of simple mid-domain effect models

Robert K. Colwell[1], Nicholas J. Gotelli[2], Carsten Rahbek[3],
Gary L. Entsminger[4], Catherine Farrell[5] and Gary R. Graves[6]

[1]*Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, Connecticut, USA,* [2]*Department of Biology, University of Vermont, Burlington, Vermont, USA,* [3]*Center of Macroecology and Evolution, Department of Biology, University of Copenhagen, Copenhagen, Denmark,* [4]*Acquired Intelligence Inc. and Rocky Mountain Biological Laboratory, Montrose, Colorado, USA,* [5]*Essex High School, Essex Junction, Vermont, USA and* [6]*Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA*

ABSTRACT

Background: Geographic ranges, randomly located within a bounded geographical domain, produce a central hump of species richness (the mid-domain effect, MDE). The hump arises from geometric constraints on the location of ranges, especially larger ones.

Questions: (1) How do patterns of species richness in one- and two-dimensional MDE models change as a function of range size? (2) How does dispersal affect these patterns?

Methods: We used a spreading dye algorithm to place assemblages of species of uniform range size in one-dimensional or two-dimensional bounded domains. In some models, we allowed dispersal to introduce range discontinuity.

Results: As uniform range size increases from small to medium, a flat pattern of species richness is replaced by a pair of peripheral peaks, separated by a valley (one-dimensional models), or by a cratered ring (two-dimensional models) of species richness. With large range sizes, the peaks or rings fuse to form a central plateau (one-dimensional) or a flat-topped mound (two-dimensional) of highest species richness. Adding dispersal to the two-dimensional model weakens the peripheral ring and introduces complex patterns for long-distance dispersal.

Conclusions: Heterogeneous range size distributions (whether theoretical or empirical) used in most MDE models produce species richness patterns dominated by wide-ranged species, hiding complex patterns formed by small- to medium-ranged species. These patterns, which are analogous for one and two dimensions, are complicated further by long-distance dispersal and discontinuous ranges. Although geometric constraints lead to classic mid-domain effects for large-ranged species and for mixed range-size frequency distributions, small- and medium-sized ranges of a uniform size generate more complex patterns, including peaks, plateaus, canyons, and craters of species richness.

Keywords: biogeography, boundary effects, doughnut, geographical range, geometric constraints, null models, spatial scale, species richness gradients, stochastic models.

Correspondence: R.K. Colwell, Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT 06269-3043, USA. e-mail: colwell@uconn.edu

INTRODUCTION

Documenting the relative importance of contemporary climate, historical effects, and geometric constraints on species richness patterns has been a major research focus in biogeography and macroecology (Willig *et al.*, 2003). Colwell and Hurtt (1994) introduced one-dimensional stochastic models of species richness in an attempt to explain Rapoport's rule, the tendency for range sizes of species to decrease towards the tropics (Stevens, 1989). While exploring these models, Colwell and Hurtt (1994) stumbled upon a surprising result that itself had nothing to do with Rapoport's rule.

Suppose a transect, or *domain*, is modelled as a one-dimensional line segment, and geographic ranges are represented as overlapping, randomly sized, smaller line segments, located at random positions within the domain. In the simplest case, the domain is a line of length one (the unit line), and geographic ranges are represented by line segments ranging in length from zero (as a limit) to one. What is the expected pattern of species richness (the number of overlapping ranges) at any point along the transect? Naively, one might expect that a random distribution of species along a bounded geographical transect would produce, on average, a uniform distribution of richness. But this turns out to be wrong. Under a wide range of model assumptions, the pattern of species richness within the domain is a hump that declines symmetrically from the centre towards the edges of the domain. Later, Colwell and Lees (2000) named this phenomenon the 'mid-domain effect', or MDE. Whereas previous explanations for species richness gradients could be broadly classified as either evolutionary (emphasizing historical factors) or ecological (emphasizing contemporary climate), the mid-domain effect suggested an entirely novel class of causal explanations: non-uniform species richness gradients might arise because of geometric constraints imposed by a bounded domain.

Is the mid-domain effect a null model or an explanatory factor?

As of the end of 2008, more than 100 papers had addressed the mid-domain effect, conceptually, theoretically or as a component of the analysis of empirical data sets. Conceptually, the issue is whether MDE explains anything we did not already know about biogeographical patterns of species richness in nature. If so, how much does it explain and under what conditions? Like any idea that calls for an entirely new way of looking at an old problem, MDE has, at times, been either too quickly embraced or too quickly dismissed, but has generally met with appropriate scepticism and gradual acceptance.

Part of the problem from the start has been that MDE models are null models, but not typical ones (Colwell *et al.*, 2004, 2005). They are null models because they deliberately exclude a key factor or mechanism believed to be important in nature, with the aim of assessing the role of the missing factor in shaping natural patterns (Gotelli and Graves, 1996). In the case of MDE, the factor left out by design is any *direct* role of environmental gradients in driving patterns of species richness. No-one doubts that environmental gradients play a strong role in adaptive evolution of populations and in shaping species distributions and range sizes (Brown *et al.*, 1996; Parmesan *et al.*, 2005). But the role of such gradients in shaping patterns of *overlap* of species ranges (species richness) – to which MDE models are blind by design – is less obvious. This leaves an opening for MDE to serve as a null model of the patterns of range overlap to be expected in the absence of any direct influence of environment on richness (e.g. Connolly *et al.*, 2003).

But what if the simple stochastic phenomenon that drives the pattern of richness in the model contributes, in it own right, to patterns in nature? MDE becomes in this case not just a null model, but a mechanistic model of the effect of geometric constraints that arise from the boundedness of domains, which can act in concert or in conflict with other causes of geographical pattern in species richness, such as primary productivity, topographic complexity or area. The studies of Jetz and Rahbek (2002) and Brehm *et al.* (2007) exemplify many others in that they show MDE acting in concert with environmental factors, whereas MDE and environmental factors were shown to be spatially discordant by, for example, McCain (2003) and Lees *et al.* (1999, with critique and response by Kerr *et al.*, 2006; Currie and Kerr, 2007; Lees and Colwell, 2007). Interactions between area and MDE emerged from other studies (e.g. Rahbek, 1997; Sanders, 2002; Bachman *et al.*, 2004; Romdal *et al.*, 2005; Romdal and Grytnes, 2007). Currie and Kerr (2008) have emphasized the problems of interpretation that arise when MDE predictions and environmental factors are collinear, although the same issues arise when environmental variables are collinear among themselves. Hawkins and Diniz-Filho (2002), Zapata *et al.* (2003, 2005), and Hawkins *et al.* (2005) have criticized MDE on several grounds, but as we have already responded in full to these critiques (Colwell *et al.*, 2004, 2005), we will not recount these discussions here.

Because any influence of MDE on spatial patterns of richness in nature is expected to interact in complex ways with other candidate explanatory variables, an appropriate framework for analysis is multivariate model selection, rather than simple hypothesis testing (Colwell *et al.*, 2004, 2005). We therefore do not see a basis for the continued treatment of MDE (most recently, Currie and Kerr, 2008) as a simple null *hypothesis*, to be rejected outright if it fails to fully explain patterns in nature.

Range size and geometric constraints in models and in nature

As Colwell and Hurtt (1994) noted, the geometric constraint on range location within a bounded domain increases linearly with range size. For a unit-line domain, the midpoint of a range of length r ($0 < r \leq 1$) is geometrically constrained to be located over a distance of $1 - r$ in the middle of the domain. For this reason, in a domain with a gradient of environmental 'favourability', a key prediction of MDE theory is that the location of smaller ranges within a domain is expected to be influenced more by the environment and less by geometric constraints, compared with larger ranges (Colwell *et al.*, 2004, 2005). Rangel and Diniz-Filho (2005) demonstrated this effect elegantly for a Hubbell-style neutral evolutionary model with spatially explicit speciation and extinction. In their model, species richness is influenced by stochastic range shifts weighted towards the 'favourable' end of a simple spatial gradient. On shallow gradients, ranges are naturally larger for a given environmental tolerance (niche breadth) than on steeper gradients (Kirkpatrick and Barton, 1997). As a result, a strong MDE develops on shallow gradients, whereas species richness accumulates in a peak near the 'favourable' end of a steep gradient.

The prediction that larger ranges are more likely to show the signature of MDE than smaller ranges has repeatedly been borne out by empirical studies of richness patterns (beginning with Pineda and Caswell, 1998; Lees *et al.*, 1999; Jetz and Rahbek, 2001). To assess the role of MDE for smaller versus larger ranges, these studies deconstructed patterns of species richness by dividing the empirical distribution of range sizes into quantiles (halves or quartiles) by rank or by absolute range size. Separate richness predictions are then generated for the ranges of each quantile. A recent review and meta-analysis of this approach by Dunn *et al.* (2007)

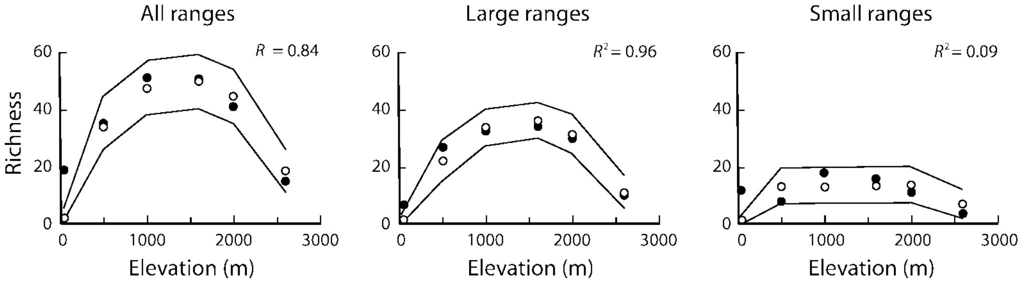

Fig. 1. Observed and predicted species richness as a function of elevation for ferns, on the Barva Transect, Costa Rica. The left-hand graph displays results for all fern species, whereas the middle and right-hand graphs show the 50% of fern species with larger and smaller ranges, respectively. Solid circles indicate observed richness for each field sampling elevation. Open circles show mean richness and fine lines give the 95% confidence intervals predicted by a mid-domain effect 'range-shuffling' model (Connolly, 2005; Colwell and Hurtt, 1994, Model 2, using empirical ranges). The coefficient of determination for the regression of observed on expected appears on each graph (data from Cardelús *et al.*, 2006).

concludes that, overall, larger ranges fit MDE predictions better than smaller ranges from the same domain. Figure 1 shows an example for fern species on a tropical elevational transect (Cardelús *et al.*, 2006).

Pursuing this approach to richness deconstruction, our aim in this study is to explore the deep structure of a widely used class of MDE models, 'spreading dye' models (Jetz and Rahbek, 2001). We do this by simplifying and then varying the data that go into them – the distribution of geographical range sizes – in relation to the size of the geographical domain. We extend this approach to examine the effect of range discontinuity by relaxing the spreading dye model assumption of strict range cohesion and introducing long-distance dispersal. We show that the classic MDE pattern, a mid-domain hump or sloped plateau, while typical for species with medium to large range sizes, is but one of many patterns that readily emerge from stochastic models of range location. At smaller range sizes, more complex landscapes of species richness – including peaks, plateaus, canyons, and craters – are generated by the spreading dye model.

Conceptual and historical context of the spreading dye model

Colwell and Hurtt (1994) explored MDE using a Monte Carlo approach. Lyons and Willig (1997), likewise pursuing a stochastic approach to modelling Rapoport's rule, independently discovered the mid-domain effect (M. Willig, personal communication). Later, they developed a simple analytical MDE model based on binomial probabilities for range limits on the unit line (Willig and Lyons, 1998). Eventually, David Lees [Box 2 in Colwell and Lees (2000)] was able to show that Colwell and Hurtt's (1994) Model 2, Willig and Lyons's (1998) binomial model, and MacArthur's (1957) Type II ('overlapping niche') broken stick distribution are mathematically identical. (Because MacArthur was modelling the relative abundance distribution of species, he was concerned with the length of the 'pieces' and did not explore the pattern of their overlap.) While studying the latitudinal distribution of seaweeds along shorelines, Pielou (1977a, 1977b) was the first to apply this model to the study of biogeographical patterns, but she considered only pairwise overlap and did not take note of the pattern of multiple-species overlap (richness) (E.C. Pielou, personal communication).

In this 'fully stochastic' model, as Colwell and Lees (2000) called it, not only does the overlap of ranges (species richness) peak in the middle of the domain, but so does the distribution of range midpoints, because larger ranges are constrained geometrically to have their midpoints towards the centre of the domain, as first illustrated by Graves (1985). Reviewing a draft of Colwell and Hurtt (1994), Michael Rosenzweig (personal communication) questioned whether the midpoint distribution by itself might be driving the humped pattern of richness. In response, Colwell and Hurtt (1994) created an alternative model (Model 3) in which range midpoints were constrained to follow a uniform distribution across the domain. In this doubly constrained 'uniform midpoint' model, range sizes were stochastic within the midpoint distribution constraint as well as within the constraint posed by the domain bounds. The shape of the hump changed somewhat, but it was still a hump (Figure 4 in Colwell and Hurtt, 1994). This model treats range size stochastically, conditional upon the specified range midpoint. Lees *et al.* (1999) [see also Box 3 in Colwell and Lees (2000)] developed the opposite model, which treats the range midpoint stochastically, conditional upon the specified range size. Colwell (2008) provides all three models (as well as others) in the freeware application *RangeModel*, for one-dimensional domains.

Rahbek (1997), Pineda and Caswell (1998), Willig and Lyons (1998), and Lees *et al.* (1999) were the first to treat MDE as a null model for biogeographical patterns in the natural world. Pineda and Caswell (1998) introduced the key idea of randomizing the location of empirical ranges within the domain, while maintaining the empirical range size frequency distribution. This approach has been applied repeatedly in later studies using both one- and two-dimensional domains (reviewed by Colwell *et al.*, 2004, 2005 and Dunn *et al.*, 2007; for more recent examples of this approach, see Cardelús *et al.*, 2006; Storch *et al.*, 2006; Brehm *et al.*, 2007; Davies *et al.*, 2007; Rahbek *et al.*, 2007; Arita and Vázquez-Domínguez, 2008).

Mid-domain effect models were originally proposed for one-dimensional domains, but were later extended to two dimensions, initially by Lees *et al.* (1999), Bokma *et al.* (2001), and Jetz and Rahbek (2001, 2002). Bokma and colleagues (2001) developed two approaches. One was an evolutionary, cellular automaton algorithm that presaged later evolutionary models by Rangel and Diniz-Filho (2005), Brayard *et al.* (2005), Rangel *et al.* (2007), and Arita and Vázquez-Domínguez (2008). In a second approach, Bokma *et al.* (2001) extended the binomial model of Willig and Lyons (1998) to two dimensions, in both cases allowing both range size and range placement to emerge from the model. Arita (2005) further explored the properties of the two-dimensional binomial model analytically, focusing on range size distributions in different regions of a hypothetical square domain.

In contrast, Jetz and Rahbek (2001) based their two-dimensional model on the strategy of Pineda and Caswell (1998) and Lees *et al.* (1999), both one-dimensional approaches, enforcing a match of the modelled range size frequency distribution (RSFD) to an *a priori* RSFD, with stochastic range location. In this model, each range has a predetermined size – the range is drawn either from an empirical RSFD or from a theoretical RSFD, rather than emerging from colonization and extinction dynamics. The approach of Jetz and Rahbek (2001) has come to be known as the 'spreading dye model' (or 'spreading dye algorithm').

The spreading dye algorithm, first suggested by Gotelli and Graves (1996, p. 256), has become the approach of choice for two-dimensional domains, and has formed the basis for more complex models of species richness that incorporate both geometric constraints and other factors (e.g. Jetz and Rahbek, 2002; Rangel and Diniz-Filho, 2003; Storch *et al.*, 2006; Davies *et al.*, 2007; Rahbek *et al.*, 2007). Connolly (2005) developed an analytical version of a spreading dye model for one-dimensional domains, and Colwell (2008) implemented a one-dimensional Monte Carlo version in the freeware application *RangeModel*. An analytical version of the spreading dye model has not been developed for two-dimensional domains and may not be possible, except for geometrically simple domains such as perfect circles, triangles, and squares.

To date, theoretical studies of the mid-domain effect have used heterogeneous distributions of range sizes (either empirical or simulated) to generate MDE patterns. These studies have emphasized the interaction between geographic range size and domain size, usually revealed by partitioning the RSFD into quantiles. We conjectured that it might be revealing to take the quantile approach to its logical limit in studying the behaviour of simple MDE models. In this study, we explore the MDE patterns that are expected for assemblages of species that have identical range sizes. Independently, Connolly (in press) has explored the spreading dye model with the same strategy.

Objectives

We use Monte Carlo versions of both one- and two-dimensional spreading dye models to explore the role of range size in driving species richness patterns. We ask two questions: (1) How do the predictions of MDE models change as a function of geographic range size, in both one and two dimensions? (2) How do the predictions of MDE models change as the assumption of strict range cohesion is relaxed and species are allowed to disperse longer distances within a bounded geographic domain?

MATERIALS AND METHODS

Domains and range sizes

We simulated the random origin and spread of species in a gridded geographic domain. The one-dimensional domain was 1 × 100 linear grid cells (line segments) and the two-dimensional domain was 100 × 100 grid cells (identical, square cells). We created assemblages of 100 species, in which all species shared an identical range size. For the one-dimensional domain, these range sizes were 1, 2, 4, 8, 16, 32, 64, 80, and 100 grid cells. For the two-dimensional domain, the range sizes were 1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048, 4096, 8192, 9000, and 10,000 grid cells. Because all range sizes are identical within each assemblage, we can directly assess the contribution of range size to MDE patterns. In contrast, in most MDE analyses, empirical or simulated range size frequency distributions are heterogeneous, with a mixture of larger and smaller ranges. When range sizes are not equal, patterns of species richness are dominated by the effects of widespread species, making it difficult to assess the spatial patterns and contributions of smaller ranges (Lees *et al.*, 1999; Jetz and Rahbek, 2001, 2002; Jetz *et al.*, 2004; Lennon *et al.*, 2004).

Simulation models: spreading dye model

We implemented a spreading dye model (Jetz and Rahbek, 2001) by randomly choosing an initial grid cell to represent the origin of a species' geographic range. Subsequent cells for the range were also chosen randomly, but range cohesion was enforced by restricting the candidate cells to those adjacent to a currently occupied cell. In the one-dimensional domain, at each time step only one or two grid cells were available for range expansion, depending on whether the species had reached the edge of the domain or not. In the two-dimensional model, a maximum of eight adjacent cells was available for range expansion, as in a king's move in chess. Qualitatively similar results were obtained for two-dimensional models in which only a maximum of four adjacent, non-diagonal cells were available (a one-square rook's move in chess). The simulation continued by randomly choosing and filling grid cells until the geographic range for a species was complete. The next species in the assemblage was then seeded in a new random grid cell, without consideration of previously placed ranges. The procedure was repeated until all 100 species ranges in the assemblage had been placed. Because no cells outside the boundaries of the domain were permitted to be occupied, an expanding range that 'hit' a boundary expanded, as necessary, towards the centre of the domain, in the one-dimensional model, or both inwards and laterally along the boundary, in the two-dimensional model.

We then determined the total species richness in each grid cell, which could theoretically range from zero (if a grid cell was never chosen) to 100 (if all species occurred in the grid cell). As range size is increased, the minimum and average species richness per grid cell also increases. In the extreme, for models in which the range size equals the domain size, observed species richness in every grid cell is inevitably 100, with no variation among runs.

We simulated each assemblage 300 times and calculated the average species richness for each grid cell. For the two-dimensional model, the results were plotted in a bivariate colour plot, in which the 'hot colours' (red and orange) indicate grid cells with large average species richness, and the 'cold colours' (blue and green) indicate grid cells with low average species richness. Colour plots were scaled internally so that variation in richness within the domain could be visualized most easily.

Stepping-stone model

For the two-dimensional domain, we built a model that relaxed the assumption of strict range cohesion in the spreading dye model. The initial cell for a range was chosen at random, just as in the algorithm specified above. To choose the second and subsequent cells, one of the cells already occupied by the species was chosen at random to be the source of a colonization event. (For choosing the second cell, this source cell is simply the initial cell.) A compass orientation (0 to 360°) was selected at random as a dispersal direction for the subsequent ('destination') cell. The dispersal distance to the destination cell from the source cell was chosen as a random variable from a Poisson distribution:

$$P(d) = \frac{\lambda^d e^{-\lambda}}{d!},$$

where $P(d)$ is the probability of dispersing distance d, and λ is the Poisson parameter. Dispersal distance d was measured in grid cell units. If the destination grid cell selected in this way was not already occupied by the species in question, it was added to the species' geographic range. If the destination grid cell was already occupied by the species, or if the dispersal distance was beyond the map boundary, the same source cell was used and a new destination grid cell was selected. This process was repeated until the geographic range was filled for a given species. All species were simulated in the same way (with identical λ values), and average species richness per grid cell was calculated and plotted as described for the spreading dye model. All simulations were carried out using *BioGeoSim*, a dedicated software package programmed in Delphi for modelling geographic ranges and species richness patterns in gridded domains (Gotelli *et al.*, 2007).

RESULTS

One-dimensional spreading dye model

Figure 2 depicts the effects of geometric constraints on ranges of different size in a one-dimensional domain. For an assemblage in which all species have a range size of only one grid cell, the pattern is purely random and uniform, because no geometric constraints are in play. As range size increases (range size $r = 2$ to 4 grid cells), the average species richness in the centre of the range shows a reduction in the variance, because ranges begin to overlap more often, so the richness count evens out spatially. Meanwhile, two symmetric peripheral 'peaks' appear near the edges of the domain, reflecting a mixture of ranges that landed within the boundaries, but extend into the area of the peaks, and ranges that originated on or near the edge of the domain and then randomly accumulated back towards the centre. As range size increases (range size $r = 8$ to 32 grid cells), the relative height of these peaks increases, with a progressive narrowing of the canyon floor of richness near the centre of the domain. This narrowing is caused by the increasing penetration into the interior of the domain by larger and larger 'reflected' ranges that were forced inwards.

With the transition from 32 to 64 grid cells per range, range size now exceeds half of the 100-cell domain, so that the canyon floor inverts to a flat-topped plateau. The plateau lies between $100 - r$ on the left and r on the right side domain, with a richness of exactly 100 species, because all ranges must overlap in this region. Thus, as range size r increases from 64 to 100 grid cells, the width of the richness plateau expands. Connolly's (2005) analytical model for the one-dimensional spreading dye model specifies the exact quantitative expectations for all regions of the richness plots in Fig. 2.

Two-dimensional spreading dye model

Figure 3 shows the colour-graded pattern of average species richness with increasing range size in a two-dimensional geographic domain. At small range sizes (1 to 8 grid cells), the pattern is virtually random, although as range size increases, more small-scale aggregations appear in the domain. (Repeated runs of the model show that these aggregations do not appear consistently in the same locations.) At larger range sizes (16 to 128 grid cells), an increasing concentration of species richness emerges in a band running parallel to the edge of the domain, concentrated even more strongly near the corners at higher range sizes. These bands correspond to the peripheral peaks in the one-dimensional model, and like them move towards the centre of the domain as range size increases. The four local richness peaks near the corners of the domain arise from the intersection of the peripheral bands, because ranges that are 'reflected' from both adjacent sides overlap in the corners. At still larger range sizes (256 to 2048 grid cells), the band expands to form a distinct ring or doughnut, with a crater of species richness in the centre of the domain. The ring contracts for the same reason the peaks converge in the one-dimensional model: larger ranges are reflected by the borders and forced further towards the centre. In the transition from 2048 to 4096 grid cells, the doughnut coalesces to a classical mid-domain peak, just as in the one-dimensional model. As the range size is increased beyond 5000 grid cells, half the size of the domain, the area of the peak becomes a plateau, and expands towards the edges of the domain (8192 to 9000 grid cells).

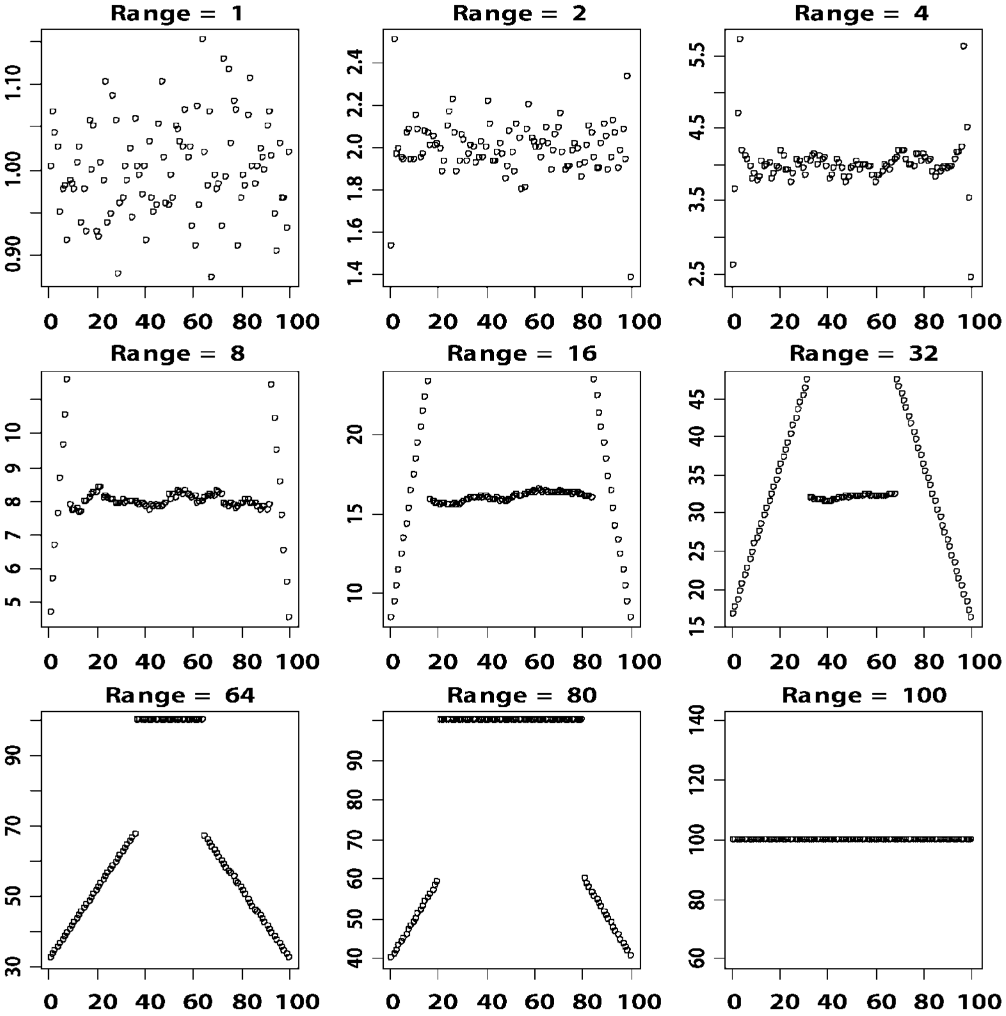

Fig. 2. One-dimensional expected species richness plots. The x-axis represents position along the number line, while the y-axis represents expected species richness. Each plot represents a different simulation for an artificial assemblage of 100 species with identical range sizes. Each point represents the average species richness in the one-dimensional domain, based on 300 stochastic runs of the model. Note that the y-axis scale varies among the plots.

Two-dimensional stepping-stone model

Figure 4 depicts the simultaneous effects of altering dispersal distance (λ) and geographic range size in the stepping-stone model. At very small dispersal distances ($\lambda = 0.5$ to 1.0), the model behaves much like the two-dimensional spreading dye model with strict range cohesion, although the formation of the ring or doughnut of species richness at intermediate range sizes is not so apparent, probably because reflected ranges are less constrained in their location. As the dispersal distance increases ($\lambda = 2$ to 16), a broad mid-domain peak emerges at intermediate to large geographic range sizes, with only a narrow band of decreased richness near the edges and corners of the domain.

As λ increases from 16 to 32 grid cell units, the patterns change abruptly, producing a second wave of peripheral richness peaks that coalesce centrally at high range sizes, but which fail to become circular. We conjecture that this abrupt shift and pattern are caused by dispersal distances being so long that many trial destination cells fall outside the domain. Repeated trials do not find all grid cells in the domain with equal frequency, imitating the pattern with no dispersal (Fig. 3). As λ increases from 32 to 128 grid cell units, a complex

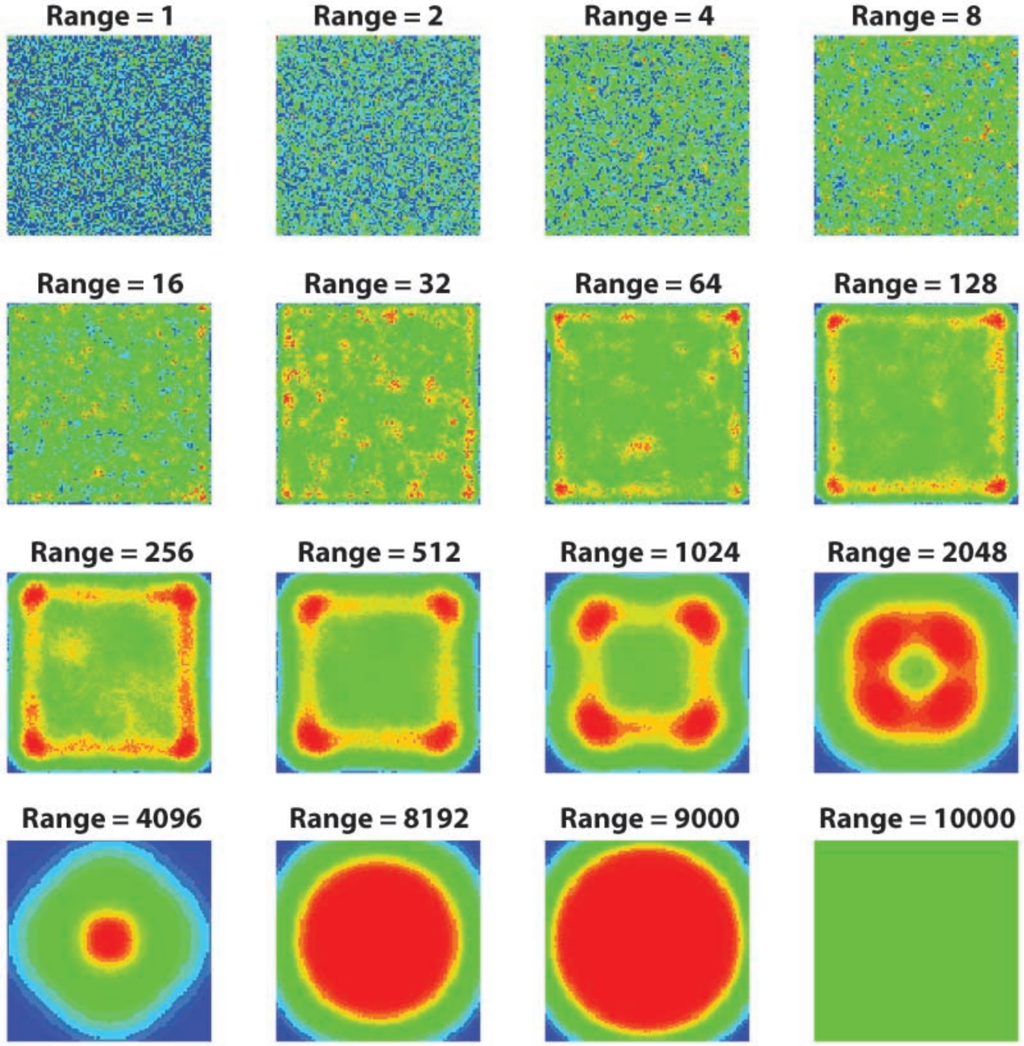

Fig. 3. Contour plots representing expected numbers of species in a 100 × 100 cell grid. These plots were generated using a simple spreading dye model for assemblages with different geographic range sizes. Blue represents the lowest species density and red represents the highest species density. Each graph represents the mean of 300 stochastic runs of the model.

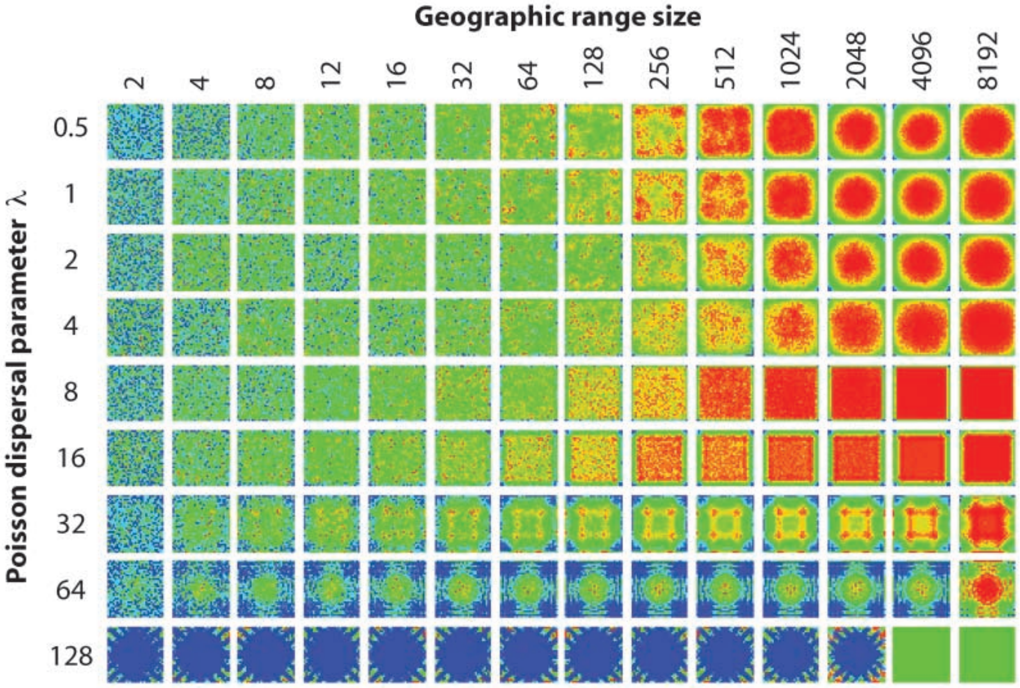

Fig. 4. Contour plots of expected species richness in a stepping-stone model. Each row represents a different value of λ, the Poisson dispersal parameter, and each column represents a different geographic range size. Domain size and colour scale as in Fig. 3. Each graph represents the mean of 300 stochastic runs of the model.

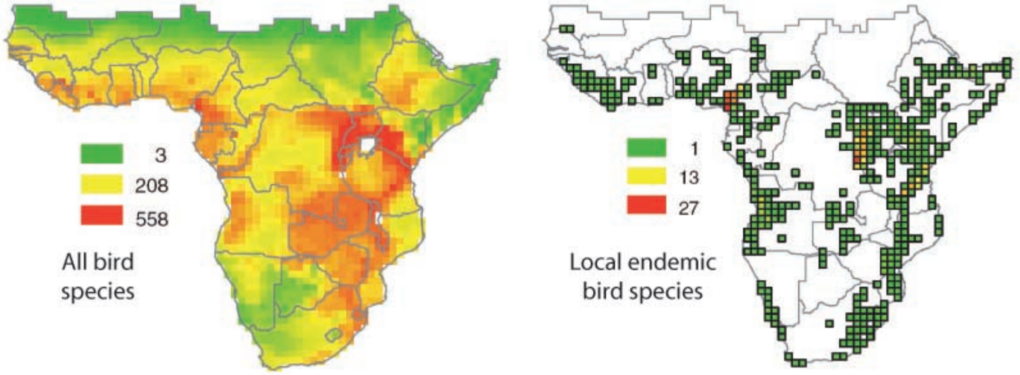

Fig. 5. Species richness for all 1599 species of birds endemic to Africa (left-hand map) and for 190 small-ranged (each occupying 10 or fewer 1 × 1 degree quadrats) local endemic bird species (right-hand map), which tend to occur along the periphery and in mountainous regions (data from Jetz *et al.*, 2004). The concentration of species along the periphery of the domain resembles some of the patterns that emerge from spreading dye simulations of small-ranged species (Figs. 2 and 3).

'wavelet' pattern emerges, and geometrically symmetric hot-spots of species richness form in the domain. With a dispersal distance of 128 grid cells and geographic range sizes of greater than 4096 cells, the pattern again returns to a simple flat plain of species richness. These complex patterns cannot be easily interpreted, and they are probably idiosyncratic to the particular combination of domain size, domain shape, range size, and dispersal distance.

DISCUSSION

Most previous stochastic models of range location in a bounded domain produce a classic mid-domain hump or, for smaller ranges, a plateau of richness [Fig. 1 shows an empirical case; see Figures 2 and 3 in Colwell and Hurtt (1994) for a theoretical example]. But, unlike our simulations here, previous models – with the exception of two of the models of Connolly (in press) – have consistently used heterogeneous range size frequency distributions (RSFDs). It has long been appreciated that large-ranged species dominate the overall pattern of species richness in gradient plots (Pineda and Caswell, 1998; Lees *et al.*, 1999) and maps (Jetz and Rahbek, 2001), 'voting' in proportion to their range size. Any effects of small and intermediate-sized ranges are masked by the strong mid-domain effect (MDE) of the large-ranged species.

In a study that developed independently of ours, Connolly (in press) used an analytical version of the one-dimensional spreading dye model (Equation 2 in Connolly, 2005) and a two-dimensional simulation model that was identical to our model with full range cohesion (Fig. 3) to examine richness patterns for assemblages with constant range size. His results (Connolly, in press, Figure 5B, 5C, and 5D and accompanying text) were identical to ours for the range sizes he explored (10, 40, and 70% of the domain for the one-dimensional model; 10 and 50% of the domain for the two-dimensional model). He points out that the abrupt discontinuities shown by simple, one-dimensional spreading dye models [e.g. Figure 5B in Connolly (in press) and our Fig. 3] are not evident in alternative models in which ranges limits are determined stochastically, with no fixed RSFD.

Understanding the distributional pattern of small-ranged species, which often differs substantially from the pattern for total species richness, requires partitioning ranges into range-size quantiles (e.g. Jetz and Rahbek, 2001) or separating out small-ranged species for special study. Major studies of continental-scale species richness in two-dimensional domains (e.g. Jetz and Rahbek, 2002; Rahbek *et al.*, 2007) and along a variety of more local, one-dimensional gradients (e.g. Cardelús *et al.*, 2006; Dunn *et al.*, 2006) reveal that the effects of environment and geometric constraints on species richness patterns differ strikingly among range-size quantiles.

The complex two-dimensional patterns of species richness that emerge from the spreading dye model for small range sizes (Figs. 3 and 4) may provide some insight into empirical richness maps. The left-hand map in Fig. 5 (based on Jetz *et al.*, 2004) shows the observed pattern of total species richness for sub-Saharan African birds, on a 1° latitude × longitude grid (1599 species, 1738 map grid cells). Completely obscured by the richness patterns of larger-ranged species is the right-hand map, which shows the species richness of the subset of 190 bird species with the very smallest geographic ranges (≤ 10 grid cells).

The concentrations of these small-ranged species (Fig. 5, right-hand map) in interior mountainous areas (Cameroon Highlands, Albertine Rift Mountains, Kenya Highlands, Eastern Zimbabwe mountains, Lesotho Highlands) is obvious, and probably reflects the effects of topography on speciation rates (Graves, 1985; Rahbek and Graves, 2000, 2001; Fjeldså and Rahbek, 2006). But a second pattern is less readily explained: small-ranged species also tend to be concentrated along the coastlines (Fig. 5, right-hand map). This pattern might be accounted for by a combination of coastal mountains, unusual coastal habitats, and narrowly distributed coastal climates. But the pattern is also reminiscent of the peripheral band (doughnut) of richness for small-ranged species that emerges in our two-dimensional null model (Fig. 3). If the doughnut effect contributes to the empirical pattern of African birds, the mechanism would likely be the same as in the spreading dye model: a range that cannot expand past a boundary tends instead to expand laterally, along the edge of the domain, producing a level of range asymmetry near domain boundaries that cannot be accounted for by topography, climate or historical factors. In the model, range size is prescribed; in nature, on the other hand, peripheral ranges that do not expand along the coast but, instead, occupy a 'truncated' range (Sandel and McKone, 2006) of arbitrarily small size may have been subject to higher probability of extinction (Colwell and Hurtt, 1994). Only carefully constructed null models and analyses of environmental and historical factors can uncover the causes of this and other intriguing patterns of small-ranged species that remain hidden under the smokescreen of wider-ranged species. Empirical and theoretical studies would benefit from even further deconstruction of species richness gradients by range size than is currently practised.

The domains illustrated in Fig. 3 can be viewed as classes of patterns representing size-based categories of ranges. Any natural range-size frequency distribution can be approximated by a mixture model composed of a weighted combination of the range sizes in the figure [or perhaps some other series of discrete range size classes (Lees *et al.*, 1999, pp. 541–546)]. In this view of the model, the resulting richness pattern for all range sizes considered together can be envisioned as a weighted sum of the patterns in the figure, or perhaps a series of layers of varying depth, representing the proportions in the mixture. In this way, the artificial quality of single-range-size models can be made realistic, while still preserving the distinct geometric effects that emerge for species of differing ranges.

The two-dimensional stepping-stone model produced an even greater wealth of pattern (Fig. 4), including the subset of patterns that arose from the strict range cohesion of the spreading dye model (Fig. 3). As Connolly (2005) noted, introducing dispersal into a range-based model converts richness from the probability of range overlap to the probability of cell co-occupancy, with increasingly porous ranges weakening the mid-domain pattern (bottom rows in Fig. 4). As dispersal distance increases, a phase of constant species richness throughout the domain reappears, resembling the pattern for small ranges with cohesive ranges, but of course with much higher richness (as illustrated in Fig. 4 by the column with range size = 2048). With very long dispersal distances ($\lambda \geq 32$), strange patterns appear, as most attempts at dispersal point outside the domain, so that trial and error redirects unsuccessful 'destinations' into the left-hand tail of a widening Poisson distribution. After many unsuccessful trials, the algorithm eventually finds a small subset of cells within the domain that meet these constraints and accumulate hot spots of species richness. These latter parameter combinations are not realistic, but they are included for the sake of completeness, curiosity, and aesthetics. No-one ever said that constrained randomness should not be beautiful, as well as instructive.

ACKNOWLEDGEMENTS

We dedicate this paper to Larry Slobodkin, who taught his students to think beyond the constraint of orthodoxy, while learning from the wisdom of the past. '*Classical*', he liked to say, 'means old, and mostly wrong, but worth studying'. We are grateful to Hector Arita, Sean Connolly, Walter Jetz, and Thiago Rangel for their critical comments and useful suggestions. R.K.C. was supported by the US National Science Foundation, DEB-0072702. N.J.G. and G.L.E. acknowledge NSF grants DEB-0107403 and DEB 05-41936 for support of modelling and null model research. G.R.G. was supported by the Alexander Wetmore fund of the Smithsonian Institution and the Danish National Science Foundation (Grant J. 21-03-0221). C.R. acknowledges the Danish National Science Foundation (Grant J. 21-03-0221) for support of macroecological research.

REFERENCES

Arita, H.T. 2005. Range size in mid-domain models of species diversity. *J. Theor. Biol.*, **232**: 119–126.

Arita, H. and Vázquez-Domínguez, E. 2008. The tropics: cradle, museum or casino? A dynamic null model for latitudinal gradients of species diversity. *Ecol. Lett.*, **11**: 653–663.

Bachman, S., Baker, W.J., Brummit, N., Dransfield, J. and Moat, J. 2004. Elevational gradients, area and tropical island diversity: an example from the palms of New Guinea. *Ecography*, **27**: 299–310.

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

Brayard, A., Escarguel, G. and Bucher, H. 2005. Latitudinal gradient of taxonomic richness: combined outcome of temperature and geographic mid-domains effects? *J. Zool. Syst. Evol. Res.*, **43**: 178–188.

Brehm, G., Colwell, R.K. and Kluge, J. 2007. The role of environment and mid-domain effect on moth species richness along a tropical elevational gradient. *Global Ecol. Biogeogr.*, **16**: 205–219.

Brown, J., Stevens, G. and Kaufman, D. 1996. The geographic range: size, shape, boundaries, and internal structure. *Annu. Rev. Ecol. Syst.*, **27**: 597–623.

Cardelús, C.L., Colwell, R.K. and Watkins, J.E. 2006. Vascular epiphyte distribution patterns: explaining the mid-elevation richness peak. *J. Ecol.*, **94**: 144–156.

Colwell, R.K. 2008. RangeModel: tools for exploring and assessing geometric constraints on species richness (the mid-domain effect) along transects. *Ecography*, **31**: 4–7.

Colwell, R.K. and Hurtt, G.C. 1994. Nonbiological gradients in species richness and a spurious Rapoport effect. *Am. Nat.*, **144**: 570–595.

Colwell, R.K. and Lees, D.C. 2000. The mid-domain effect: geometric constraints on the geography of species richness. *Trends Ecol. Evol.*, **15**: 70–76.

Colwell, R.K., Rahbek, C. and Gotelli, N. 2004. The mid-domain effect and species richness patterns: what have we learned so far? *Am. Nat.*, **163**: E1–E23.

Colwell, R.K., Rahbek, C. and Gotelli, N. 2005. The mid-domain effect: there's a baby in the bathwater. *Am. Nat.*, **166**: E149–E154.

Connolly, S.R. 2005. Process-based models of species distributions and the mid-domain effect. *Am. Nat.*, **166**: 1–11.

Connolly, S.R. in press. Macroecological theory and the analysis of species richness gradients. In *Marine Macroecology* (J.D. Witman and K. Roy, eds.). Chicago, IL: University of Chicago Press.

Connolly, S.R., Bellwood, D.R. and Hughes, T.P. 2003. Indo-Pacific biodiversity of coral reefs: deviations from a mid-domain model. *Ecology*, **84**: 2178–2190.

Currie, D.J. and Kerr, J.T. 2007. Testing, as opposed to supporting, the Mid-domain Hypothesis: a response to Lees and Colwell (2007). *Ecol. Lett.*, **10**: E9–E10.

Currie, D.J. and Kerr, J.T. 2008. Tests of the mid-domain hypothesis: a review of the evidence. *Ecol. Monogr.*, **78**: 3–18.

Davies, R.G., Orme, C.D.L., Storch, D., Olson, V.A., Thomas, G.H., Ross, S.G. *et al.* 2007. Topography, energy and the global distribution of bird species richness. *Proc. R. Soc. Lond. B*, **274**: 1189–1197.

Dunn, R.R., Colwell, R.K. and Nilsson, C. 2006. The river domain: why are there more species halfway up the river? *Ecography*, **29**: 251–259.

Dunn, R.R., McCain, C.M. and Sanders, N. 2007. When does diversity fit null model predictions? Scale and range size mediate the mid-domain effect. *Global Ecol. Biogeogr.*, **3**: 305–312.

Fjeldså, J. and Rahbek, C. 2006. Diversification of Tanagers, a species rich bird group, from lowlands to montane regions in South America. *Integr. Comp. Zool.*, **46**: 72–78.

Gotelli, N.J. and Graves, G.R. 1996. *Null Models in Ecology*. Washington, DC: Smithsonian Institution Press.

Gotelli, N.J., Entsminger, G.L., Rahbek, C., Colwell, R.K. and Graves, G.R. 2007. *BioGeoSim. Version 1.0.* Burlongton, VA: Acquired Intelligence Inc. & Kesey-Bear (http://www.garyentsminger.com/biogeosim/biogeosim.htm).

Graves, G.R. 1985. Elevational correlates of speciation and intraspecific geographic variation in plumage in Andean forest birds. *Auk*, **102**: 556–579.

Hawkins, B.A. and Diniz-Filho, J.A.F. 2002. The mid-domain effect cannot explain the diversity gradient of Nearctic birds. *Global Ecol. Biogeogr.*, **11**: 419–426.

Hawkins, B.A., Diniz-Filho, J.A.F. and Weis, A.E. 2005. The mid-domain effect and diversity gradients: is there anything to learn? *Am. Nat.*, **166**: E140–E143.

Jetz, W. and Rahbek, C. 2001. Geometric constraints explain much of the species richness pattern in African birds. *Proc. Natl. Acad. Sci. USA*, **98**: 5661–5666.

Jetz, W. and Rahbek, C. 2002. Geographic range size and determinants of avian species richness. *Science*, **279**: 1548–1551.

Jetz, W., Rahbek, C. and Colwell, R.K. 2004. The coincidence of rarity and richness and the potential signature of history in centres of endemism. *Ecol. Lett.*, **7**: 1180–1191.

Kerr, J.T., Perring, M. and Currie, D.J. 2006. The missing Madagascan mid-domain effect. *Ecol. Lett.*, **9**: 149–159.

Kirkpatrick, M. and Barton, N.H. 1997. Evolution of a species' range. *Am. Nat.*, **150**: 1–23.

Lees, D.C. and 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). *Ecol. Lett.*, **10**: E4–E8.

Lees, D.C., Kremen, C. and Andriamampianina, L. 1999. A null model for species richness gradients: bounded range overlap of butterflies and other rainforest endemics in Madagascar. *Biol. J. Linn. Soc.*, **67**: 529–584.

Lennon, J.J., Koleff, P., Greenwood, J.J.D. and Gaston, K.J. 2004. Contribution of rarity and commonness to patterns of species richness. *Ecol. Lett.*, **7**: 81–87.

Lyons, S.K. and Willig, M.R. 1997. Latitudinal patterns of range size: methodological concerns and empirical evaluations for New World bats and marsupials. *Oikos*, **79**: 568–580.

MacArthur, R.H. 1957. On the relative abundance of bird species. *Proc. Natl. Acad. Sci. USA*, **43**: 293–295.

McCain, C.M. 2003. North American desert rodents: a test of the mid-domain effect in species richness. *J. Mammal.*, **84**: 967–980.

Parmesan, C., Gaines, S., Gonzalez, L., Kaufman, D.M., Kingsolver, J., Peterson, A.T. *et al.* 2005. Empirical perspectives on species borders: from traditional biogeography to global change. *Oikos*, **108**: 58–75.

Pielou, E.C. 1977a. The latitudinal spans of seaweed species and their patterns of overlap. *J. Biogeogr.*, **4**: 299–311.

Pielou, E.C. 1977b. The statistics of biogeographic range maps: sheaves of one-dimensional ranges. *Bull. Int. Stat. Inst.*, **47** (Book 2): 111–122.

Pineda, J. and Caswell, H. 1998. Bathymetric species-diversity patterns and boundary constraints on vertical range distributions. *Deep-Sea Res. II*, **45**: 83–101.

Rahbek, C. 1997. The relationship between area, elevation and regional species richness in Neotropical birds. *Am. Nat.*, **149**: 875–902.

Rahbek, C. and Graves, G.R. 2000. Detection of macro-ecological patterns in South American hummingbirds is affected by spatial scale. *Proc. R. Soc. Lond. B*, **267**: 2259–2265.

Rahbek, C. and Graves, G.R. 2001. Multiscale assessment of patterns of avian species richness. *Proc. Natl. Acad. Sci. USA*, **98**: 4534–4539.

Rahbek, C., Gotelli, N., Colwell, R.K., Entsminger, G.L., Rangel, T.F.L.V.B. and Graves, G.R. 2007. Predicting continental-scale patterns of bird species richness with spatially explicit models. *Proc. R. Soc. Lond. B*, **274**: 165–174.

Rangel, T.F.L.V.B. and Diniz-Filho, J.A.F. 2003. Spatial patterns in species richness and the geometric constraint simulation model: a global analysis of mid-domain effect in Falconiformes. *Acta Oecol.*, **24**: 203–207.

Rangel, T.F.L.V.B. and Diniz-Filho, J.A.F. 2005. An evolutionary tolerance model explaining spatial patterns in species richness under environmental gradients and geometric constraints. *Ecography*, **28**: 253–263.

Rangel, T.F.L.V.B., Diniz-Filho, J.A.F. and Colwell, R.K. 2007. Species richness and evolutionary niche dynamics: a spatial pattern-oriented simulation experiment. *Am. Nat.*, **170**: 602–616.

Romdal, T.S. and Grytnes, J.A. 2007. The indirect area effect on elevational species richness patterns. *Ecography*, **30**: 440–448.

Romdal, T.S., Colwell, R.K. and Rahbek, C. 2005. The influence of band sum area, domain extent, and range sizes on the latitudinal mid-domain effect. *Ecology*, **86**: 235–244.

Sandel, B.S. and McKone, M.J. 2006. Reconsidering null models of diversity: do geometric constraints on species ranges necessarily cause a mid-domain effect? *Divers. Distrib.*, **12**: 467–474.

Sanders, N. 2002. Elevational gradients in ant species richness: area, geometry, and Rapoport's rule. *Ecography*, **25**: 25–32.

Stevens, G.C. 1989. The latitudinal gradient in geographical range: how so many species coexist in the tropics. *Am. Nat.*, **133**: 240–256.

Storch, D., Davies, R.G., Zajicek, S., Orme, C.D.L., Olson, V., Thomas, G.H. *et al.* 2006. Energy, range dynamics and global species richness patterns: reconciling mid-domain effects and environmental determinants of avian diversity. *Ecol. Lett.*, **9**: 1308–1320.

Willig, M.R. and 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., Kaufmann, D.M. and Stevens, R.D. 2003. Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. *Annu. Rev. Ecol. Syst.*, **34**: 273–309.

Zapata, F.A., Gaston, K.J. and Chown, S.L. 2003. Mid-domain models of species richness gradients: assumptions, methods and evidence. *J. Anim. Ecol.y*, **72**: 677–690.

Zapata, F.A., Gaston, K.J. and Chown, S.L. 2005. The mid-domain effect revisited. *Am. Nat.*, **166**: E144–E148.