ECOLOGY LETTERS

Ecology Letters, (2016) **19**: 1009–1022 doi: 10.1111/ele.12640

IDEA AND PERSPECTIVE

Midpoint attractors and species richness: Modelling the interaction between environmental drivers and geometric constraints

Robert K. Colwell,[1,2,3]* Nicholas J. Gotelli,[4] Louise A. Ashton,[5,6] Jan Beck,[3,7] Gunnar Brehm,[8] Tom M. Fayle,[9,10,11] Konrad Fiedler,[12] Matthew L. Forister,[13] Michael Kessler,[14] Roger L. Kitching,[5] Petr Klimes,[9] Jürgen Kluge,[15] John T. Longino,[16] Sarah C. Maunsell,[5] Christy M. McCain,[3,17] Jimmy Moses,[18,19] Sarah Noben,[14] Katerina Sam,[9] Legi Sam,[5,9] Arthur M. Shapiro,[20] Xiangping Wang[21] and Vojtech Novotny[9,18]

Abstract

We introduce a novel framework for conceptualising, quantifying and unifying discordant patterns of species richness along geographical gradients. While not itself explicitly mechanistic, this approach offers a path towards understanding mechanisms. In this study, we focused on the diverse patterns of species richness on mountainsides. We conjectured that elevational range midpoints of species may be drawn towards a single *midpoint attractor* – a unimodal gradient of environmental favourability. The midpoint attractor interacts with geometric constraints imposed by sea level and the mountaintop to produce taxon-specific patterns of species richness. We developed a Bayesian simulation model to estimate the location and strength of the midpoint attractor from species occurrence data sampled along mountainsides. We also constructed *midpoint predictor* models to test whether environmental variables could directly account for the observed patterns of species range midpoints. We challenged these models with 16 elevational data sets, comprising 4500 species of insects, vertebrates and plants. The midpoint predictor models generally failed to predict the pattern of species midpoints. In contrast, the midpoint attractor model closely reproduced empirical spatial patterns of species richness and range midpoints. Gradients of environmental favourability, subject to geometric constraints, may parsimoniously account for elevational and other patterns of species richness.

Keywords

Bayesian model, Biogeography, elevational gradients, geometric constraints, mid-domain effect, midpoint predictor model, stochastic model, truncated niche.

Ecology Letters (2016) **19**: 1009–1022

INTRODUCTION

The search for a unified mechanistic understanding of repeated, global and regional patterns of species richness has long been frustrated by taxonomic and geographical idiosyncrasies, lack of reliable climatic data on appropriate spatial scales, and reliance on case studies built from statistical correlation and *post hoc* conjecture. We offer no cures for these many ills, but instead, propose a novel conceptual approach. While not itself mechanistic, by unifying and

[1]*Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT 06269, USA*
[2]*Departmento de Ecologia, Universidade Federal de Goiás, CP 131, Goiânia, GO 74.001-970, Brasil*
[3]*University of Colorado Museum of Natural History, Boulder, CO 80309, USA*
[4]*Department of Biology, University of Vermont, Burlington, VT 05405, USA*
[5]*Environmental Futures Research Institute, Griffith University, Nathan, Qld 4111, Australia*
[6]*Life Sciences Department, Natural History Museum, South Kensington, London, SW7 5BD, UK*
[7]*Department of Environmental Science (Biogeography), University of Basel, Basel, Switzerland*
[8]*Phyletisches Museum, Friedrich-Schiller Universität, Jena 07743, Germany*
[9]*Institute of Entomology, Biology Centre of the Czech Academy of Sciences and Faculty of Science, University of South Bohemia, Branišovská 31, 370 05 České Budějovice, Czech Republic*
[10]*Forest Ecology and Conservation Group, Imperial College London, Silwood Park Campus, Buckhurst Road, Ascot, Berkshire, SL5 7PY, UK*
[11]*Institute for Tropical Biology and Conservation, Universiti Malaysia Sabah, 88400 Kota Kinabalu, Sabah, Malaysia*
[12]*Department of Botany & Biodiversity Research, University of Vienna, Rennweg 14, 1030 Vienna, Austria*
[13]*Program in Ecology, Evolution, and Conservation Biology, Department of Biology, University of Nevada, Reno, NV 89557, USA*
[14]*Institute of Systematic Botany, University of Zurich, 8008 Zurich, Switzerland*
[15]*Department of Geography, University of Marburg, 35032 Marburg, Germany*
[16]*Department of Biology, University of Utah, Salt Lake City, UT 84112, USA*
[17]*Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, CO 80309, USA*
[18]*New Guinea Binatang Research Center, P.O. Box 604, Madang, Papua New Guinea*
[19]*School of Natural and Physical Sciences, University of Papua New Guinea, P.O. Box 320, National Capital District, Papua New Guinea*
[20]*Center for Population Biology, University of California, Davis, CA 95616, USA*
[21]*College of Forestry, Beijing Forestry University, Beijing 100083, China*
**Correspondence: E-mail: robertkcolwell@gmail.com*

quantifying discordant patterns of species richness, this framework offers a path towards understanding mechanism. We develop and illustrate this approach for terrestrial, elevational gradients. However, the framework and statistical model are quite general, could easily be applied to other habitats, and could be extended from one-dimensional to two- or even three-dimensional spatial domains.

Along any continental latitudinal transect, species richness for most higher taxa peaks in the tropics, where mean annual temperature is the highest and annual variability in temperature is lowest (Wright *et al*. 2009; Chan *et al*. 2016). Regardless of latitude, temperature on most mountainsides declines steadily with elevation, driven by adiabatic cooling, so that the warmest temperatures usually prevail at the bottom of elevational gradients (Ahrens 2013; Fan & van den Dool 2008). Net primary productivity (NPP), although crucially dependent on precipitation, is strongly driven by temperature. Thus, if radiant energy or NPP are fundamentally responsible for the latitudinal richness pattern, as many ecologists suggest (Currie *et al*. 2004; Allen *et al*. 2007), species richness for higher taxa along elevational transects in humid climates should be expected to peak at the lowest elevations.

However, in a review of hundreds of published examples, Rahbek (1995, 2005) showed that species richness usually does *not* peak at the bottom of elevational gradients. For the preponderance (70%) of studies that encompassed complete elevational gradients and accounted for sampling effects, species richness peaked, instead, at intermediate elevations. Declining richness with elevation was the second most-common pattern, but was found in < 20% of studies (Rahbek 2005). Among other things, these meta-analyses imply that, for most terrestrial taxa, local species richness peaks at intermediate tropical elevations, rather than in the tropical lowlands.

Many explanations have been proposed for mid-elevation richness peaks, and surely no single factor is responsible. For some clades, intermediate climatic conditions at these elevations may be more suitable for survival and reproduction: lower elevations may be too hot or too dry (McCain 2007) and higher elevations too cold, too wet or too cloudy (Longino *et al*. 2014). A history of speciation (or more precisely, net diversification) within a clade that is constrained by its heritable environmental tolerances to a specific range of elevations, can lead to a build-up of species at intermediate elevations (Graham *et al*. 2014; Wu *et al*. 2014). In the tropics, a history of mountaintop extinctions during glacial minima and sea level extinctions during glacial maxima could also produce or enhance mid-elevation richness peaks (Colwell & Rangel 2010). Spatially structured dispersal within an elevational domain, such as source-sink dynamics (Grytnes 2003; Grytnes *et al*. 2008) or ecotonal mixing (Lomolino 2001), could also lead to peaks of species richness at intermediate elevations.

Geometric constraints

In addition to these ecological and historical explanations, Colwell & Hurtt (1994) showed, with a simple stochastic model, that a mid-elevation richness peak might be expected even in the absence of climatic drivers or historical forces. In their model, a mid-elevation richness peak arises from the tendency of larger species ranges to overlap more at mid-elevations than at high or low elevations, when they are *geometrically constrained* by the hard boundaries (sea level and the mountaintop) of an elevational domain. Figure 1a offers a physical analogy (a pencil-box) for this phenomenon, which later became known as the *mid-domain effect* (Colwell & Lees 2000) or MDE, because, in a simple 1-dimensional domain, the expected distribution of species richness under this model is exactly symmetrical about the centre of the domain. Geometric constraints have been generalised to other bounded spatial (Storch *et al*. 2006) and non-spatial (Letten *et al*. 2013) domains at the assemblage level, as well as to studies of home ranges (Prevedello *et al*. 2013) and the movement of individuals within a population (Tiwari *et al*. 2005).

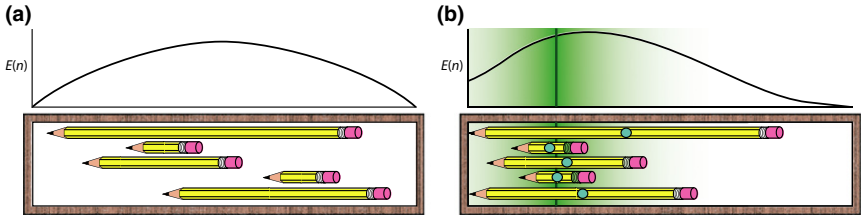


Figure 1 Geometric constraint models. (a) The *classic geometric constraint model* illustrated by a physical analogy: a set of pencils (species), some shorter and some longer (narrower and wider elevational ranges), stored in a schoolchild's old-fashioned pencil-box (the bounded elevational domain) (Colwell *et al*. 2004). If the box is shaken end to end, horizontally, so that the position of each pencil is randomised, the expected number $E(n)$ of pencils that overlap (species richness) near the middle of the box is inevitably greater than the number that overlap nearer the ends of the box, a pattern that is symmetric around the centre of the box. But the constraint does not act uniformly on the pencils as the box is shaken: the shorter pencil stubs move more widely and freely than the longer pencils. By analogy, the distribution of small-ranged species is less constrained by geometry than the distribution of large-ranged species (Colwell & Lees 2000; Dunn *et al*. 2007). (b) A physical analogy for the *midpoint attractor model*. Suppose that each pencil has a steel ball bearing embedded at its midpoint (blue circles). A magnetic field, the attractor, is applied across the pencil box (green). As the box is shaken end to end, the pencils tend to collect near the attractor, as their midpoint ball bearings are drawn towards the magnet. If the attractor is located near one end of the box, as illustrated, the expected number of pencils $E(n)$ that stack up at any location along the length of the pencil box is asymmetric. However, because the midpoints of the longer pencils cannot align with the magnet (since longer pencils abut the end of the box), the peak of $E(n)$ does not coincide with the centre of the attractor. Thus $E(n)$ is influenced jointly by the attractor (the magnet) and the constraint (the limits of the pencil box). The pattern of $E(n)$ is narrow when the attractor is strong, broad when the attractor is weak.

Early studies treated geometric constraints as a stand-alone hypothesis, subject to falsification if it failed to fully explain patterns of richness (see Colwell *et al*. 2004, 2005), or strictly as an alternative hypothesis to environmental explanations (Currie & Kerr 2008). But this either/or perspective misses the point that constraints and drivers do not operate independently, but instead interact. It has proven challenging to integrate geometric constraints with environmental and historical explanations for patterns of species richness. We review the history of these efforts in *Appendix 2*, *Supplemental Introduction*.

A Bayesian midpoint attractor model

The likelihood that several different mechanisms contribute to elevational richness peaks calls for a conceptually and methodologically unifying approach to these patterns at a different level. We introduce the idea that species elevational ranges, which underlie elevational richness gradients, can be treated and modelled as if responding, independently, to a single environmental *attractor* that operates within the geometric constraints of an elevational domain and is specific to a taxon-based assemblage. We develop this approach as a simulation model, apply it to a diverse group of data sets, and then discuss it from the broader perspective of biogeographical gradients.

We take a novel approach to integrating environment with geometric constraints over elevational gradients. Inspired by Wang & Fang's (2012) evidence that large- and small-ranged species respond similarly to environmental drivers and by Rangel & Diniz-Filho's (2005) mechanistic model, we postulate the presence of an underlying unimodal 'favourability' gradient, specific to each elevational transect and to each taxon or functional group.

We modelled the simplest possible pattern of environmental favourability – a unimodal peak – on the simplest possible domain – the unit line. The model is general, but in this study, we assume that the one-dimensional unit domain represents an elevational transect from low elevation (sea level, for all our data sets) to the highest habitable point on a mountain massif. Somewhere along this elevational domain lies a unimodal *midpoint attractor*, specific to the locality and taxon, representing a gradient of 'attraction' for species range midpoints. A continuous function describes the relative strength of the attractor at every point within the domain (Fig. 2).

We model the midpoint attractor as a normal (Gaussian) probability density function $N(A, B)$ with two parameters: its mean location A ($0 < A < 1$) on the unit-line domain, and its standard deviation B ($0 < B < 1$) around the attractor, an inverse measure of attractor strength (Fig. 2). Because the unit domain is bounded at 0 and 1, A and B determine not only the location and shape of the attractor, but also jointly define the upper and lower bounds of the attractor distribution, which is truncated at the domain limits (Fig. 2). To simulate a bounded elevational richness pattern driven by the midpoint attractor, we place the empirical elevational ranges (transformed to unit-line equivalents) on the domain stochastically, sampling their midpoints from the modelled attractor probability density function (which we will henceforth call, simply, *the attractor*). Figure 1b updates the pencil-box analogy for the classic MDE by adding an off-centre attractor for pencil midpoints.

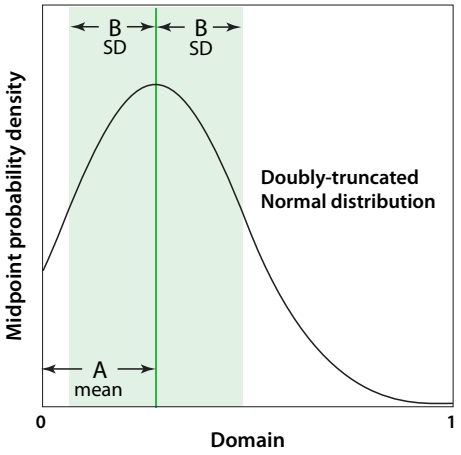


Figure 2 The midpoint attractor modelled with a doubly truncated Gaussian probability density function with mean A and standard deviation B. Parameter A controls the position of the attractor on the gradient. Parameter B controls the strength of the attractor (small B = a strong attractor, large B = a weak attractor).

We developed a Bayesian model to estimate the optimum shape and position of the midpoint attractor for a particular taxon on a particular elevational gradient. The model aims to explain the empirical distribution of species elevational ranges (as indexed by their elevational midpoints), and thus to account for empirical patterns of richness on mountainsides, under geometric constraints. With a centred Gaussian distribution as the starting point, the model employs a simple Gibbs sampler (Gelman *et al*. 2013) to find the posterior distributions of parameter values for the attractor (its location, A, and strength, B), that are most probable (P(model | data)), given the observed elevational pattern of species richness and the empirical *range size frequency distribution* (RSFD).

The midpoint attractor model does not incorporate any environmental data into the estimation of these parameters. It makes no assumptions or *a priori* hypotheses about which environmental or biotic factors might be driving the attractor and the favourability gradient it represents. Instead, once a well-fitting attractor model has been identified using this approach, we subsequently attempt to interpret the attractor statistically in terms of empirically measured environmental variables.

Although the midpoint attractor model maximises P(model | data), most previous attempts to interpret richness patterns have, instead, been conducted in a traditional, frequentist framework, estimating the probability of the data (observed richness), given a specified multivariate statistical model (P(data | model)). The statistical model usually takes the form of a regression of species richness on environmental variables, with (Longino & Colwell 2011) or without (Hawkins *et al*. 2003) a predictor variable for geometric constraints. To compare the results from our Bayesian analyses with these traditional, correlative analyses, we carried out multiple regressions of species richness over elevational gradients, as a

direct function of the same environmental variables that we used to interpret the attractors.

Midpoint predictor models

In addition to the midpoint attractor model, we built two alternative, stochastic, *midpoint predictor models* – one with and one without geometric constraints – that directly assessed environmental variables as predictors of midpoint density (not species richness) over the elevational gradient. In these models, as in the midpoint attractor model, each empirical range midpoint is placed on the domain stochastically. However, range placement is not driven by a hypothetical attractor, as it is in the Bayesian midpoint attractor model. Instead, at each point in the domain, the probability of midpoint placement is directly and linearly proportional to the value of a single, measured, environmental variable (e.g. temperature or precipitation), regardless of the elevational pattern of the variable. Although the midpoint *attractor* model seeks an optimal location and optimal strength for a hypothetical attractor, the midpoint *predictor* models assess the fit of the empirical midpoint data to a probability distribution directly defined by a measured environmental variable. This approach is somewhat akin to the models of Storch *et al.* (2006) and Rahbek *et al.* (2007), but contrasts with the traditional MDE model, in which the probability of midpoint occurrence is constant across the domain.

Application of the models

We applied the midpoint attractor model and the two midpoint predictor models to 16 high-quality data sets that recorded the elevational distribution of more than 4500 species of ferns, insects, mammals or birds in globally distributed localities, mostly in the tropics (Table S1, *Appendix 1*). As we will demonstrate, with or without geometric constraints, the midpoint predictor models generally provide a poor fit to the observed pattern of range midpoints. In contrast, the Bayesian midpoint attractor model simulations consistently produce a good fit to both species richness and midpoint distributions of empirical data sets.

MATERIALS AND METHODS

Empirical data sets and data representation

We applied the midpoint attractor model and the two midpoint predictor models to the 16 data sets detailed in Table S1 (*Appendix 1*). Three groups of data sets included multiple taxa studied on the same gradients: northern Costa Rica, Mt. Wilhelm in Papua New Guinea and the Border Ranges in Australia. To label the individual data sets, we preface the name of the taxonomic group with the name of the geographical location of the gradient (e.g. 'New Guinean ants', 'Costa Rican ferns', etc.). The biogeographical data from these studies consist of species occurrences recorded at a variable number of sampling elevations (5–70 elevations, median = 8) along each gradient. Each data set also included measurements for two or more environmental factors along the gradient (Table S1, *Appendix 1*). We rescaled each elevational domain to the [0, 1] unit line. Within this domain, we standardised sampling points and converted species occurrence records into an estimated elevational range and midpoint for each species, following data preparation protocols detailed in *Supplemental Materials and Methods* (*Appendix 2*). Each data set was represented in two ways: A *midpoint-range plot* (Colwell & Hurtt 1994), with range size as the ordinate and range midpoint as the abscissa for each range in a data set (Fig. 3, *right panel*, grey-scale dots and horizontal line segments), and a corresponding *species richness plot*, showing the number of overlapping ranges at each of a sequence of sampling locations (elevations) spanning the domain (Fig. 3, *left panel*, black dots).

The Bayesian midpoint attractor model

As outlined in the Introduction, we modelled the midpoint attractor as a Gaussian probability density function $N(A, B)$ with two parameters: its mean location A ($0 < A < 1$) on the unit-line domain, and its standard deviation B ($0 < B < 1$) around the attractor (Fig. 2). Because a Gaussian distribution extends from negative to positive infinity, the attractor distribution is truncated at the lower (0) and upper (1) bounds of the domain.

The choice of a unimodal midpoint attractor distribution was based on the empirical prevalence in the published literature of unimodal peaks of species richness (Rahbek 2005), which in turn suggest unimodal midpoint patterns. Our choice of a doubly truncated Gaussian distribution to represent the attractor, rather than a probability distribution that declines to zero at the domain limits (e.g. the beta distribution), was based on biological grounds: many species are regularly present at either sea level or mountaintop, their realised distributions directly abutting a domain limit. Such geographical distributions suggest that these species could readily tolerate more extreme conditions than those at domain limits, on a particular elevational gradient. For example species living at sea level on a mid-latitude elevational gradient might well tolerate even warmer temperatures at a lower latitude. Fundamental niches may fail to be fully expressed for many reasons, but we suggest that elevational domain limits may often impose environmental niche truncation (Colwell & Rangel 2009; Feeley & Silman 2010).

To model the expected pattern of species richness under the influence of the attractor, each of the empirical ranges in a data set is placed on the domain stochastically, without replacement, with its midpoint drawn at random from a proposed attractor distribution $N(A, B)$. To enforce the geometric constraint (Fig. 3, *right panel*) and maintain the empirical RSFD, the midpoint is sampled from this distribution only over the interval of feasible midpoints, given the size of each range, such that the range does not extend beyond either the lower or upper domain limit (Colwell & Lees 2000). For a range of length R, this means that the midpoint must lie in the interval $[R/2, 1 - R/2]$. For these stochastic range simulations, we explored two alternative algorithms for placing ranges within the domain, within this constraint. The two algorithms differ only in how the placement constraint is achieved.

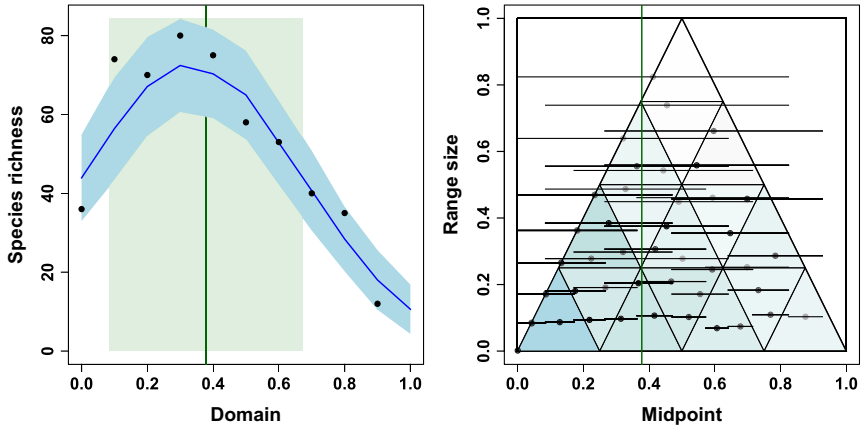


Figure 3 The Bayesian midpoint attractor model applied to the Costa Rican arctiine moth data set (222 species sampled across a 2906 m elevational domain, rescaled to a [0,1] unit line, where 0 represents sea level). *Left panel*: Mean species richness (dark blue line) and 95% confidence interval (light blue band) for 100 simulations. The simulation is driven by a midpoint attractor (dark green vertical line) at 0.378, with a standard deviation (light green rectangle) of 0.294. These parameter values were chosen to maximise the fit of modelled species richness (blue line) to empirical species richness pattern (black dots), using a simple MCMC Gibbs sampler. Empirical range sizes are maintained in the simulation. *Right panel:* Midpoint-range plot for the same data. The *x*-axis is the location of the range midpoint for each species on the elevational domain, and the *y*-axis plots the elevational span of the range (range size). The large triangle sets the geometrically feasible midpoint limits for ranges of a given size. Black and grey points and associated horizontal line segments illustrate the empirical midpoint and range values for the 222 species of moths. Because many species have identical ranges and midpoints in this data set, the shading of each point is proportional to the number of coincident species midpoints. The white-to-blue colour scale in the 16 small triangles is proportional to the mean number of modelled points falling in each triangle, averaged over the 100 runs of the simulation. The correspondence between the number of empirical points (the density of black points and their grey saturation) and the average number of modelled points (blue saturation) among the 16 small triangles is significant at $P < 0.001$ for this data set (Table 1). (See *Materials and Methods* for details of the test.) Note that both empirical and modelled midpoint density is stronger to the left of the attractor than to its right, reflecting the build-up of ranges near sea level, constrained by the domain boundary.

In Algorithm 1, for a species with an empirical range of length R, a midpoint is simply drawn from $N(A, B)$ on the interval $[R/2, 1 - R/2]$ and assigned to the species. Biologically, this algorithm assumes that the elevational distribution of a typical species reaches the limits of its environmental niche within the scope of the gradient, because neither its upper nor its lower range limit is likely to reach a domain limit. This algorithm is the equivalent, for the midpoint attractor model, of the *classic MDE model* of Colwell & Hurtt (1994, their Model 2).

In Algorithm 2, a candidate midpoint is drawn from $N(A, B)$ on the full domain interval [0, 1]. If the candidate midpoint lies within the interval $[R/2, 1 - R/2]$, it is assigned to the species and the next species is considered. If it lies to the left of the interval $[R/2, 1 - R/2]$, then $R/2$ is assigned as the midpoint, whereas if the midpoint lies to the right of the interval $[R/2, 1 - R/2]$, then $1 - R/2$ is assigned. The result is that each such shifted range exactly abuts a domain limit. Because it preserves the empirical RSFD, while allowing a range to reach a domain limit (Connolly 2005), this algorithm is the equivalent, for a one-dimensional domain, of the classic two-dimensional *spreading dye* model of Jetz & Rahbek (2001). Biologically, it captures the idea that the limits of environmental niches of species on ecological gradients are often not fully expressed, so that observed distributions are based on truncated niches. Hence, a better fit to Algorithm 2 than to Algorithm 1 would suggest the prevalence of truncated niches.

By design, these stochastic placement algorithms preserve the empirical RSFD, whereas empirical midpoints are completely ignored. Thus, given the RSFD, correspondence between modelled and empirical patterns of richness, and between empirical and modelled patterns of midpoints, is driven by the location and strength of the attractor.

Just as for empirical richness patterns, the modelled richness at sampling points on the domain is simply the number of stochastically placed ranges that overlap at each sampling point. Because range midpoints are assigned from a statistical distribution (the midpoint attractor), however, each run (realisation) of the midpoint attractor simulation yields a somewhat different pattern of richness over the domain. As illustrated in Fig. 3 (*left panel*), over many runs (e.g. 100), a mean result (dark blue line) and a 95% confidence interval (light blue band) can be defined and plotted to compare with empirical richness (black dots).

In an Approximate Bayesian Computational framework (Marjoram *et al*. 2003; Hartig *et al*. 2011), we used a simple, custom Monte Carlo Markov Chain (MCMC) Gibbs sampler to seek the posterior distribution of model parameters A and B (which, together, fully define the location, shape and truncation points of the Gaussian attractor) that maximised the probability of the model, given the empirical species richness pattern and the empirical RSFD for each data set. In other words, this procedure finds the location and shape of the midpoint attractor that provides the best fit between modelled richness and empirical richness. The details of the ABC and MCMC procedures appear in *Supplemental Materials and Methods* (*Appendix 2*).

In summary, the midpoint attractor model simulates the interaction between a simple, unimodal environmental gradient (the attractor) and the geometric constraints imposed by domain limits. As in the pencil-box analogy (Fig. 1b), because

of the constraint, the distribution of predicted midpoints in the model will not always centre on the attractor. Therefore, the closer the modelled attractor lies to one of the two domain limits, the greater the expected discrepancy between the location of the attractor and the mean location of range midpoints on the domain. Because of this discordance, if the model fitting procedure is successful, we expected that empirical species richness should correlate more strongly with modelled species richness, as simulated by the midpoint attractor model, than with the attractor itself, for communities with off-centre attractors.

Statistical comparison between modelled and empirical midpoint densities

Conceivably, the midpoint attractor model could provide a good fit to the empirical species richness pattern, but fail to produce an underlying pattern of range midpoints within the domain that resembles the corresponding empirical pattern of midpoints: the right answer for the wrong reasons. As an additional, more-detailed assessment of fit, we devised a statistical measure of the correspondence between the modelled and empirical patterns of midpoints and ranges, which we applied to the results of the Bayesian model.

We divided the constraint triangle of the midpoint-range plot evenly into 16 smaller triangles (Fig. 3, *right panel* and Fig. S3, *Appendix 1*) (Laurie & Silander 2002). As a statistic of correspondence between empirical and modelled midpoint density distributions in the 16 sub-triangles, we used the rank of the observed OLS R^2, computed for the 16 sub-triangles, among 999 R^2 values generated by bootstrap resampling. Raw R^2 is inflated by the fact that the total number of points within each of the four rows of smaller triangles (triangle 1, triangles 2–4, 5–9 and 10–16 in Fig. S3) is identical for modelled and empirical distributions. These numbers are identical because the empirical RSFD is used, for each data set, to construct the modelled distribution.

To establish an unbiased sampling distribution, the midpoints within each of the lower three rows of triangles were shuffled at random among the triangles within each row (e.g. among triangles 5–9) and R^2 was computed between the empirical counts and the shuffled counts for all 16 triangles, 999 times. Triangle 1 is constrained to have exactly the same number of points for modelled and empirical data, so no shuffling can be done. The ordinal *P*-value for the modelled vs. empirical R^2 was then based on its rank among the 999 bootstrapped values of R^2.

To assess the prediction (Wang & Fang 2012) that species with small ranges and species with large ranges respond to the same attractor (an assessment not possible with the Bayesian model alone), we repeated the bootstrap procedure separately for larger ranged species (range size > 0.25 of the domain) and for smaller ranged species (range size ≤ 0.25 of the domain).

Mapping midpoint attractors onto environmental variables

The Bayesian model optimises the location and shape of a simple midpoint attractor, without reference to environmental variables measured along each of the gradients. In fact, we know from many sources of evidence that species and species groups respond in complex and often idiosyncratic ways to environmental and elevational gradients (Gotelli *et al.* 2009; Newbery & Lingenfelder 2009; Albert *et al.* 2010; McCain & Grytnes 2010; Presley *et al.* 2011; Sundqvist *et al.* 2011). As typical of most field studies, only limited environmental data were available for the elevational transects in our data sets, and data for different sets of environmental variables were available for different transects.

In an attempt to characterise attractors statistically in terms of underlying available environmental variables, we carried out (linear) multiple regressions, with AIC-based model selection, for each data set on each gradient. At each of a series of evenly spaced elevations, we treated the magnitude of the fitted attractor function as the response variable and the smoothed, interpolated environmental variables as candidate predictor variables. The multiple regression models were fitted using the application Spatial Analysis in Macroecology, version 4.0 (Rangel *et al.* 2010). The data points (elevations) for regression were the same, evenly spaced points across the unit-line domain that were used to fit each midpoint attractor (see *Supplemental Materials and Methods* in *Appendix 2*).

For comparison with traditional correlative approaches applied to explain species richness patterns, we carried out additional multiple regressions, in a model selection framework, with empirical richness as the response variable and environmental variables as candidate predictor variables. We also carried out simple linear regressions with empirical richness as the response variable and the magnitude of the fitted attractor function as the only predictor variable (visualising the results of the Bayesian fitting procedure).

Midpoint predictor models

The midpoint attractor model is, by design, an indirect approach to understanding the drivers of species richness over elevational gradients. As an alternative, direct approach, we designed two explicit *midpoint predictor models*, one with and one without geometric constraints, for the placement of empirical range midpoints within a domain as a direct function of measured environmental variables. For each of the two midpoint predictor models and each of the 16 elevational data sets, we assessed, statistically, the degree of correspondence between the empirical distribution of range midpoints within a domain and the midpoint distribution predicted by a stochastic simulation. In each simulation, range midpoints were placed stochastically on the domain, with the probability of placement at each location directly proportional to the magnitude of a measured environmental variable. In contrast with most other studies, including our midpoint attractor model, the midpoint predictor models consider only the frequency distribution of species midpoints along the elevational gradient, and not the resulting species richness arising from the overlap of species ranges. Details of the two midpoint predictor models and our approach to model evaluation appear in *Supplemental Materials and Methods* (*Appendix 2*).

RESULTS

Midpoint attractors and geometric constraints

Figure 3 shows the empirical data and the fitted midpoint attractor model for the Costa Rican arctiine moth data set. The corresponding graphical results for the other 15 data sets appear in Figs 4 and 5, organised by locality and arranged in the figures to facilitate comparisons among taxonomically and geographically related data sets. We emphasise that the graphs for each data set show the results of 100 simulations using a single pair of values of the midpoint attractor parameters, A and B. These 'best' parameter values were chosen from the Bayesian posterior distribution for the corresponding data set (Fig. S2). For each data set, nearby values of these parameters produce similar graphs. The spreading dye algorithm (Algorithm 2) consistently yielded a fit between modelled and empirical richness that was at least as good, and often better, than the classic approach (Algorithm 1). Consequently, we used the spreading dye algorithm for all data sets in the final models (Table 1).

Table 1 displays the quantitative results for midpoint attractor parameters, and for each, the results for the statistical comparisons between modelled and empirical midpoint density patterns within the geometric constraint triangle (*right panel* for each data set in Figs 3–5). For 14 of the 16 data sets, the test affirms a highly significant (mean $P < 0.002$) correspondence between empirical and modelled midpoint density patterns. The two exceptions (Costa Rican ferns and North American butterflies), instructive in their own right, are discussed below (*Centred midpoint attractors*).

The comparison of modelled and empirical midpoint densities for large-ranged vs. small-ranged species confirmed the expectation that both large and small ranges are equally well fitted by the same midpoint attractor model for most data sets (11 of 16 data sets; Table 1). For a few data sets, a single attractor may not be an appropriate model. Bornean geometrid moths and perhaps North American mammals (Fig. 5) show signs of multimodal attractors, although the fit for a simple, unimodal attractor is nonetheless significant.

The quantitative results in Table 1 demonstrate the key role of geometric constraints in the modelled patterns of richness. As expected (*Materials and Methods*), the closer the modelled attractor is to a domain limit, the greater the discrepancy between the location of the attractor and the mean location of range midpoints on the domain (Fig. 6). In terms of the pencil-box analogy (Fig. 1b), the closer the magnet is set to one end of the box, the further the average pencil midpoint is forced away from the box end.

Moreover, the shift of mean midpoint locations towards mid-domain for ranges on gradients with off-centre attractors (Fig. 6) perhaps reconciles our results with the finding of several previous studies that species richness for small- and large-ranged species is correlated with different environmental factors (e.g. Dunn *et al*. 2006). Instead, the same environmental attractors may act differently on small- and large-ranged species to generate differing distributions. With off-centre attractors, the discordance between attractor and range midpoint increases with range size (e.g. Bornean geometrid moths, Fig. 5). Thus, for larger ranges, patterns of population density or other indicators of performance or fitness may be asymmetrical around the range midpoint, with the performance or fitness peak lying closer to the attractor than to the range midpoint.

The fitted standard deviation of the midpoint attractor (parameter B in the simulations), an inverse measure of the strength of the attractor, varied from 0.023 (strong attractor) for Costa Rican ants to 0.476 (weak attractor) for North American butterflies (Table 1). The location of the midpoint attractor (parameter A) on the unit-line domain ranged from 0.065 for Costa Rican ants, with nearly monotonically declining richness with elevation, to several data sets with A near 0.5 (Costa Rican ferns and geometrid moths, North American butterflies and Australian moths and leaf-miner parasitoids) to 0.742 (Australian leaf miners, on a short, 1100 m gradient). When translated to absolute elevation, A and B vary even more strikingly, because the data sets vary from 1100 m to 4095 m in elevational scope (Table S1, *Appendix 1*).

How well did the model perform in simulating empirical richness? The first two graphs for each data set in Fig. S1 (*Appendix 1*) show: (1) the regression of empirical richness on the magnitude of the modelled midpoint function, and (2) the regression of empirical richness on modelled richness. Table S2 (*Appendix 1*) provides the corresponding statistical results. From these results, we can assess the expectation (*Materials and Methods*), based on the modelled interaction between the attractor and geometric constraints and the fitting method itself, that empirical species richness should correlate more strongly with modelled species richness than with the attractor itself. This expectation was borne out in 12 of the 16 data sets. In all but one of the exceptions, the fit of empirical richness to modelled richness did not differ, by AIC, from the fit of empirical richness to the attractor. In one data set with relatively low species richness (Australian leaf-miner parasitoids), empirical richness was significantly more strongly correlated with the attractor than with modelled richness.

Centred midpoint attractors

When the best-fit attractor lies near the centre of the domain, as it does for Costa Rican ferns and geometrid moths (Fig. 4), North American butterflies (Fig. 5) and Australian moths and leaf-miner parasitoids (Fig. 5) (all with $0.45 < A < 0.55$), the modelled pattern of richness may be quite symmetrical – but so is the expected pattern from a simpler MDE model of geometric constraints with no environmental drivers. For Costa Rican ferns, for example the prediction of the MDE model differs little from the corresponding plot with an optimised midpoint attractor (Fig. 7). The sub-triangle statistical test for the Costa Rican ferns and North American butterfly data sets yields no evidence of an attractor ($P > 0.994$) (Table 1), nor do the tests for large and small ranges for these two data sets ($P > 0.983$). Although the modelled and empirical midpoint densities correspond closely in these two data sets, neither differs from a random distribution of midpoints (given the empirical RSFD), necessarily the baseline for judging the presence of an attractor (*SI Materials and Methods*). Costa Rican geometrid moths show this same result for small-ranged species.

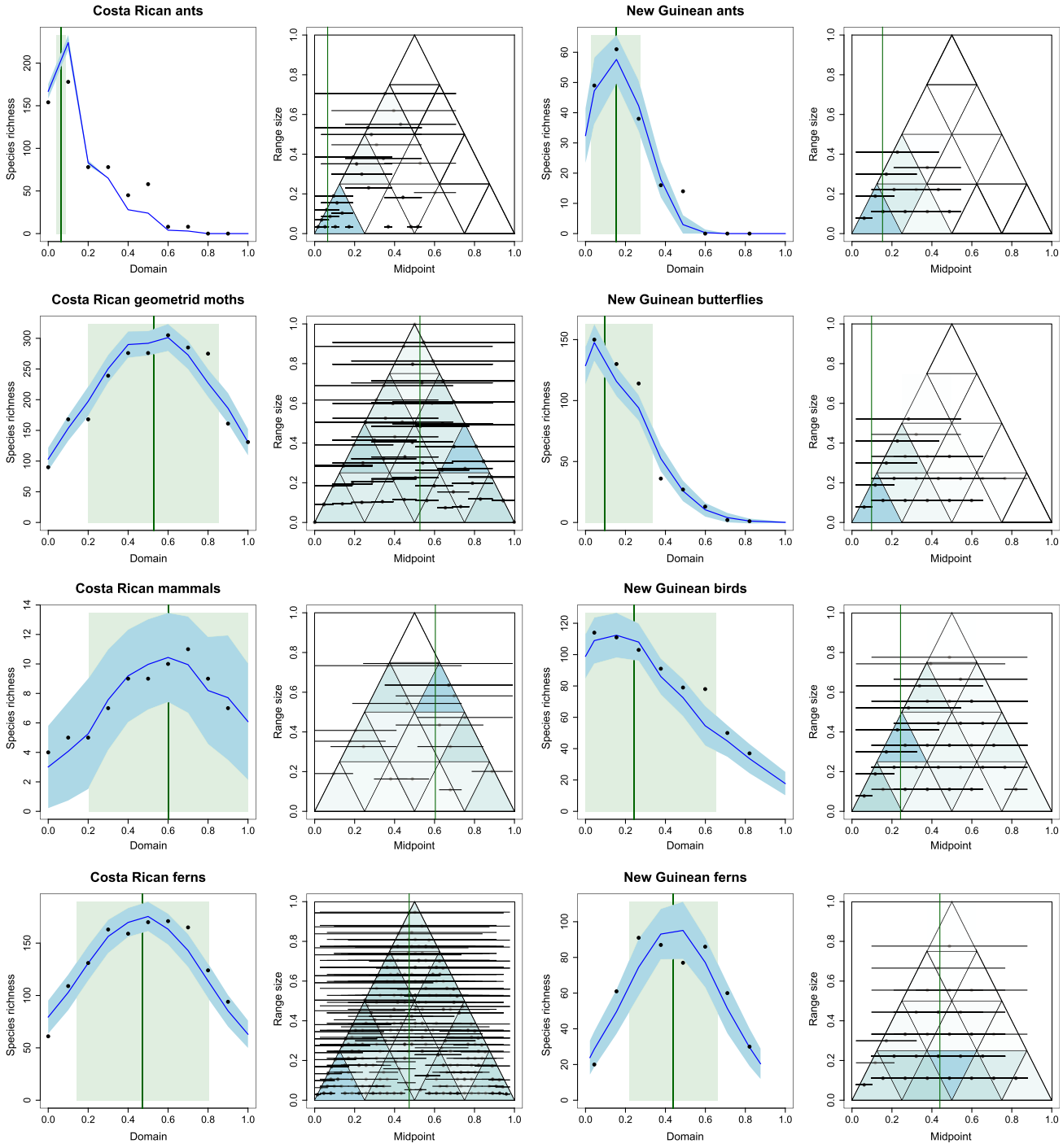


Figure 4 The Bayesian midpoint attractor model applied to four data sets from the same elevational gradient (or, for mammals, a nearby gradient) in Costa Rica (*panel columns 1 and 2*) and four data sets from a single elevational gradient in Papua New Guinea (*panel columns 3 and 4*). The number of empirical points (the density of black points and their grey saturation) and the average number of modelled points (blue saturation) among the 16 small triangles is significant at $P < 0.001$ for seven of the eight data sets (Costa Rican ferns are the exception; see *Centred Midpoint Attractors* in *Results*). A fifth data set from the same Costa Rican gradient appears in Fig. 3, and Fig. 5 shows seven additional data sets. See Fig. 3 for graphical details, Table 1 for statistical results and Table S1 (*Appendix 1*), for details of the data sets.

In such cases, the most conservative conclusion is that we cannot distinguish between pure geometric constraints and a broad (but not too broad) environmental attractor with a peak near the centre of the domain. Although the pure geometric constraints model has two fewer parameters and would thus be favoured in a strict model selection approach, it seems more parsimonious, overall, to adopt a single model of interaction between attractor and constraints for all data sets. Other data sets with attractors closely centred on the domain (e.g. Costa Rican geometrid moths, for large ranges, Fig. 4,

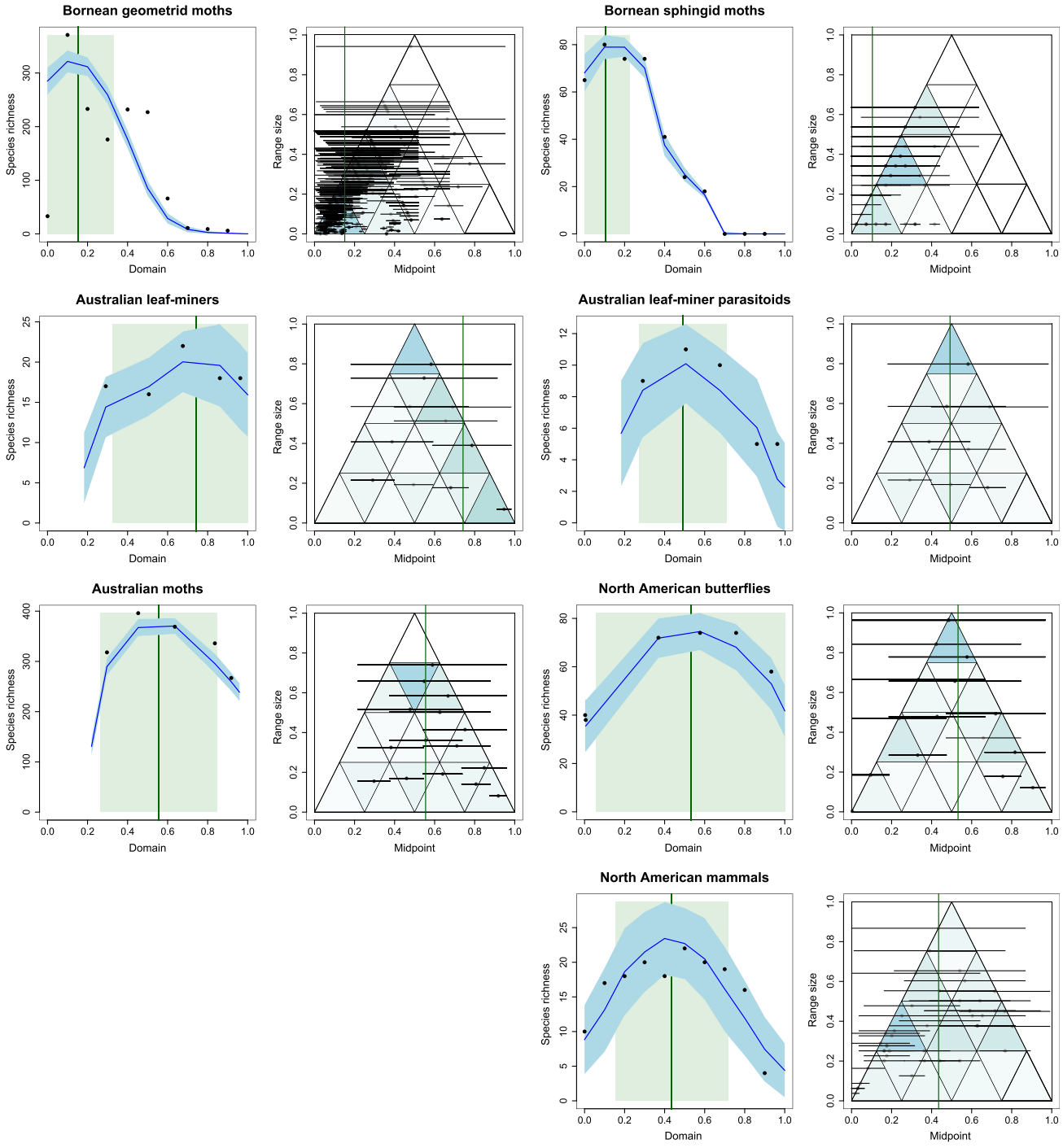


Figure 5 The Bayesian midpoint attractor model applied to seven data sets from Borneo, Australia and North America. See the caption of Fig. 3 for graphical details, Table 1 for statistical results and Table S1 (*Appendix 1*) for details of the data sets.

or Australian moths or leaf-miner parasitoids, Fig. 5) differ from random midpoint locations enough that the test detects the close correspondence between model and data (Table 1).

Mapping midpoint attractors onto environmental variables

Using the results from the midpoint attractor model, the third and fourth graph for each data set in Fig. S1 (*Appendix 1*) present the results for all 16 data sets from the AIC-guided analyses of (1) the regression of the magnitude of the modelled midpoint attractor function on environmental variables, and (2) the regression of empirical richness on environmental variables. In all plots, the points represent the magnitude of the X and Y variables at evenly spaced elevations. Neither axis represents elevation itself. Table S2 (*Appendix 1*) provides the corresponding statistical results and comparisons.

Table 1 Midpoint attractors on a rescaled unit domain (0 = sea level, 1 = mountaintop) for 16 elevational data sets. The fitted attractor mean (*A*) and attractor standard deviation (*B*) define the environmental attractors that drive the modelled species richness patterns in Figs 3–5. Mean midpoint values were computed from the modelled midpoint distributions. Mean range size was computed from the empirical (= modelled) range sizes (RSFD). The statistical correspondence (assessed by R^2) between the midpoint density arising from the midpoint attractor model and empirical midpoint density was tested for significance for all ranges, for large ranges (≥ 0.25 of the unit domain) and for small ranges (< 0.25 of the unit domain) by a bootstrap procedure. Insignificant tests ($P > 0.05$) are reported in boldfaced italics. See Materials and Methods for details.

	Attractor mean (*A*)	Attractor SD (*B*)	Mean midpoint	Mean range	R^2 all ranges	*P* all ranges	R^2 large ranges	*P* large ranges	R^2 small ranges	*P* small ranges
Costa Rica data sets										
Ants	0.065	0.023	0.196	0.181	0.949	0.001	0.968	0.001	0.955	0.001
Arctiine moths	0.378	0.294	0.332	0.228	0.770	0.001	0.747	0.001	0.768	0.001
Geometrid moths	0.527	0.327	0.492	0.306	0.650	0.002	0.762	0.001	***0.155***	***0.999***
Ferns	0.473	0.331	0.479	0.303	***0.466***	***0.999***	***0.617***	***0.997***	***0.368***	***0.999***
Mammals	0.604	0.401	0.521	0.425	0.675	0.001	0.744	0.001	0.007	0.001
Papua New Guinea data sets										
Ants	0.153	0.123	0.199	0.156	0.867	0.001	0.995	0.001	0.832	0.001
Butterflies	0.098	0.239	0.205	0.180	0.954	0.001	0.998	0.001	0.941	0.001
Birds	0.243	0.411	0.330	0.285	0.918	0.001	0.899	0.001	0.968	0.001
Ferns	0.440	0.222	0.447	0.156	0.801	0.001	0.548	0.001	0.246	0.001
Australia border ranges data sets										
Moths	0.555	0.291	0.583	0.419	0.912	0.001	0.939	0.001	0.831	0.001
Leaf miners	0.742	0.418	0.630	0.426	0.555	0.015	0.542	0.013	***0.569***	***0.999***
Leaf-miner parasitoids	0.492	0.219	0.534	0.493	0.551	0.003	0.553	0.003	***0.436***	***0.758***
Borneo data sets										
Geometrid moths	0.151	0.181	0.226	0.173	0.840	0.001	0.626	0.001	0.868	0.001
Sphingid moths	0.104	0.120	0.214	0.342	0.993	0.001	1.000	0.001	0.966	0.001
North America data sets										
Butterflies	0.532	0.476	0.502	0.486	***0.632***	***0.996***	***0.635***	***0.997***	***0.448***	***0.983***
Mammals	0.435	0.282	0.381	0.353	0.435	0.001	0.371	0.009	0.724	0.001

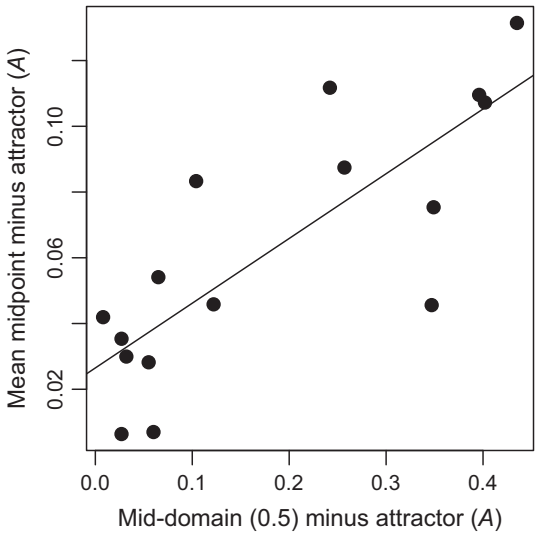


Figure 6 The signature of geometric constraints in the modelled patterns of species richness. The closer the modelled attractor lies to a domain limit, the greater the discrepancy between the location of the attractor and the mean location of range midpoints on the domain. The graph shows the relationship between |(*mean midpoint* − *attractor*)| and |(*0.5* − *attractor*)|. Each point represents a different data set ($n = 16$, *slope* = 0.592, $P < 0002$). See Table 1 for attractor mean (parameter *A*) and mean midpoint values.

The environmental variables that best explained the modelled midpoint attractor often differed from the environmental variables that best predicted observed species richness. Only three of the 16 data sets yielded an identical statistical model (or model group, when ΔAIC was < 3 between alternatives), in terms of the predictor variables included, for attractor and for species richness. However, the model with the lowest absolute AIC matched in 10 of the 16 data sets, if ΔAIC-grouped models were ignored (Table S2, illustrated in Fig. S1, *Appendix 1*).

Midpoint predictor models

For each data set, the same environmental variables assessed in interpreting midpoint attractors (Table S2 and Fig. S1, *Appendix 1*) were tested for the two midpoint predictor models, one with and the other without geometric constraints. In these models, an environmental variable determined the stochastic placement of range midpoints at locations across the domain. Across all data sets, 98 of 112 statistical tests (two models, 56 data set-variable combinations) strongly rejected the null hypothesis that modelled midpoints resemble the empirical ones, with $P < 0.001$ in nearly every case (Table S3, *Appendix 1*). Only four of the 16 data sets showed an acceptable fit ($P > 0.05$) to either of the midpoint predictor models. But these data sets were, not coincidentally, the four smallest, in terms of number of species (Australian leaf miners and their parasitoids, Costa Rican and North American mammals; Table S1, *Appendix 1*), and thus had the weakest statistical power to reject the null hypothesis.

DISCUSSION

By modelling and quantifying repeated underlying structures, which we call *attractors*, we offer a fresh approach to

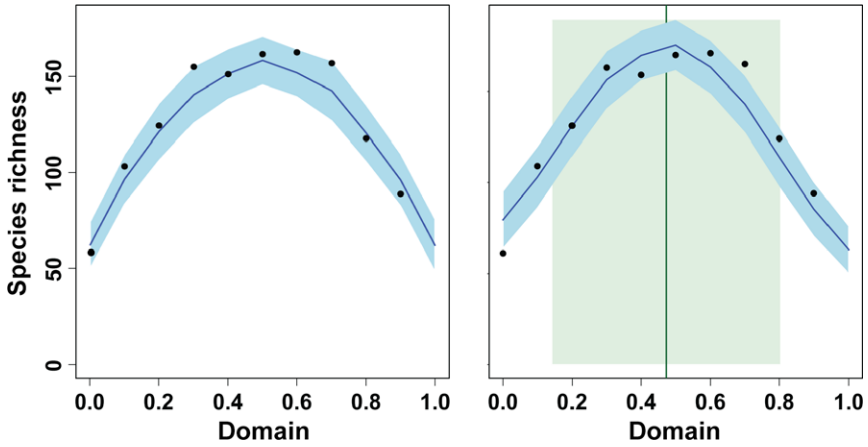


Figure 7 Costa Rican fern data set with no attractor (pure geometric constraints) (*left panel*) and with the best-fit midpoint attractor (*right panel*). The modelled curves differ slightly in shape, but the overall fit is quite similar. Empirical richness values are the black points, identical in the two plots.

characterising biogeographical patterns of species richness. In a broader perspective, we view the attractor not as a mechanism in itself, but as a unifying, intermediate link or layer, between data and mechanism. With two simple parameters, the Bayesian midpoint attractor model characterised the shape and location of a broad spectrum of species richness patterns for a wide variety of taxa along mountain slopes. Just as in other realms of biology, modelling function requires modelling constraints. In our model, these constraints were imposed by the domain boundaries and the range size frequency distribution. By characterising an underlying gradient of favourability, our midpoint attractor model offers a unifying approach to elevational richness gradients that has not been achieved by traditional, *ad hoc* interpretations of correlations of species richness with environmental variables (Gotelli *et al*. 2009).

The stochastic simulations used in the midpoint attractor model are simply a means of producing statistical distributions of richness and range location, under specified conditions, for comparison with corresponding empirical patterns. The algorithms used in these simulations are not, in themselves, intended to represent mechanistic processes in any literal sense. Even the notion of a midpoint attractor, representing a gradient of environmental 'favourability', is a conceptual stand-in for the unspecified biological mechanisms that ultimately lead to concentrations of ranges in certain regions of the domain: adaptive range shifts, diversification of lineages, differential extinction and other forces (Colwell & Rangel 2010). The pencilbox analogy (Fig. 1), likewise, is intended to demonstrate patterns, not processes.

Although the elevational richness patterns successfully modelled in this study varied widely in shape and location on the domain, the midpoint attractor model successfully reproduced not only taxon-specific peaks of species richness but also their underlying empirical midpoint distributions (Figs 3–5). The strong signature of geometric constraints in these results (Fig. 6) shows that the midpoint attractor, alone, is not responsible for the excellent fit of model to data. Instead, the seamless integration of attractor and constraints allows the model to generate patterns ranging from nearly monotonic declines of species richness to perfectly symmetric mid-elevation humps.

It might be tempting to dismiss these results as a needlessly elaborate method of descriptive curve-fitting. A polynomial regression, for example would do an excellent job (in fact, a perfect job, with enough parameters) of fitting any of the distributions of sampling point empirical richness in Figs 3–5. Of course, the 16 data sets would vary in the number of parameters required (assuming model selection was applied), yielding a large table of fitted parameters for the data sets, many or most of them uninterpretable. Instead, the midpoint attractor efficiently unifies all data sets with the same two parameters of the attractor for every data set. Moreover, the sub-triangle statistical analysis of the joint distribution of midpoints and range sizes provides an additional, more-detailed assessment of fit, because it relies on the underlying empirical range and midpoint data. In contrast, using a polynomial regression (or direct fitting of a statistical distribution, such as the gamma distribution) to describe observed richness ignores the underlying data: the elevational ranges of the individual species. Furthermore, these methods provide no avenue to explore or document the role of geometric constraints (Fig. 6).

Constructing the midpoint attractor model in a Bayesian framework was not a matter of convenience, interpretation or fashion, but rather a logical necessity. Given the conjecture that a taxon-specific, location-specific, underlying gradient of favourability, interacting with geometric constraints, could explain elevational richness patterns, the appropriate way forward was to maximise the probability of a general, underlying model, challenged with a plethora of contrasting data sets – a fundamentally Bayesian approach.

Data sets with many ranges abutting the low-elevation domain limit (e.g. Costa Rican ants, Fig. 4, and Bornean geometrid and sphingid moths, Fig. 5) or the high-elevation domain limit (Australian leaf-miner parasitoids and North American butterflies, Fig. 5) strongly suggest an unexpressed potential for some species to prosper in environmental conditions more extreme than conditions at the lower or upper domain limit. In other words, range limits in geographical space, forced by the domain boundaries (e.g. sea level or

mountaintop), may not coincide with niche limits in niche space for such species (Colwell & Rangel 2010). The excellent performance of the doubly truncated Gaussian attractor and our finding that Algorithm 2 (spreading dye) provided a better fit than Algorithm 1 (classic), considered together, offer support for the inference that ranges that abut domain boundaries represent niches truncated by the limits of elevational gradients. In contrast, a species range that reaches neither of the domain limits on the gradient may – or may not – fully express the species' fundamental niche.

Shuffling the observed ranges (the RSFD) within a bounded domain, with or without a midpoint attractor, assumes that the RSFD is *representative* of the size distribution of elevational ranges for a particular taxon on a particular elevational gradient at the particular time that the data were taken (Colwell *et al.* 2004). Given that ranges are drawn without replacement from the RSFD and placed randomly on the domain (within geometric constraints), whereas observed ranges tend to be truncated at domain boundaries, the question then arises: does the midpoint attractor model produce a deficit of small ranges near the domain boundary and an excess of small ranges in mid-domain? If there were such an effect, we would expect it to be stronger for attractors located nearer a domain limit. We tested for this bias by comparing empirical to modelled midpoint density in sub-triangles 10 (near sea level), and 7 (mid-domain) (Fig. S3, *Appendix 1*), as a function of attractor location (*A*), for the 16 data sets. We found no evidence of any pattern of deficiency or excess in modelled midpoint density. If there is any bias, it is slight enough to be completely masked by the heterogeneous size and placement of ranges, both empirical and modelled. These results are consistent with the models of (Colwell & Hurtt 1994), who simulated range truncation for a classical MDE model, and found very little decrease in mean range size as the domain boundary was approached.

With or without geometric constraints, the midpoint predictor models, which assessed empirical environmental factors as candidate midpoint predictors, fitted observed elevational midpoint distributions very poorly (Table S3, *Appendix 1*), despite incorporating the empirical RSFD (in one variant) and having two free parameters, just like the midpoint attractor model. For the data sets in this study, the seemingly intuitive hypothesis that environmental conditions should predict the location of species range midpoints failed to account for most observed midpoint patterns. In contrast, the midpoint attractor model, which, by design, ignores environmental variables, yielded midpoint distributions very close to the empirical midpoint distributions. How can we reconcile this failure of the midpoint predictor model with the success of the midpoint attractor model? At least three, non-exclusive explanations are possible: (1) We may have used the 'wrong' environmental variables in the midpoint predictor models. Although the midpoint attractors, together with geometric constraints, produced a good fit to empirical species richness, the fit of the attractors themselves to environmental variables was often rather poor (Table S3; third panel in each graph in Fig. S1). The original investigators for our data sets measured important aspects of temperature, precipitation and other variables (such as plant cover) that are believed to affect species richness on elevational gradients. Primary productivity is thought to be a key correlate of species richness for many groups (Storch *et al.* 2006). However, primary productivity is difficult to measure directly, is difficult to estimate accurately on small spatial scales from remotely sensed data, and is missing from all our data sets. (2) We might have analysed the right variables, but we had the wrong functional form. In preliminary analyses, however, alternative functional forms (e.g. logarithmic, exponential) did not improve the fit. For many of our data sets, such as Borneo geometrid moths and New Guinea butterflies, the high concentration of species range midpoints in the lower elevations of the domain cannot be accounted for by any univariate or multivariate transformation of the available environmental variables. (3) Lineage diversification with strong niche conservatism may have produced spatial concentrations of range midpoints in narrowly defined environments – a sort of theme-and-variations. Concentrations of elevational range midpoints may arise from 'colonisation' of new environments (e.g. transitions from lowland to montane specialists) followed by net diversification, with little divergence in environmental tolerances (e.g. Graham *et al.* 2014; Wu *et al.* 2014). A search for multimodal attractors and alignment with phylogenetic structure would be a fruitful area of future research.

Like *niche*, or *community*, or *ecosystem*, the idea of an environmental *attractor* reifies an abstract construct. Such constructs endure only if they prove adaptable and useful. In this study, we began with the idea of an attractor, treating it in a Bayesian framework as a model to be challenged by elevational data. But the idea of a range attractor model need not be limited to one-dimensional gradients, nor to terrestrial environments. The location and shape of midpoint attractors within a particular domain arise from the interactions between taxon, climate and history. Comparative study of the relative influence of these factors can be made rigorous and quantitative by fitting attractors to multiple data sets, as we have done in this study. The environmental and historical factors defining midpoint attractors in nature are likely to be complex, presenting a challenge for future research. But our approach, in which a modelled midpoint attractor drives the location of species ranges placed stochastically within a bounded domain, may prove more fruitful than further attempts to correlate patterns of species richness along bounded gradients with environmental factors.

ACKNOWLEDGEMENTS

We are grateful to the Editors and to four dedicated, anonymous reviewers for their incisive and challenging questions and suggestions, which greatly improved this manuscript. We thank Paul Lewis for consultation on Bayesian methods. This study had its origins at a workshop in Ceske Budejovice, Czech Republic, in August, 2013, organised by V. Novotny and funded by the Czech Ministry of Education and the European Social Fund (CZ.1.07/2.3.00/20.0064). Author support: CAPES Ciência sem Fronteiras (Brazil) (RKC); U. S. NSF DEB 1257625, DEB 1144055 and DEB 1136644 (NJG); NSF DEB 1354739 (Project ADMAC) (JTL); NSF DEB 841885 (JM and VN); Griffith University and Australian

Postgraduate Research Awards (LAA and SCM); Australian Research Council DP140101541 (RLK and TMF) and Yayasan Sime Darby (TMF); German DFG Br 2280/1-1, Fi547/5-1 and FOR 402/1-1 (GB and KF); DFG and German Academic Exchange Service DAAD (JK); DFG, Swiss National Fund, & Claraz Schenkung (MK and SN); Czech Science Foundation 14-36098G, 14-32302S & 14-32024P (TMF, PK and KS), 13-10486S (JM and VN); National Natural Science Foundation of China 31370620 (XW); UK Darwin Initiative 19-008 (JM and VN); and IBISCA (TMF, KS and LS).

AUTHORSHIP

RKC and NJG conceived and implemented the models, carried out the analyses, and drafted the manuscript. RKC prepared the figures. LAA, JB, GB, TMF, MLF, KF, MK, RLK, PK, JK, JTL, SCM, CMM, JM, SN, KS, LS and AMS collected or provided data. All authors contributed substantially to the development of ideas, the interpretation of results and revision of the manuscript.

REFERENCES

Ahrens, C.D. (2013). *Meteorology today*, (11th edn.). Brooks/Cole Publishing, Belmont CA, USA.

Albert, C.H., Thuiller, W., Yoccoz, N.G., Soudant, A., Boucher, F., Saccone, P. *et al*. (2010). Intraspecific functional variability: extent, structure and sources of variation. *J. Ecol*., 98, 604–613.

Allen, A.P., Gillooly, J.F. & Brown, J.H. (2007). Recasting the species-energy hypothesis: the different roles of kinetic and potential energy in regulating biodiversity. In: *Scaling Biodiversity* (eds. Storch, D, Marquet, PA & Brown, JH). Cambridge University Press Cambridge, UK, pp. 283–299.

Chan, W.-P., Chen, I.-C., Colwell, R.K., Liu, W.-C., Huang, C.-Y. & Shen, S.-F. (2016). Seasonal and daily climate variation have opposite effects on species elevational range size. *Science*, 351, 1437–1439.

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

Colwell, R.K. & 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. & Rangel, T.F. (2009). Hutchinson's duality: the once and future niche. *PNAS*, 106, 19651–19658.

Colwell, R.K. & Rangel, T.F. (2010). A stochastic, evolutionary model for range shifts and richness on tropical elevational gradients under Quaternary glacial cycles. *Philos. Trans. R. Soc. Lond. B Biol. Sci*., 365, 3695–3707.

Colwell, R.K., Rahbek, C. & 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. & 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 middomain effect. *Am. Nat*., 166, 1–11.

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

Currie, D., Mittelbach, G., Cornell, H., Field, R., Guegan, J., Hawkins, B. *et al*. (2004). Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecol. Lett*., 7, 1121–1134.

Dunn, R.R., Colwell, R.K. & 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. & Sanders, N. (2007). When does diversity fit null model predictions? Scale and range size mediate the mid-domain effect. *Glob. Ecol. Biogeogr*., 3, 305–312.

Fan, Y. & van den Dool, H. (2008). A global monthly land surface air temperature analysis for 1948-present. *J. Geophys. Res*., 113, D01103.

Feeley, K.J. & Silman, M.R. (2010). Biotic attrition from tropical forests correcting for truncated temperature niches. *Global Change Biol*., 16, 1830–1836.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. & Rubin, D.B. (2013). *Bayesian Data Analysis*. CRC Press, Boca Raton, FL, USA.

Gotelli, N., Anderson, M.J., Arita, H.T., Chao, A., Colwell, R.K., Connolly, S.R. *et al*. (2009). Patterns and causes of species richness: a general simulation model for macroecology. *Ecol. Lett*., 12, 873–886.

Graham, C.H., Carnaval, A.C., Cadena, C.D., Zamudio, K.R., Roberts, T.E., Parra, J.L. *et al*. (2014). The origin and maintenance of montane diversity: integrating evolutionary and ecological processes. *Ecography*, 37, 711–719.

Grytnes, J. (2003). Ecological interpretations of the mid-domain effect. *Ecol. Lett*., 6, 883–888.

Grytnes, J.A., Beaman, J.H., Romdal, T.S. & Rahbek, C. (2008). The mid-domain effect matters: simulation analyses of range-size distribution data from Mount Kinabalu. *Borneo. J. Biogeogr*., 35, 2138–2147.

Hartig, F., Calabrese, J.M., Reineking, B., Wiegand, T. & Huth, A. (2011). Statistical inference for stochastic simulation models – theory and application. *Ecol. Lett*., 14, 816–827.

Hawkins, B.A., Field, R., Cornell, H.V., Currie, D.J., Guégan, J.-F., Kaufmann, D.M. *et al*. (2003). Energy, water, and broad-scale geographic patterns of species richness. *Ecology*, 84, 3105–3117.

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

Laurie, H. & Silander Jr., J.A. (2002). Geometric constraints and spatial pattern of species richness: critique of range-based null models. *Divers. Distrib*., 8, 351–364.

Letten, A.D., Kathleen Lyons, S. & Moles, A.T. (2013). The mid-domain effect: it's not just about space. *J. Biogeogr*., 40, 2017–2019.

Lomolino, M.V. (2001). Elevational gradients of species density: historical and prospective views. *Global Ecol. Biogeogr*., 10, 3–13.

Longino, J.T. & Colwell, R.K. (2011). Density compensation, species composition, and richness of ants on a neotropical elevational gradient. *Ecosphere*, 2(3), art. 29. doi:10.1890/ES10-00200.1.

Longino, J.T., Branstetter, M.G. & Colwell, R.K. (2014). How ants drop out: ant abundance on tropical mountains. *PLoS ONE*, 9, e104030.

Marjoram, P., Molitor, J., Plagnol, V. & Tavaré, S. (2003). Markov chain Monte Carlo without likelihoods. *Proc. Natl. Acad. Sci. U. S. A*., 100, 15324.

McCain, C.M. (2007). Could temperature and water availability drive elevational species richness patterns? A global case study for bats. *Glob. Ecol. Biogeogr*., 16, 1–13.

McCain, C.M. & Grytnes, J.A. (2010). Elevational gradients in species richness. In: *Encyclopedia of Life Sciences (ELS)* (ed. Janson, R.). John Wiley & Sons, Ltd, Chichester, UK.

Newbery, D. & Lingenfelder, M. (2009). Plurality of tree species responses to drought perturbation in Bornean tropical rain forest. *Forest Ecology* 201, 147–167.

Presley, S.J., Willig, M.R., Bloch, C.P., Castro-Arellano, I., Higgins, C.L. & Klingbeil, B.T. (2011). A complex metacommunity structure for gastropods along an elevational gradient. *Biotropica*, 43, 480–488.

Prevedello, J.A., Figueiredo, M.S., Grelle, C.E. & Vieira, M.V. (2013). Rethinking edge effects: the unaccounted role of geometric constraints. *Ecography*, 36, 287–299.

Rahbek, C. (1995). The elevational gradient of species richness: a uniform pattern? *Ecography*, 19, 200–205.

Rahbek, C. (2005). The role of spatial scale in the perception of large-scale species-richness patterns. *Ecol. Lett*., 8, 224–239.

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

Rangel, T.F.L.V.B. & 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., Diniz Filho, J.A.F. & Bini, L.M. (2010). SAM: a comprehensive application for Spatial Analysis in Macroecology. *Ecography*, 33, 46–50.

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.

Sundqvist, M.K., Giesler, R., Graae, B.J., Wallander, H., Fogelberg, E. & Wardle, D.A. (2011). Interactive effects of vegetation type and elevation on aboveground and belowground properties in a subarctic tundra. *Oikos*, 120, 128–142.

Tiwari, M., Bjorndal, K.A., Bolten, A.B. & Bolker, B.M. (2005). Intraspecific application of the mid-domain effect model: spatial and temporal nest distributions of green turtles, *Chelonia mydas*, at Tortuguero, Costa Rica. *Ecol. Lett.*, 8, 918–924.

Wang, X. & Fang, J. (2012). Constraining null models with environmental gradients: a new method for evaluating the effects of environmental factors and geometric constraints on geographic diversity patterns. *Ecography*, 35, 1147–1159.

Wright, S.J., Muller-Landau, H. & Schipper, J. (2009). The future of tropical species on a warmer planet. *Conserv. Biol.*, 23, 1418–1426.

Wu, Y., Colwell, R.K., Han, N., Zhang, R., Wang, W., Quan, Q. *et al.* (2014). Understanding historical and current patterns of species richness of babblers along a 5000-m subtropical elevational gradient. *Global Ecol. Biogeogr.*, 23, 1167–1176.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

Editor, Nathan Swenson
Manuscript received 20 January 2016
First decision made 28 February 2016
Manuscript accepted 18 May 2016

Midpoint attractors and species richness: Modeling the interaction between environmental drivers and geometric constraints

APPENDICES

APPENDIX 1: SUPPLEMENTAL TABLES AND FIGURES

APPENDIX 2: SUPPLEMENTAL TEXT

REFERENCES FOR THE APPENDICES

APPENDIX 1: SUPPLEMENTAL TABLES AND FIGURES

SUPPLEMENTAL TABLES S1–S3

Table S1. The datasets and their characteristics. *Sampling limits* represent the lowest and highest occurrence on a unit-line transect, after range adjustments described in the *Supplemental Materials and Methods* (Dataset Selection and Preparation). *Sampling scope* is the difference between the sampling limits.

Dataset	Locality	Transect coordinates	Sampling elevations	Sampling limits	Sampling scope	Domain limits (m asl)	Sampling method	Species	Environmental variables and their units	Data provider	Collection dates	Published references to the dataset
Costa Rica Datasets												
Ants	Barva Transect (Prov. Heredia)	10°08'N–10°26'N, 84°00'W–84°07'W	7	0.004, 0.705	0.701	0, 2900	Miniwinkler extractors	332	MAT (°C), Mean RH (%), MAP (mm), Area (% of total per 100m band)	John T. Longino	2001-2007	(Longino & Colwell 2011; Longino *et al.* 2014)
Arctiine moths	Barva Transect (Prov. Heredia)	10°08'N–10°26'N, 84°00'W–84°07'W	12	0.013, 0.940	0.927	0, 2900	Light traps, manual collection	222	MAT (°C), Mean RH (%), MAP (mm), Area (% of total per 100m band)	Gunnar Brehm	2003–2004	None
Geometrid moths	Barva Transect (Prov. Heredia)	10°08'N–10°26'N, 84°00'W–84°07'W	12	0.013, 0.940	0.927	0, 2900	Light traps, manual collection	739	MAT (°C), Mean RH (%), MAP (mm), Area (% of total per 100m band)	Gunnar Brehm	2003–2004	(Brehm *et al.* 2007)

Dataset	Locality	Transect coordinates	Sampling elevations	Sampling limits	Sampling scope	Domain limits (m asl)	Sampling method	Species	Environmental variables and their units	Data provider	Collection dates	Published references to the dataset
Ferns	Barva Transect (Prov. Heredia)	10°08'N–10°26'N, 84°00'W–84°07'W	29	0.011, 0.986	0.975	0, 2900	Plot-based (20x20m²)	434	MAT (°C), Mean RH (%), MAP (mm), Area (% of total per 100m band)	Jürgen Kluge	2002-2003	(Kluge *et al.* 2006)
Mammals	Tilarán Mt. Range	10°23'N–10°17'N, 84°47'W–84°26'W	18	0.000, 0.998	0.989	0, 1840	Live traps, kill traps, and pitfall traps	18	Average Temperature (°C), Annual Precipitation (mm) [worldclim], elevational area (km2 per 100m elevational band) [DEM, ArcGIS]	Christy McCain	2000-2002	(McCain 2004; McCain 2005)
Papua New Guinea Datasets												
Ants	Mt. Wilhelm Transect	5°44'S–5°47'S, 145°03'E–145°20'E	8	0.007, 0.822	0.815	0, 4509	Pitfall trapping and hand-collecting Yusah et al.(2012)	116	MAT (°C), Mean RH (%), MAP (mm), Area (% of total per 100m band)	Jimmy Moses, Tom M. Fayle, Petr Klimes	2012	(Moses 2015)
Butterflies	Mt. Wilhelm Transect	5°44'S–5°47'S, 145°03'E–145°20'E	8	0.022, 0.876	0.854	0, 4509	Modified Pollard transects (Caldas & Robbins 2003)	264	MAT (°C), Mean RH (%), MAP (mm), Area (% of total per 100m band)	Legi Sam	2009	(Sam 2011)

Dataset	Locality	Transect coordinates	Sampling elevations	Sampling limits	Sampling scope	Domain limits (m asl)	Sampling method	Species	Environmental variables and their units	Data provider	Collection dates	Published references to the dataset
Birds	Mt. Wilhelm Transect	5°44'S–5°47'S, 145°03'E–145°20'E	8	0.022, 0.876	0.854	0, 4509	Point-counts, mist-netting	245	MAT (°C), Mean RH (%), Mean Tree Height (m), Mean Tree Basal Area (cm2)	Katerina Sam	2010-2012	(Tvardikova 2013; Sam & Koane 2014)
Ferns	Mt. Wilhelm Transect	5°44'S–5°47vS, 145°03'E–145°20'E	8	0.022, 0.876	0.854	0, 4509	Plot-based (20x20m²)	359	MAT (°C), Mean RH (%), MAP (mm), Area (% of total per 100m band)	D.N. Karger, S. Noben, M. Lehnert, M.S. Sundue	2014	None
Australia Datasets												
Moths (macromoths + Pyraloidea)	Border Ranges (NSW)	28°24'S–28°22'S, 153°1'E–153°5'E	5	0.220, 0.959	0.739	0, 1100	Light traps	612	°C min, max median, average plant richness	Louise Ashton, Roger Kitching	2009-2010	(Ashton *et al.* 2015)
Leaf miners (Lepidoptera, Coleoptera, Diptera, Hymenoptera)	Border Ranges (NSW)	28°24'S–28°22'S, 153°1'E–153°5'E	5	0.183, 0.981	0.798	0, 1100	Hand collecting and rearing	34	Average Temperature (°C), Annual Precipitation (mm), Vegetation cover (log of cm's intercepted)	Sarah Maunsell	2011 - 2012	(Maunsell *et al.* 2016)

Dataset	Locality	Transect coordinates	Sampling elevations	Sampling limits	Sampling scope	Domain limits (m asl)	Sampling method	Species	Environmental variables and their units	Data provider	Collection dates	Published references to the dataset
Leaf miner parasitoids (Hymenoptera)	Border Ranges (NSW)	28°24'S–28°22'S, 153°1'E–153°5'E	5	0.183, 0.981	0.798	0, 1100	Hand collecting and rearing	14	Average Temperature (°C), Annual Precipitation (mm), Vegetation cover (log of cm's intercepted)	Sarah Maunsell	2011 - 2012	(Maunsell *et al.* 2015)
Borneo Datasets												
Geometrid moths	NE Borneo	1°28'N-6°16'N, 112°06'E-117°53'E	70	0.000, 0.958	0.958	0, 4095	Light traps	775	Average Temperature (°C), Annual Precipitation (mm) [worldclim], forest stratum, vegetation type [field descriptions]	Jan Beck, Jeremy Hollo-way, Chey Vun Khen	1965-2003	(Beck *et al.* 2012) (undis-turbed habitats only)
Sphingid moths	NE Borneo	0°05'S-6°18'N, 109°43'E-118°10'E	19	0.000, 0.958	0.958	0, 4095	Light traps	102	Average Temperature (°C), Annual Precipitation (mm) [worldclim], area [of 200m bands], vegetation type [globcov]	Jan Beck, Ian Kitching et al.	1965-2005	(Beck & Kitching 2009)

North America Datasets												
Butterflies	Califor-nia	38°34'N-39°20'N, 120°20'-121°25'W	6	0.001, 0.966	0.965	0, 2775	Pollard walk, presence/ absence	129	Average Max Daily Temperature (°C), Average Min Daily Temperature (°C), Annual Precipitation (mm)	Arthur Shapiro	1973-2014	(Forister *et al.* 2010).
Mammals	Yosemite NP (Califor-nia)	37°30'N–37°59'N, 118°56'–120°28'W	40	0.000, 0.990	0.990	0, 3997	Live traps, kill traps, hunting, visual observations	46	Average Temperature (°C), Annual Precipitation (mm) [worldclim], elevational area (km per 100m elevational band) [DEM, ArcGIS]	Joseph Grinnell & Tracy Storer	1914-1916, 1919.	(Grinnell & Storer 1924; McCain 2005)

Table S2. Midpoint attractors in relation to environmental variables and observed species richness, analyzed by multiple regression with AIC-based model selection. Values of $R^2 < 0.5$ are set in italics. Values of delta AIC > 3 are boldfaced. The corresponding scatterplots appear in Fig. S1. Because the attractor is a continuous function (a doubly-truncated Gaussian distribution) and the other variables are spatially autocorrelated, significance probabilities cannot be assigned to R^2 values, which are thus best viewed as comparative. Dark grey fill indicates datasets for which the same environmental predictor (or predictor set) best explains both the Attractor and Empirical richness, by a strict ΔAIC criterion (ΔAIC > 3). Light grey fill indicates datasets for which the same environmental predictor (or predictor set) best explains both the Attractor and Empirical richness, disregarding the ΔAIC criterion.

Dataset	**Response Variable**	**Predictor Variables**	***n***	***R*²**	**Condition Number**	**Delta AIC**
Costa Rica Datasets						
Ants	Empirical richness	Modeled richness	10	0.942	1	0
		Attractor	10	0.845	1	**9.563**
	Attractor	Temperature & Relative humidity	10	0.971	2.354	0
		Temperature & Area	10	0.971	4.625	1.332
		Temperature & Precipitation	10	0.972	2.548	1.857
	Empirical richness	Temperature	10	0.856	1	0
Arctiine moths	Empirical richness	Modeled richness	10	0.879	1	0
		Attractor	10	0.762	1	**6.77**
	Attractor	Precipitation & Area	10	0.852	1.616	0
		Relative humidity & Precipitation	10	0.833	1.04	1.217
		Temperature & Precipitation	10	0.825	2.548	1.712
		Precipitation	10	0.678	1	1.793
	Empirical richness	Temperature & Area	10	0.927	4.625	0
		Precipitation	10	0.839	1	1.896
Geometrid moths	Empirical richness	Modeled richness	10	0.898	1	0
		Attractor	10	0.869	1	2.448
	Attractor	Relative humidity	10	0.623	1	0

Dataset	Response Variable	Predictor Variables	*n*	R^2	Condition Number	Delta AIC
	Empirical	Area	10	0.647	1	0
		Temperature & Area	10	0.788	4.625	0.899
		Relative humidity	10	0.608	1	1.037
		Precipitation & Area	10	0.744	1.616	2.377
Ferns	Empirical richness	Modeled richness	10	0.898	1	0
		Attractor	10	0.883	1	1.312
	Attractor	Relative humidity	10	0.666	1	0
	Empirical richness	Precipitation & Area	10	0.785	1.616	0
		Temperature & Area	10	0.770	4.625	0.69
		Relative humidity	10	0.560	1	1.165
Mammals	Empirical richness	Attractor	10	0.630	1	0
		Modeled richness	10	0.611	1	0.538
	Attractor	Area	10	*0.483*	1	0
	Empirical richness	Area	10	*0.043*	1	0
		Precipitation	10	*0.020*	1	0.267
		Temperature	10	*0.013*	1	0.348
Papua New Guinea Datasets						
Ants	Empirical richness	Modeled richness	8	0.894	1	0
		Attractor	8	0.861	1	2.142
	Attractor	Temperature	8	0.586	1	0
		Tree height	8	0.539	1	0.859
	Empirical richness	Temperature	8	0.867	1	0
Butterflies	Empirical richness	Modeled richness	8	0.975	1	0
		Attractor	8	0.950	1	**5.461**
	Attractor	Temperature	8	0.925	1	0

Dataset	**Response Variable**	**Predictor Variables**	***n***	***R^2***	**Condition Number**	**Delta AIC**
		Temperature & Relative humidity	8	0.968	1.812	2.486
	Empirical richness	Temperature	8	0.842	1	0
Birds	Empirical richness	Modeled richness	8	0.935	1	0
		Attractor	8	0.876	1	**5.222**
	Attractor	Temperature	8	0.804	1	0
		Tree height	8	0.731	1	2.53
	Empirical richness	Temperature	8	0.958	1	0
		Temperature & Basal area	8	0.985	1.36	1.299
Ferns	Empirical richness	Modeled richness	8	0.813	1	0
		Attractor	8	0.810	1	0.137
	Attractor	Basal area	8	*0.447*	1	0
		Humidity	8	*0.272*	1	2.207
	Empirical richness	Basal area	8	*0.442*	1	0
		Humidity	8	*0.236*	1	2.518
Australia Datasets						
Moths*	Empirical richness	Attractor	10	0.926	1	0
		Modeled richness	10	0.907	1	1.966
	Attractor	Temperature-Precipitation PCA	10	*0.139*	1	0
		Tree Richness	10	*0.078*	1	0.625
	Empirical richness	Temperature-Precipitation PCA	10	*0.123*	1	0
		Tree Richness	10	*0.122*	1	0.007
Leaf-miners	Empirical richness	Modeled richness	10	*0.342*	1	1
		Attractor	10	*0.163*	1	2.162
	Attractor	Temperature-Precipitation PCA	10	0.560	1	0
	Empirical richness	Temperature-Precipitation PCA & Tree richness	10	0.704	1.357	0

Dataset	Response Variable	Predictor Variables	*n*	R^2	Condition Number	Delta AIC
		Tree Richness	10	*0.338*	1	0.032
		Temperature-Precipitation PCA	10	*0.164*	1	2.131
Leaf-miner parasitoids	Empirical richness	Attractor	10	0.939	1	0
		Modeled richness	10	0.770	1	**11.878**
	Attractor	Temperature-Precipitation PCA	10	*0.442*	0.939	0
	Empirical richness	Temperature-Precipitation PCA	10	*0.476*	1	0
Borneo Datasets						
Geometrid Moths	Empirical richness	Attractor	10	*0.469*	1	0
		Modeled richness	10	*0.461*	1	0.152
	Attractor	Temperature	10	0.680	1	0
	Empirical richness	Temperature	10	*0.188*	1	0
		Precipitation	10	*0.068*	1	1.337
Sphingid moths	Empirical richness	Modeled richness	10	0.994	1	0
		Attractor	10	0.713	1	**38.012**
	Attractor	Temperature	10	0.702	1	0
		Cover Classes	10	0.683	1	0.614
	Empirical richness	Temperature & Area	10	0.944	2.034	0
North American Datasets						
Butterflies	Empirical richness	Modeled richness	11	0.968	1	0
		Attractor	11	0.936	1	**7.506**
	Attractor	Precipitation	11	*0.404*	1	0
		Precipitation & Minimum temperature	11	0.624	2.324	0.17
		Precipitation & Maximum temperature	11	0.590	2.265	1.12
	Empirical richness	Precipitation	11	0.533	1	0

Dataset	**Response Variable**	**Predictor Variables**	***n***	***R^2***	**Condition Number**	**Delta AIC**
Mammals	Empirical richness	Modeled richness	10	0.725	1	0
		Attractor	10	0.697	1	**4.655**
	Attractor	Precipitation	10	*0.429*	1	0
	Empirical richness	Precipitation	10	*0.154*	1	0
		Area	10	*0.140*	1	0.163
		Temperature	10	*0.034*	1	1.327

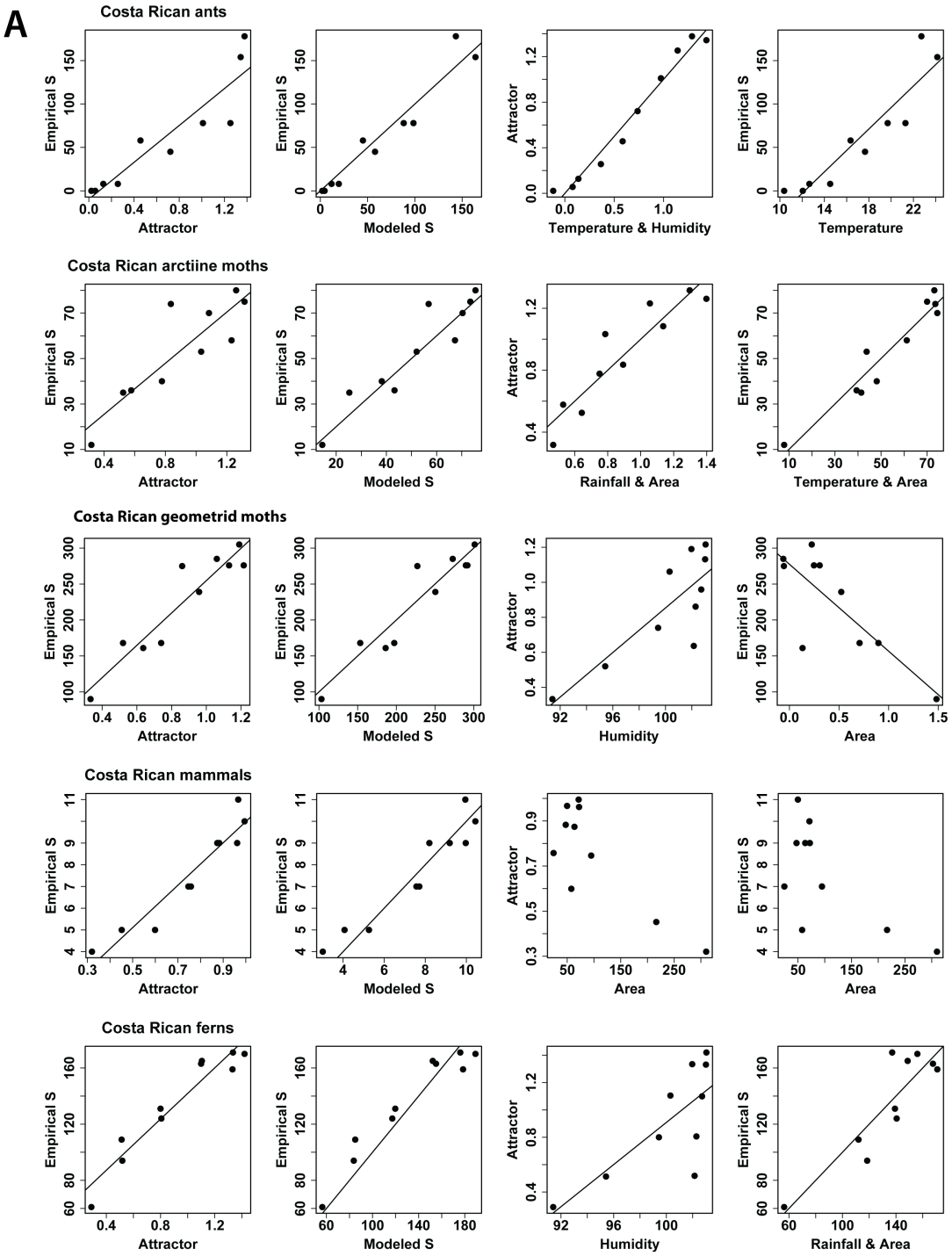
*Temperature and precipitation were highly (inversely) correlated for the Australian moths dataset (Condition Number = 21.696). PCA was extracted to reduce the effects of collinearity.

Table S3. Analysis of midpoint predictor models for range midpoint locations. Each row represents a different environmental variable used to model probabilities of midpoint occurrence along the domain. A plus sign (+) indicates $P < 0.05$, meaning that the results were improbable relative to a particular model ($P(data|model)$). Numerical entries indicate one-tailed P values, based on 1000 simulations, for which $P > 0.05$ indicates that the data were not improbable, given the model. See *Materials and Methods* in the main text for the algorithms of the two midpoint predictor models.

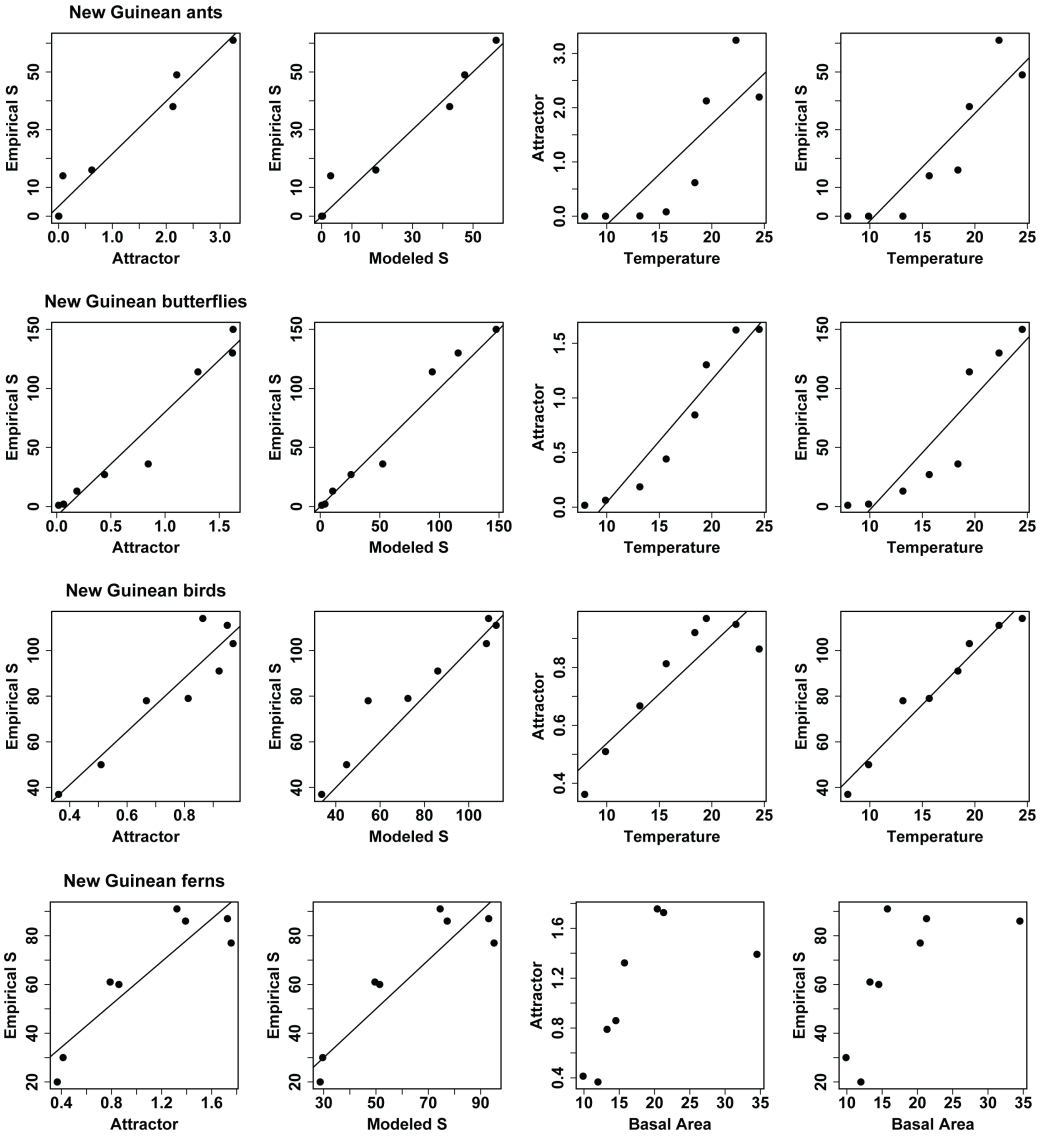
Dataset	Environmental Variable	Model 1	Model 2
Costa Rica Datasets			
Ants	Temperature	+	+
	Precipitation	+	+
	Relative humidity	+	+
	Area	+	+
Arctiine moths	Temperature	+	+
	Precipitation	+	+
	Relative humidity	+	+
	Area	+	+
Geometrid moths	Temperature	+	+
	Precipitation	+	+
	Relative humidity	+	+
	Area	+	+
Ferns	Temperature	+	+
	Precipitation	+	+
	Relative humidity	+	+
	Area	+	+
Mammals	Temperature	0.277	0.294
	Precipitation	0.300	0.305
	Area	+	+
Papua New Guinea Datasets			
Ants	Temperature	+	+
	Relative humidity	+	+
	Tree height	+	+
	Basal area	+	+
Butterflies	Temperature	+	+
	Relative humidity	+	+
	Tree height	+	+
	Basal area	+	+
Birds	Temperature	+	+
	Relative humidity	+	+
	Tree height	+	+

Dataset	Environmental Variable	Model 1	Model 2
	Basal area	+	+
Ferns	Temperature	+	+
	Relative humidity	+	+
	Tree height	+	+
	Basal area	+	+
Australia Datasets			
Moths	Temperature	+	+
	Precipitation	+	+
	Tree richness	+	+
Leaf-miners	Temperature	+	+
	Precipitation	0.132	0.083
	Tree richness	+	+
Leaf-miner parasitoids	Temperature	+	0.053
	Precipitation	+	0.056
	Tree richness	0.074	0.153
Borneo Datasets			
Geometrid moths	Temperature	+	+
	Precipitation	+	+
Sphingid moths	Temperature	+	+
	Precipitation	+	+
	Area	+	+
	Cover classes	+	+
North American Datasets			
Butterflies	Minimum temperature	+	+
	Maximum temperature	+	+
	Precipitation	+	+
Mammals	Temperature	0.266	0.069
	Precipitation	0.230	+
	Area	0.154	+

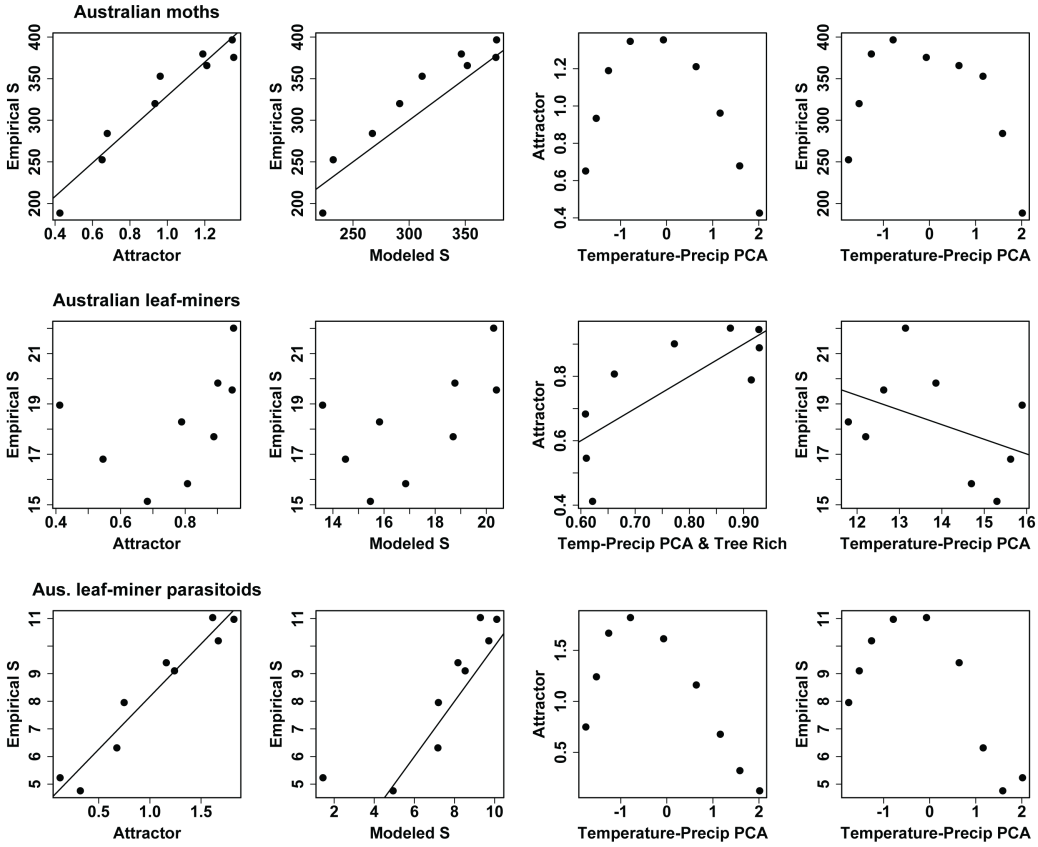
SUPPLEMENTAL FIGURES S1–S3



B


New Guinean ants
Empirical S
Attractor
Modeled S
Temperature
New Guinean butterflies
New Guinean birds
New Guinean ferns
Basal Area

C


Australian moths
Empirical S
Attractor
Modeled S
Temperature-Precip PCA
Australian leaf-miners
Temp-Precip PCA & Tree Rich
Aus. leaf-miner parasitoids

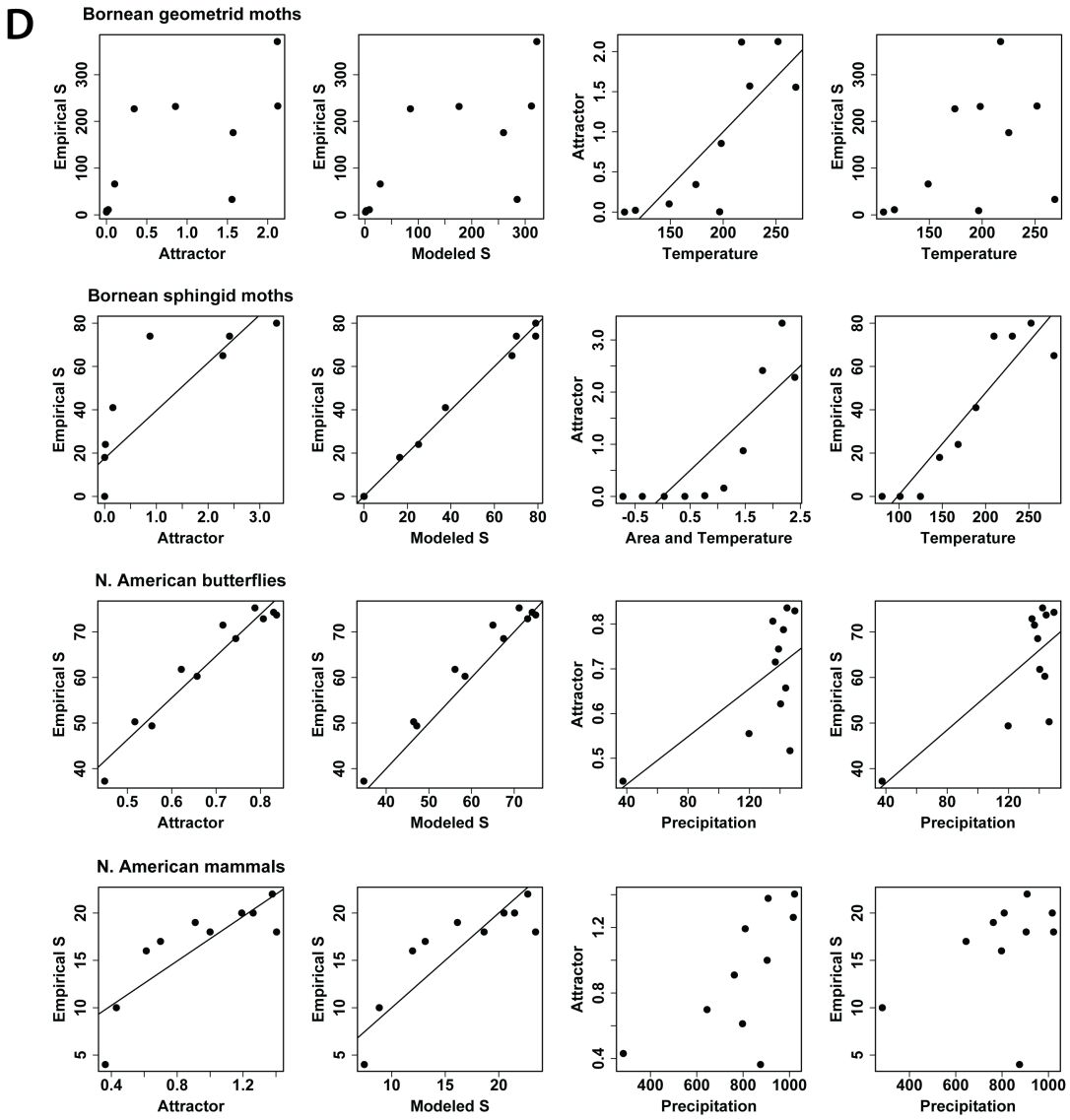


Fig. S1, A-D. Relationships between the modeled attractor, simulated species richness, empirical species richness, and measured environmental variables for each of the 16 datasets (in four geographical groups). Each dataset is represented by the four panels in a row. Within a panel, each point represents one of 9 or 10 elevations within the (rescaled) domain at which variables were evaluated. ***First panel:*** the regression of empirical richness *vs.* the magnitude of the modeled midpoint attractor function. ***Second panel:*** unity-line regression (*slope* = 1, Romdal *et al.* 2005) of modeled richness *vs.* empirical richness. ***Third panel*****:** regression of the magnitude of the modeled midpoint attractor function *vs.* the best-fitting (by AIC) environmental variables. ***Fourth panel:*** the regression of empirical species richness *vs.* the best-fitting (by AIC) environmental variables. See Table S2 for statistical results. As explained in the caption for Table S1, the regressions plotted in this figure cannot be assessed for statistical significance, because the points are not independent. Nevertheless, R^2 is as an appropriate measure of

linear goodness-of-fit between variables, sensitive both to linearity and scatter. In the plots here, we have set an arbitrary lower threshold of $R^2 = 0.5$ for display of regression lines.

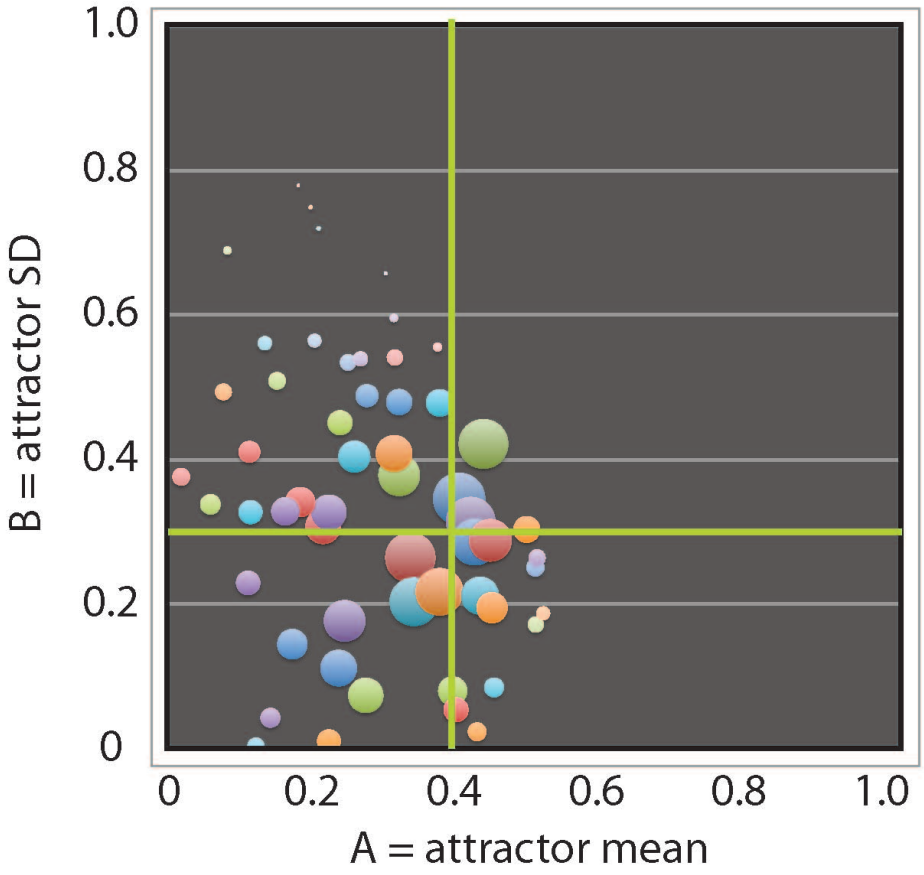


Fig. S2. Sampled (*A*, *B*) pairs of midpoint attractor parameters generated by the MCMC Gibbs sampler for the Costa Rican arctiine moth dataset. Point width is proportional to the coefficient of determination (R^2) between modeled and observed species richness across the elevational domain. Point color is arbitrary. The green lines indicate an optimized pair of parameter values ($A = 0.378$, $B = 0.294$), centered in the cluster of coordinate pairs with highest R^2, which was used to produce the model for the arctiine moth dataset in Fig. 3 (main text).

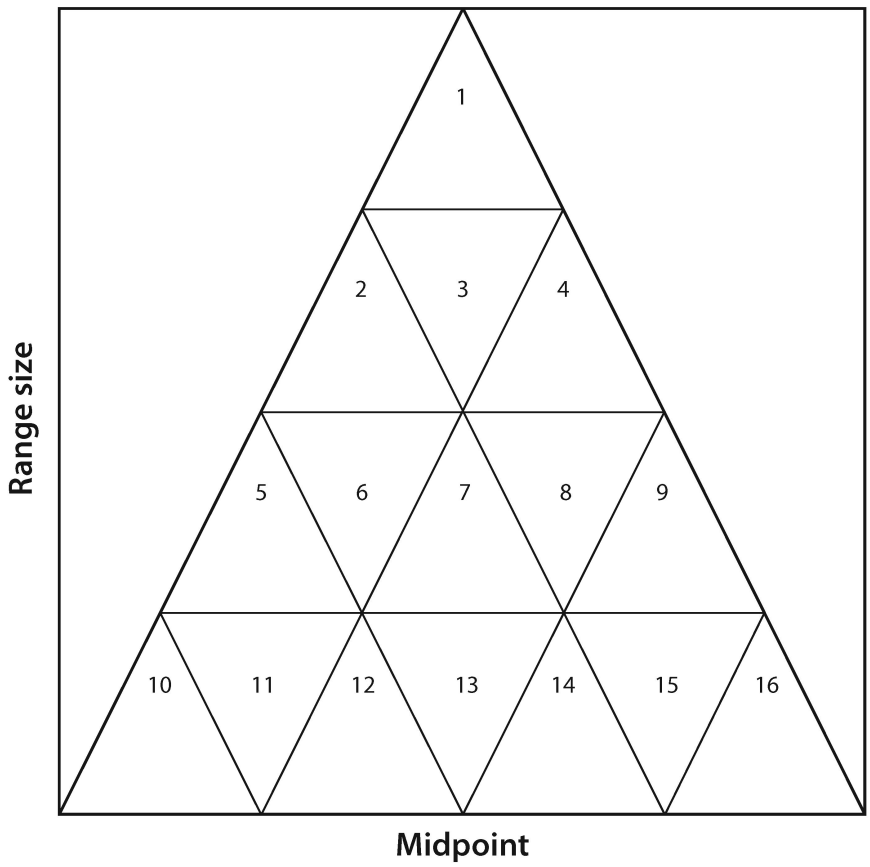


Fig. S3. The geometric constraint triangle, subdivided into 16 smaller, equal-sized triangles.

APPENDIX 2: SUPPLEMENTAL TEXT

SUPPLEMENTAL INTRODUCTION

Beginning with Lees *et al.* (1999) and Jetz and Rahbek (2001), many authors have taken a statistical approach to integrating geometric constraints with environmental variables, treating "pure" MDE model predictions as candidate predictor variables. In most of these studies, the observed range-size frequency distribution (RSDF) was sampled without replacement to generate the MDE model predictions of expected species richness at each location in the domain (Colwell *et al.* 2004, 2005). The MDE predictions and standard environmental variables were then used together in traditional correlative modeling of species richness patterns. Increasingly rigorous versions of this statistical approach have incorporated formal model selection, spatial statistics, and assessment of multicollinearity (Bellwood *et al.* 2005; Davies *et al.* 2007; Wu *et al.* 2012).

Several studies have integrated constraints and drivers directly, incorporating the interacting effects of geometric constraints and environmental drivers on species richness (Gotelli *et al.* 2009), using environmental variables to condition probabilities of range placement and expansion within a spatially bounded domain (Storch *et al.* 2006; Rahbek *et al.* 2007), thus relaxing the assumption of a pure MDE model that all parts of the domain are environmentally identical. These models were also conditioned on the empirical range size frequency distribution (RSFD). In contrast, Grytnes et al. (2008) modeled plant species richness on a bounded elevational gradient by drawing range sizes from theoretical distributions and range midpoints from a probability distribution fitted directly to the observed richness gradient.

Rangel and Diniz-Filho (2005) built a stochastic, mechanistic model that integrates speciation, range expansion, and extinction on a bounded, monotonic environmental "favorability" gradient, without reference to empirical data. The model is effectively a spatially explicit version of the neutral model (Hubbell 2001) in a one-dimensional bounded domain, but with an underlying environmental gradient. The Rangel and Diniz-Filho (2005) model generated off-center species richness peaks that emerged from the interaction between the gradient and the geometric constraints (Colwell & Rangel 2009). Without the environmental gradient—or with a very weak gradient—Rangel and Diniz-Filho's model generated a peak of species richness in the center of the domain that was qualitatively similar to the predictions of a simple MDE model.

Wang and Fang (2012) developed a third approach. They fitted a multiple regression model of species richness as a response to environmental variables, but they used only the subset of species with the smallest geographic ranges to parameterize the model. They reasoned that the placement of small-ranged species within a bounded domain is little affected by the location of range boundaries, so that, for this subset of taxa, correlations between species richness and environmental variables would not be distorted by geometric constraints. They then used the resulting model coefficients, together with the empirical RSFD, to simulate the placement of range midpoints of the larger-ranged species within the bounded domain. They showed that a single

environmental model, combined with strong geometric constraints, best explains the species richness of both small- and large-ranged plant species along elevational gradients in China.

SUPPLEMENTAL MATERIALS AND METHODS

Dataset selection and preparation

As a criterion for inclusion in this study, we applied the rule (McCain 2007; McCain 2009) that at least 70% of the physical gradient between sea level and mountaintop must have been sampled and at least four environmental variables had been reported for the gradient.

Each of the 16 datasets (Table S1) was prepared in the same way. Domain limits were defined as sea level and the highest elevation on the mountain massif upon which the gradient was located. This domain was converted to the unit line, and all empirical sampling elevations were proportionally scaled within this [0,1] domain. Environmental variables (Table S1) were resampled, as necessary, after smoothing with cubic spline interpolation, using the *splinefun* function in R, version 3.1.1 (R Core Team 2014).

If the highest elevation at which a species was recorded was not at the highest sampling location, the upper boundary for that species range was estimated to occur halfway between the highest elevation of recorded occurrence and the next higher sampling elevation. If the highest elevation at which a species was recorded at was the highest sampling elevation, the upper boundary of that species range was estimated to occur halfway between that sampling elevation and the upper limit of the domain. The lower boundary for each range was treated analogously, being extended halfway to the next lower sampling elevation or halfway to the lower domain limit (sea level), if a species was recorded at the lowest sampling elevation, but that sampling elevation was not the domain limit. The ranges of each species found at only one sampling elevation were treated similarly; otherwise, these point ranges would have had a zero range, and would have been lost from the model. We assumed that the occurrence of each species was continuous between its estimated upper and lower recorded range boundaries. These range-adjustment procedures and assumptions have been widely used in previous studies (e.g., Cardelús *et al.* 2006; Longino *et al.* 2014).

The protocol for range adjustment, described above, leaves most datasets without any empirical ranges that actually reach the domain boundaries, resulting in zero estimated empirical richness at one or both limits of the domain. A few zeroes are real (e.g., ants do not occur at very high elevations in the Costa Rica and New Guinea gradients), but most others are artifacts of the location of original sampling elevations and the range estimation protocol. Data providers (Table S1) were asked in each case whether such zeroes in their data sets were real or artifactual. If real, zero richness at the domain endpoint (and in some cases adjacent sampling points) was plotted and included in analyses; if artifactual, we proportionally adjusted all empirical range midpoints so that ranges nearest to the domain limit exactly reached it. The shifts needed to achieve this

adjustment, which effectively shifts the domain boundary slightly, were consistently very small (0.002 to 0.02 on the unit line).

To cope with the wide variation among datasets in number and spacing (often not uniform) of empirical sampling points, we took a mixed approach. For fitting the attractor (see below), we used a series of 11 evenly spaced sampling locations across the entire unit line (domain), including both ends of the domain (0 and 1), for all datasets except the New Guinea group. The New Guinea transect was sampled in the field at 8 evenly-spaced elevations, so with the domain ends added, we used 10 sampling points for fitting the attractor in those datasets. For plotting model results (main text Figs. 3, 4, and 5), we used the original sampling points for datasets with fewer than 11 original points (eight points for the four Papua New Guinea datasets, five for the three Australia datasets, and six for North American butterflies), and 11 points for all other datasets.

The Bayesian Midpoint Attractor model

The MCMC sampler and richness pattern simulation. We designed a simple MCMC Gibbs sampler (Gelman *et al.* 2013) to select (*A*, *B*) pairs for the mean (*A*) and standard deviation (*B*), the parameters of the Gaussian midpoint attractor, with the objective of simulating the richness pattern over the domain for a particular empirical dataset, using *only* the range-size frequency distribution (RSFD) as input. Empirical midpoints were completely ignored for the simulations. The goodness of fit between modeled and empirical richness was then assessed for each simulation, as detailed below.

Running the simulation. For each candidate (*A*, *B*) pair, each empirical range was placed stochastically on the domain, without replacement, using either Algorithm 1 or 2 (*Main text, Materials and Methods*). The modeled richness was recorded for *L* (10 or 11, see above) evenly spaced sampling locations across the domain, always including both ends of the domain (0 and 1). The process was repeated *M* (= 100) times, for the same (*A*, *B*) pair. The mean richness for each of the *L* sampling points on the domain was then computed, among the *M* runs, to estimate the expected richness pattern, given the (*A*, *B*) pair and the empirical RSFD.

Measuring goodness-of-fit. The next step in the MCMC procedure assessed the goodness-of-fit (GOF) between the empirical richness pattern and the mean modeled richness pattern, for a given candidate (*A*, *B*) pair, at the *L* sampling points. We applied three alternative GOF measures: (1) *r*, the Pearson product-moment correlation coefficient (but only when positive), squared; (2) the chi-squared statistic computed on standardized richness (the richness at each sampling point, divided by total richness at all *L* points), treating the empirical richness as "expected" and the modeled richness as "observed" (as is customary in Bayesian modeling); and (3) the two-sample Kolmogorov-Smirnov (K-S) statistic. Note that none of these measures can be used in this way to yield a probability test of significance; they are simply mathematically suitable measures of GOF for richness patterns. The protocol for choosing the best GOF for each dataset is described, in context, in the next section.

Sampling the parameter space. Using the procedure just described, the MCMC sampler tested a series of (*A*, *B*) pairs. At each step in this process, a candidate (*A*, *B*) pair was

proposed by drawing a new value for A and a new value for B from uniform distributions $[0 < A < 1]$ and $[0 < B < 1]$. In Bayesian terms, A was a flat prior, with the full [0,1] domain sampled uniformly for the location of the mean (A). For the standard deviation (B), we also set the upper limit at 1 because this value produces a spatial pattern of richness broader and flatter than any empirical richness pattern we have seen; thus the prior distribution of B incorporated this information. (An even higher limit for B could have been used, but the results would not have changed.)

The candidate (A, B) pair was evaluated by running the simulation (M times) and assessing goodness-of-fit (GOF) between the mean modeled richness (averaged among M runs) and empirical richness (as described above). If the GOF for the candidate (A, B) pair was better, or not much worse, than the GOF for the previous pair, the new pair was added to the chain and the process repeated. The criterion for "not much worse" is important. If only parameter sets (A, B pairs) that yield a better fit than the previous step are kept, the chain may become stuck on a local GOF "peak" in the parameter space, and fail to detect a higher peak nearby.

The criterion for accepting a candidate (A, B) pair in our model was the *threshold-for-acceptance ratio* T, between the GOF of the candidate (A, B) pair and the GOF of the previous (A, B) pair in the chain. The ratio T was compared to a uniform random number on the interval [0,1] (Gelman *et al.* 2013). If T was greater than this number, the candidate (A, B) pair was accepted and the chain continued; if T was smaller than this number, the candidate pair was rejected, and a new candidate pair was proposed. In this way, better pairs ($T > 1$) were always accepted, and some not-as-good pairs ($T < 1$) were also accepted, ensuring a better sampling of the parameter space.

For each dataset, $C = 200$ to 500 candidate pairs were tried, and the accepted (A, B) pairs (the chain) were tabulated, each with its GOF and step number in the chain. When the process was complete, the accepted (A, B) pairs were plotted (Fig. S2), and ranked by their GOF (largest to smallest for Pearson and Kolmogorov-Smirnov GOFs, smallest to largest for the chi-squared GOF).

For each dataset, when results differed substantially between the two stochastic range placement algorithms in the Bayesian attractor model (*Main text, Materials and Methods*), GOF measures were used to choose the better of the two algorithms. When results differed substantially among GOF measures for a given algorithm for a particular dataset, choice of GOF was based on minimizing overall deviation of empirical points from the 95% confidence intervals of the model. On the basis of this procedure, Pearson correlation emerged as the most successful GOF (13 of 16 datasets), with chi-squared providing a better result in two cases (Australian leaf-miners and Bornean geometrid moths), and Kolmogorov-Smirnov in one case (North American mammals).

Midpoint predictor models

For each of the two midpoint predictor models, we assessed the same set of environmental variables used to interpret modeled attractors in the Bayesian midpoint attractor model (Table S1), one variable at a time. To construct the probability density functions, the [0,1] domain was divided into 1000 bins. For each bin, the magnitude of the environmental variable was approximated by linear interpolation between measured values at sampling locations on the elevational gradients (Table S1). Next, probabilities

for each bin were assigned proportional to these measured values. Finally, a range midpoint representing each empirical species was placed stochastically in the domain in proportion to these values. For Model 1, no geometric constraints were enforced. In Model 2, range placement was constrained by the domain boundaries.

Midpoint predictor model evaluation. For each midpoint predictor model, we calculated the cumulative distribution function (cdf) of species range midpoints across the domain, averaged over 1000 simulations. Steeply rising sections of this cdf indicate elevations with a high concentration of species range midpoints, whereas flatter sections of the cdf indicate elevations where few or no species range midpoints occur. We refer to this averaged cdf as the *model reference cdf*.

We next constructed the cdf for the empirical midpoint data and calculated the maximum difference between this curve and the model reference cdf. This difference is the traditional Kolmogorov-Smirnov test statistic. To generate a null distribution and estimate the tail probability for the empirical data, we generated 1000 additional midpoint distributions with the midpoint predictor model, and for each of these we calculated the K-S test statistic between the cdf of the single simulated midpoint distribution and the model reference cdf.

We then compared the histogram of the 1000 simulated K-S differences with the observed K-S difference between the empirical data and the model reference cdf. A non-significant one-tailed value ($P > 0.05$) indicates an adequate fit with the data. In contrast, unusually large K-S values for the observed data would suggest that the midpoint predictor model does not successfully reproduce the pattern of midpoints in the data.

Software

The midpoint attractor simulator and the MCMC sampler were implemented in 4th Dimension, in an extension of the RangeModel application (Colwell 2008) that is available from the authors. The midpoint predictor models were programmed in R version 3.1.1 (R Core Team 2014), with base functions from the EcoSimR development package (https://github.com/GotelliLab/EcoSimR), which is available as an R package. R scripts for the midpoint predictor model analyses and for plotting the graphics for the midpoint attractor models (main text Figs. 3, 4, and 5) are available from the authors.

REFERENCES FOR THE APPENDICES

Ashton, L.A., Odell, E.H., Burwell, C.J., Maunsell, S.C., Nakamura, A., McDonald, W.J.F. *et al.* (2015). Altitudinal patterns of moth diversity in tropical and subtropical Australian rainforests. *Austral Ecol.*, 41, 197–208,.

Beck, J., Holloway, J.D., Khen, C.V. & Kitching, I.J. (2012). Diversity partitioning confirms the importance of beta components in tropical rainforest Lepidoptera. *Am. Nat.*, 180, E64-E74.

Beck, J. & Kitching, I.J. (2009). Drivers of moth species richness on tropical altitudinal gradients: a cross-regional comparison. *Global Ecol. Biogeogr.*, 18, 361-371.

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

Brehm, G., Colwell, R.K. & Kluge, J. (2007). The role of environment and mid-domain effect on moth species richness along a tropical elevational gradient. *Global Ecology & Biogeography*, 16, 205-219.

Cardelús, C.L., Colwell, R.K. & Watkins, J.E. (2006). Vascular epiphyte distribution patterns: explaining the mid-elevation richness peak. *J. Ecology*, 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., Rahbek, C. & 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. & Gotelli, N. (2005). The mid-domain effect: there's a baby in the bathwater. *Am. Nat.*, 166, E149–E154.

Colwell, R.K. & Rangel, T.F. (2009). Hutchinson's duality: the once and future niche. *PNAS* 106, 19651-19658.

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. B-Biol. Sci.*, 274, 1189-1197.

Forister, M.L., McCall, A.C., Sanders, N.J., Fordyce, J.A., Thorne, J.H., O'Brien, J. *et al.* (2010). Compounded effects of climate change and habitat alteration shift patterns of butterfly diversity. *Proc. Natl. Acad. Sci. U. S. A.*, 107, 2088-2092.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. & Rubin, D.B. (2013). *Bayesian data analysis*. CRC press.

Gotelli, N., Anderson, M.J., Arita, H.T., Chao, A., Colwell, R.K., Connolly, S.R. *et al.* (2009). Patterns and causes of species richness: a general simulation model for macroecology. *Ecol. Lett.*, 12, 873-886.

Grinnell, J. & Storer, T.I. (1924). *Animal life in the Yosemite: an account of the mammals, birds, reptiles, and amphibians in a cross-section of the Sierra Nevada*. University of California Press.

Grytnes, J.A., Beaman, J.H., Romdal, T.S. & Rahbek, C. (2008). The mid-domain effect matters: simulation analyses of range-size distribution data from Mount Kinabalu, Borneo. *J. Biogeogr.*, 35, 2138-2147.

Hubbell, S.P. (2001). *The unified theory of biodiversity and biogeography*. Princeton University Press, Princeton, N. J.

Jetz, W. & Rahbek, C. (2001). Geometric constraints explain much of the species richness pattern in African birds. *PNAS*, 98, 5661-5666.

Kluge, J., Kessler, M. & Dunn, R.R. (2006). What drives elevational patterns of diversity? A test of geometric constraints, climate and species pool effects for pteridophytes on an elevational gradient in Costa Rica. *Global Ecology & Biogeography*, 15, 358-371.

Lees, D.C., Kremen, C. & 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.

Longino, J.T., Branstetter, M.G. & Colwell, R.K. (2014). How Ants Drop Out: Ant abundance on tropical mountains. *PLoS One*, 9, e104030.

Longino, J.T. & Colwell, R.K. (2011). Density compensation, species composition, and richness of ants on a neotropical elevational gradient. *Ecosphere*, 2(3):art29, doi:10.1890/ES10-00200.1.

Maunsell, S.C., Burwell, C.J., Morris, R.J., McDonald, W.J., Edwards, T., Oberprieler, R. *et al.* (2016). Elevational turnover in the composition of leaf miners and their interactions with host plants in Australian subtropical rainforest. *Austral Ecology*, 41, 238-247.

Maunsell, S.C., Kitching, R.L., Burwell, C.J. & Morris, R.J. (2015). Changes in host–parasitoid food web structure with elevation. *J. Anim. Ecol.*, 84, 353-363.

McCain, C. (2005). Elevational gradients in diversity of small mammals. *Ecology*, 86, 366-372.

McCain, C.M. (2004). The mid-domain effect applied to elevational gradients: species richness of small mammals in Costa Rica. *J. Biogeogr.*, 31, 19-31.

McCain, C.M. (2007). Could temperature and water availability drive elevational species richness patterns? A global case study for bats. *Global Ecology & Biogeography*, 16, 1-13.

McCain, C.M. (2009). Vertebrate range sizes indicate that mountains may be 'higher' in the tropics. *Ecol. Lett.*, 12, 550-560.

Moses, J. (2015). Tropical elevational gradient in ants (Hymenoptera: Formicidae): Diversity patterns, food preferences and nutrient redistribution rates on Mt Wilhelm, Papua New Guinea. MSc Thesis, University of Papua New Guinea, Port Moresby., p. 85 pp.

R Core Team (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. , URL http://www.R-project.org/.

Rahbek, C., Gotelli, N., Colwell, R.K., Entsminger, G.L., Rangel, T.F.L.V.B. & Graves, G.R. (2007). Predicting continental-scale patterns of bird species richness with spatially explicit models. *Proceedings of the Royal Society of London Series B*, 274, 165-174.

Rangel, T.F.L.V.B. & 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.

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

Sam, K. & Koane, B. (2014). New avian records along the elevational gradient of Mt. Wilhelm, Papua New Guinea. Bulletin of the British Ornithologists' Club 134, 116-133.

Sam, L. (2011). Responses of butterfly (Lepidoptera) communities along an altitudinal forest gradient in Papua New Guinea. MSc Thesis, University of Papua New Guinea, Port Moresby, p. 66 pp.

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.

Tvardikova, K. (2013). Trophic relationships between insectivorous birds and insect in Papua New Guinea. Thesis Series, No. 9. University of South Bohemia, Faculty of Science, School of Doctoral Studies in Biological Sciences, České Budějovice, Czech Republic, p. 184 pp.

Wang, X. & Fang, J. (2012). Constraining null models with environmental gradients: a new method for evaluating the effects of environmental factors and geometric constraints on geographic diversity patterns. *Ecography*, 35, 1147–1159.

Wu, Y., Yang, Q., Wen, Z., Xia, L., Zhang, Q. & Zhou, H. (2012). What drives the species richness patterns of non-volant small mammals along a subtropical elevational gradient? *Ecography*, 36, 185-196.