The mid-domain effect: geometric constraints on the geography of species richness

Robert K. Colwell and David C. Lees

Geographic patterns of species richness are influenced by many factors, but the role of shared physiological and physiological boundaries in relation to range-size distributions has been surprisingly neglected. In spite of the fact that such geometric constraints lead to mid-domain richness peaks even without environmental gradients (the mid-domain effect), relying on null models, several recent studies have begun to quantify this problem using simulated and empirical data. This approach promises to transform how we perceive geographic variation in diversity, including the long unresolved latitudinal gradient in species richness. The question is not whether geometry affects such patterns, but by how much.

Robert Colwell is at the Dept of Ecology and Evolutionary Biology, U-43, University of Connecticut, Storrs, CT 06269-3043, USA (colwell@uconnvm.uconn.edu); David Lees is at the Dept of Palaeontology and the Biogeography and Conservation Laboratory (Dept of Entomology), Natural History Museum, Cromwell Road, South Kensington, London, UK. SW7 5BD (dcl@nhm.ac.uk).

The latitudinal gradient in species richness, perhaps the most conspicuous feature of global biogeography, has intrigued ecologists and biogeographers continuously since the times of de Candolle (Ref. 1, pp. 1270–1276) and Wallace (Ref. 2, pp. 65–66). In spite of this long history, consensus on the causes of tropical peaks in species richness remains elusive1–3. The overwhelming range of hypotheses proposed to account for changes of species richness (Palmer6 lists 120 named hypotheses for variation in species richness or coexistence, and Rohde7 identifies 28 specifically applied in species richness or coexistence, and lists 120 named hypotheses for variation of species richness) has, until recently, been almost completely ignored: the geometry of the latitudinal gradient in species richness is inevitable for virtually any set of ranges, theoretical or empirical, when these ranges are randomly placed within a bounded geographical domain, in the complete absence of any supposition of environmental gradients within the domain. Qualitatively, this result survives a wide range of changes in model details and assumptions. Quantitatively, it explains a surprisingly large proportion of geographic variation in species richness for diverse groups of organisms, for the empirical cases so far examined.

In spite of conceptual antecedents in the ecological literature dating back more than 40 years10,11, and a substantial literature on formally analogous problems in niche overlap and phenology (Box 1), this important result seems to have gone unnoticed until 1994 (Ref. 11).

What do the geometric models predict?

There is a growing consensus that the regional pool of species is the key determinant of the species composition of more or less unsaturated local communities12–14. Based on regional biota, geometric models11–14 demonstrate that the stochastic placement of species ranges between shared geographic boundaries can generate precise predictions of species richness at points between the boundaries. For example, in the case of random, one-dimensional placement of ranges between two boundaries, the null models predict a convex, symmetrical pattern of species richness: this pattern is either parabolic12–13 or quasi-parabolic11,14 depending upon alternative distributions of range sizes and of midpoints (Fig. 1; Boxes 2 and 3).

An instructive way to grasp this idea quickly is to experiment with a null model that simulates range size and range dominance placement within differently defined boundaries (Fig. 1). A graphical simulation program, that not only implements the principal stochastic range model variants, but also allows input of empirical range-size frequency

References

40 Rausher, M.D. (1992) The measurement of selection on quantitative traits: biases due to environmental covariance between traits and fitness. Evolution 46, 816–826
44 Clausen, J. et al. (1940) Experimental Studies on the Nature of Plant Species. I. Effect of Varied Environment on Western North American Plants, Carnegie Institution
The species range has long been studied as the fundamental unit of species richness gradients \(29-33\). Although some modeling approaches have incorporated aspects of range size evolution, distribution or overlap \(29-33\), species richness has traditionally been viewed as a function of area, history and climatic or biogeographic gradients, rather than geometric constraints. Colwell and Hurtt \(11\) were apparently the first to show that mid-domain species richness peaks can arise from geometric constraints alone, although analogous problems in ecology have been studied for more than 40 years.

In his classic paper on the relative abundance of species, MacArthur \(11\) proposed two ‘broken stick’ models. The first of these (independent, random model) divides a single unit ‘stick’ into \(n\) pieces, resulting in a strong mid-domain effect (Boxes 2 and 3). The second model was a random ‘overlapping stick’ model: two independent, uniform random sticks are placed along a unit stick, defined by the endpoints of each species in pairs of points for each species. However, in spite of the niche overlap metaphor, MacArthur was concerned only with the distribution of the length of the pieces, not the patterns of overlap among them along the unit line. Quickly dismissed for its poor fit to real data, it was resurrected by Pielou \(13\) to assess the randomness of distributional overlap among species along gradients. Later, Pielou \(13\) extended this approach, theoretically and empirically, to an analysis of latitudinal overlap among congeneric seaweed species \(14-16\). But neither MacArthur nor Pielou examined the expected distribution of range overlaps as a function of location on the domain, thus they failed to discover the mid-domain peak of richness.

The overlap of phenological events within a temporal domain is precisely analogous to one-dimensional geographical range overlap within a spatial domain. To test for non-random spacing of flowering periods, Cole \(24\) and others \(25\) (reviewed by Ostioli and Graves \(26\) computed expected pairwise overlaps and randomized empirical flowering periods within the flowering season. Cole and Sugihara \(26\) extended Pielou's analytical approach to phenology and to utilization spectra in general. Nonetheless, no one in the phenology debate examined the expected number of simultaneously flowering species as a function of time of year for randomized flowering periods, for which the mid-domain effect predicts a mid-season flowering peak.

Perhaps the most obvious manifestation of geometric constraints on patterns of geographic range on a bounded domain is the limiting placed on the relation between range size and range midpoint (Fig. 1b). Species with midpoints near a boundary necessarily have relatively small ranges or, conversely, species with increasingly larger ranges must have their midpoints increasingly nearer the centre of the domain. Pielou \(13\) was apparently first to recognize (and plot) this constraint, followed by Graves \(14\) (elevation), Rhodes \(27\) (latitude), and Pineda \(27\) (depth). Although their contributions pointed out the boundary constraint on range placement – ultimately the cause of richness peaks in stochastic range models – none of these authors noted the expected mid-domain peak.

**Box 1. Antecedents and analogues**

The richness gradient predicted by the geometric models essentially matches the observed variation in latitudinal richness for some taxa \(11\). The models explain a remarkably high proportion of empirical variation in latitudinal richness for some taxa \(11\). A striking prediction of these models is that richness patterns on elevational and on bathymetric gradients should also be unimodal, with a mid-gradient richness peak \(11,14,27\). This prediction of a mid-gradient richness peak derives, in large part, from the observation that species richness decreases monotonically with increasing elevation and depth \(14,27\). In fact, most data and recent meta-analysis suggest that richness patterns with a medial peak might be more the rule than the exception for both the elevational gradient \(11,14,27\) and the bathymetric gradient \(14,27\), although the

**Fig. 1. A fully stochastic null model for species richness-gradients within a bounded domain \(11\).** For particular sets of species, the unit domain might represent elevation from sea level to mountain peak \(9\), ocean depth from the surface to the abyss \(14\), distance from one end of a large island to another \(9\), or latitude from the northern to the southern end of the continental New World \(9\), or latitude from the northern to the southern limit of the distribution of a clade \(2\). In (a), the range size for each species is plotted against its range midpoint (100 species shown). In this model (Box 2), midpoints and range values are generated as a uniform random coverage of feasible values. In (b), the ranges for a subset (50 species) of the points in (a) are shown as horizontal lines centred on their midpoints. Because the domain is bounded at 0 and 1, all midpoint-range coordinate pairs fall on the isosceles triangle. For any point x in the domain, richness is computed as the number of horizontal range lines that a vertical line at x (the broken line) would intersect. In (c), the closed circles plot the species richness across the domain for the points in (a) and (b). The open circles plot species richness when maximum range size is limited to half the domain (0.5) and the crosses show richness for a maximum range size of 0.25. The ordinate in (c) scales richness as a proportion of all species in the simulation. In all cases, the richness peaks at the domain midpoint. Only the top curve is parabolic and peaks at a proportional richness of 0.5. Note the more pronounced mid-domain effect when larger ranges are permitted. Adapted, with permission, from Ref. 11.
Box 3. Constrained null models for species richness gradients

In addition to the fully stochastic model discussed in Box 2, we can distinguish two other fundamental geometric null models for one dimension, constrained according to the dependence structure of either range size or range placement. These correspond to the models of Lees et al. (Model 1) and Model 3 of Colwell and Hurtt (Model 2).

As for the fully stochastic case (Box 2), each of these two models generates a convex, symmetrical pattern of species richness owing to the mid-domain effect, differing only quantitatively. This simulation program of Colwell implements all these models: http://viceroy.eeb.uconn.edu/rangemodel.

The model of Lees et al. (Model 1) effectively randomizes range placements within a given range size distribution. This probabilistic, analytical null model assumes a uniform distribution of range sizes, but can be partitioned to approximate empirical range size distributions (see Box 2). Here, we seek probability densities derived for the triangles in the figure in Box 2, whose integration for this distribution of r evaluates to 2x ≈ 2x.

The ongoing debate over Rapoport’s rule and the idea that the evolution of broader physiological tolerance favours larger range sizes towards the harsher extremes of environmental gradients has motivated much of the recent work discussed here. The relationship between Rapoport’s rule and the geometric models is complex (Box 4), but the models clarify the constraints on range size variation within domains and highlight the limitations of Rapoport’s rule as an explanation for richness patterns.

The mid-domain effect: a geometry of ranges and richness

We call the geometry theory of species richness gradients the ‘mid-domain effect’. We define the mid-domain effect as the increasing overlap of species ranges towards the centre of a shared geographic domain due to geometric boundary constraints in relation to the distribution of species range sizes and midpoints. Here, we place the limits of a richness peak often not precisely in the middle of the gradient.

Another pattern predicted by the models (but in this case predicted by no other hypothesis) is that, regardless of latitude, the richness of terrestrial groups should peak in the middle of large, isolated, continuous biomes to which they are endemic, as a result of the constraints imposed by biome boundaries. Lees et al. documented a striking pattern for Madagascar rainforest (Fig. 3).

Naturally, boundaries vary in their potential to limit range sizes. For example, mountaintops set absolute elevational limits for non-airborne organisms, and land-use limits fortuitous by deeper ocean trenches offer stronger resistance to range expansion within most terrestrial clades. Other limits might be more yielding.

Lyons and Willig included physiological frontiers along with topographic features, such as orographic barriers, in the category of ‘soft’ boundaries, whereas Colwell and Hurtt used ‘hard’ to refer to any naturally definable biogeographic barrier that presents some degree of resistance to dispersal. Thus, defining, quantifying or even ranking barrier resistance is likely to remain a challenge, because the effectiveness of boundaries greatly depends on the temporal scale on which they are considered and on the phylogenetic constraints. At least, physiological limits to the geographic distribution of a clade might be just as effective as sharp physiological boundaries on the shorter time scales most relevant to macroecological patterns.

The mid-domain effect: a geometry of ranges and richness

We call the geometric theory of species richness gradients the ‘mid-domain effect’. We define the mid-domain effect as the increasing overlap of species ranges towards the centre of a shared geographic domain due to geometric boundary constraints in relation to the distribution of species range sizes and midpoints. Here, we place the limits of a
domain (in its mathematical sense of a union of a connected set and its boundary points) in a geographic context, either physiographically (e.g. the land-sea interface), biologically (e.g. biome interfaces, such as tree-line or the limits of rainforest in Madagascar), as the limits of a region of endemism (e.g. the northern and southern range limits of New World marsupials) or arbitrarily (e.g. 20°N to 20°S latitude). Although we focus on one-dimensional views of geographical domains and their geometric representation on a line, the concept applies equally to geographical areas and to volumes. Although clearly related to earlier models of geometric constraints (Box 1), the mid-domain effect focuses on the emergent, macroecological pattern that these constraints predict, rather than on the effects of constraints on individual ranges. (The effect has also been referred to, without explicit definition, as the 'geometric constraint model' or 'the Périnet effect').

In assessing the role of the mid-domain effect in specific cases, an important element of this definition is that the species considered must, collectively, share the same boundaries. In other words, species ranges that are not fully contained in the geographical domain under consideration are excluded (although such taxa can still be analysed using different domain limits). The resulting shape, intensity and, thus, detectability of the mid-domain effect depends on the taxa involved; the distribution of their range sizes and range midpoints; the existence, location and limiting potential of the boundary constraints; the sampling position or region; and on countervailing or intensifying biological, climatic or historical determinants of species richness. Thus, confirmation or rejection of the mid-domain effect in any particular case requires statistical evaluation of an appropriate null model in relation to observed patterns.

---

**Fig. 2.** Comparison of actual latitudinal patterns of species richness for New World marsupials (points) with patterns predicted by the fully stochastic null model (solid curves) (Fig. 1; Box 2). In each graph, the central curve represents the predicted mean richness for each latitude; the outer lines define the 95% confidence bands. The same empirical data points are shown, with model predictions, for three different latitudinal domains: (a) the continental New World, (b) the distributional limits of all New World marsupials and (c) the smallest domain within which 95% of all New World marsupial species occur. Reproduced, with permission, from Ref. 13.

**Fig. 3.** Latitudinal proportional species richness, based on interpolated ranges, for species strictly endemic to the Madagascan rainforest biome. In (a) observed richness for ten faunal groups (carnivore beetles, tiger beetles, butterflies, ctenuchine moths, chameleons, frogs, birds, tenrecs, rodents and lemurs) is compared with an analytical null model (Boxes 3 and 5). The model computes expected richness based on either two or four equal partitions of range size. The middle curve represents all 637 species (closed squares, observed; solid curve, 4-partition analytical model). The top curve represents all 182 wide-ranging species – species with ranges that cover at least half the domain (open squares, observed; solid curve, 2-partition analytical model). The lower curve represents all 455 narrow-ranging species – species with ranges that cover at most half the domain (open circles, observed; solid curve, 2-partition analytical model). Modified, with permission, from Ref. 15. In (b) the same data are plotted as for the middle curve in (a) for all 637 Madagascan rainforest species, showing the fit to 4-partition analytical geometric null model, which explains 85% of the variance (uncompensated for spatial autocorrelation).
produce a model so constrained that statistical comparison with empirical richness patterns would
within boundary constraints, to each empirical midpoint. We suspect that incorporating both separate the mid-domain effect (inevitably present) from biological and historical causes of rich-
model, Lyons and Willig12 (in their ‘pseudorandom’ simulations) assigned a range at random,
use a similar method for range placements. In an empirical version of the constrained midpoint
approach for empirical range sizes; the simulations of Pineda and Caswell14 (their Figs 5 and 6)
on the domain at random, given the midpoint constraints imposed by the size of the range. The
history of null models in other areas of ecology45 suggests that no single model of Lees (see above) along the smaller (longitudinal) axis, given regional latitudinal range
constraint, an analytical method based on partitioning of energy (Fig. 3; Boxes 3 and 5). Latitudi-
al dimensions) than either area or distance. To date, four major studies have assessed this heretical conjecture.
In their study of the distribution of bats and marsupials throughout the continental New World, Wilig and Lyons13 found that, when latitudinal distribu-
tional limits were used to approximate geographic limits for these groups as a whole, a geometric null model (Box 2) can explain 69-84% of variation in species richness patterns over the entire transect (Fig. 2). Richness for marsupials reaches its peak well south of the equa-
tor (at about 20°S), out of phase with climatic gradients, thus weakening explanations based on climate or energy.
In an analysis of the distribution of nearly 1200 species of insects and verte-
brates, including 637 species endemic to the Madagascan rainforest (a biome encompassing about 13 degrees of lati-
tude and 2100 m of elevation entirely south of the equator), Lees et al17 iden-
tified geometry as the most important cause of species richness gradients. A geometric null model explained far more of the variation (85% for latitude alone, 75% for latitude and longitude as orthog-
nonal dimensions) than either area or energy (Fig. 3, Boxes 3 and 5). Latitudi-
ally and longitudinally, species richness peaked at approximately the (mid-
domain) position and the level expected from the model, given the location of biome boundaries, the empirical species range size distributions and the size of the regional species pool. How-
ever, for at least one large adaptive radi-
dation examined, the level of richness peaked somewhat below the middle of the el-
vational domain (as is common for elevational richness gradients55), sug-
ning a significant modifying influence of other factors.
In a comprehensive study of the elevational distribution of 2800 species of birds in all tropical biomes of South
America (including both western and eastern slopes of the Andes), Rahbek56 found that, once effects of area are fac-
toced, geographical range size of species richness emerges – as predicted by geo-
metric models, although at an elevation below the true elevational midpoint. He con-
cluded that the mid-elevation peak is best explained by boundary constraints
In summary, work completed so far indicates a strong influence of the mid-domain effect in shaping geographic patterns of richness, when evaluated against competing biological and climatic hypotheses. The components of these patterns that cannot be predicted by geometric null models (e.g. off-centre elevational or bathymetric richness peaks or large residuals for individual sampling sites) stand out as appropriate targets of research for non-geometric explanations.

**Stochastic versus deterministic perspectives**

Although comparison with an appropriately constructed and parameterized geometric null model is necessary to detect the probable influence of the mid-domain effect, the demonstration of such influence neither assumes nor proves that species’ ranges, or the boundaries that constrain them, are individually random in shape and in placement, with respect to ecological, evolutionary and historical causes. Of course, they are not, but the seeming paradox is a familiar one. No biologist would consider it disturbingly paradoxical that genes and environment largely determine the individual heights of a sample of 1000 adult women, whereas the distribution of their heights nearly fits a normal curve based on mean and variance in height. Analogously, the mid-domain effect is a macroecological outcome of geometric constraints that, for large numbers of taxa with their own history and adaptations, shape richness patterns along geographic gradients. These constraints arise from boundaries (both geographic limits and biological thresholds) and from an emergent statistical property: demographically defined biotas: their range size and midpoint distributions.

How, then, are we to view the real world relevance of a geometric null model? There seem to be two, non-exclusively alternative approaches to interpreting empirical species richness patterns. On the one hand, we can assume that range sizes and range placements are governed by a strong element of chance. On the other hand, we could view a good fit between model and observed species richness patterns from a deterministic perspective. Thus, ranges largely determined by the diverse adaptations and the histories of species in a regional assemblage might nonetheless fit a geometric null model, rather than yielding an emergent richness response to an environment gradient. Lyons and Willig express this view: ‘Within a larger taxon such as bats or marsupials, no single factor may attain hegemony in limiting the distribution of constituent species, and as a consequence, range boundaries may have a geographic distribution that is quite similar to those produced by chance, even though different deterministic factors account for the limitations of each species.’

A specific way to reconcile the mid-domain effect with determinism is to view range overlap from a phylogeographic perspective at the species level or above. Sister populations that have diverged allopatrically or parapatrically might emerge over evolutionary time with gene flow from source populations at the centre of the ancestral species’ range. Eventually, the two ranges might overlap, thus enriching a local biota. Averaged over many such cladistic events, the most probable area of overlap will be at mid-domain. Indeed, if each of the two (continuous) ranges expands to at least half the available domain, both must occur over mid-domain regions, the essence of the mid-domain effect. Such wide-ranging species tend to mask any environmental influence on local species richness [4, 15]. However, in the case of two species whose ranges each span at most half the domain (narrow-ranging species), the contribution to richness at mid-domain will average less than two. The probability of two narrow-ranging species both occurring at mid-domain therefore decreases from one towards zero as the range size becomes smaller [15] (Boxes 2 and 3). Thus, the geometric theory of species richness generates a specific and highly testable prediction: that wide-ranging species or higher taxa within a regional assemblage are considerably more likely to show patterns in accordance with geometric theory than narrow-ranging taxa [15, 12]. (Fig 3a) Geographic patterns of species richness in narrow-ranging taxa, less constrained by geometry, are more likely to reflect environmental and historical factors. Therefore, we could use a way to test for geometric effects is to categorize data by range size class [14, 17, 20].

**Prospects**

Incorporating the geometry of ranges and of richness into local, regional and global perspectives on biodiversity promises far-reaching consequences. Geographic patterns are capable of precisely predicting the expected shape and local magnitude of the interaction between geographic domains and species ranges for classes of different range size. Thus, by considering deviations from null model expectations, the prospects are good for disentangling biological, climatic and historical factors that affect species richness at different spatial and temporal scales (Box 5). The mid-domain effect emerges as a compellingly parsimonious foundation for a general theory of patterns of species richness along different gradients at a range of scales.

In effect, biogeographers and ecologists have simply had the wrong null model at the back of their minds for the past 150 years, by assuming that, were there no climatic, physical or biological gradients, species richness would be the same at all latitudes, elevations and depths. The mid-domain effect shows that mid-domain richness peaks are to be expected in the absence of such gradients. Departure from the expected richness peak, under an appropriate null model, but not the peak itself, requires biological or historical explanation at geographic scales.

**Acknowledgements**

We thank Hal Caswell, Elaine Cole, Pan Eter, Steve Gaines, Nick Gotelli, Gary Graves, Chris Humphreys, George Hurt, Kate Lyons, Clive Moncrieff, Mike Palmer, Chris Peluso, Jesús Pineda, Carston Rabbeke, Brian Roan, Sasha Specter, Dick Vane-Wright, Joseph Veitch, Paul Williams, Michael Willig and anonymous referees for consultation, for comments on this article and for providing articles in press. Support was provided by the US National Science Foundation (DEB-9706976), the University of Connecticut (Ecolaw), the Biotechnology and Biological Science Research Council, the Leverhulme Foundation and the National Environmental Research Council (Ecolaw).

**References**


14. Caswell H compared the parabolic richness gradient for northwest Atlantic gastropods and polychaetes, Pineda and Caswell compared the parabolic richness gradient for these groups with patterns constructed and parameterized geometric null model. 15. Caswell14 compared the parabolic richness gradient for northwest Atlantic gastropods and polychaetes, Pineda and Caswell compared the parabolic richness gradient for these groups with patterns constructed and parameterized geometric null model. 16. Caswell14 compared the parabolic richness gradient for northwest Atlantic gastropods and polychaetes, Pineda and Caswell compared the parabolic richness gradient for these groups with patterns constructed and parameterized geometric null model. 17. Caswell14 compared the parabolic richness gradient for northwest Atlantic gastropods and polychaetes, Pineda and Caswell compared the parabolic richness gradient for these groups with patterns constructed and parameterized geometric null model. 18. Caswell14 compared the parabolic richness gradient for northwest Atlantic gastropods and polychaetes, Pineda and Caswell compared the parabolic richness gradient for these groups with patterns constructed and parameterized geometric null model. 19. Caswell14 compared the parabolic richness gradient for northwest Atlantic gastropods and polychaetes, Pineda and Caswell compared the parabolic richness gradient for these groups with patterns constructed and parameterized geometric null model. 20. Caswell14 compared the parabolic richness gradient for northwest Atlantic gastropods and polychaetes, Pineda and Caswell compared the parabolic richness gradient for these groups with patterns constructed and parameterized geometric null model.
PERSPECTIVES

References
2 Wallace, A.R. (1878) Tropical Nature and Other Essays ed), MacMillan
10 Rosenzweig, M.L. (1992) Species diversity gradients: we know more and less than we thought. J. Anim. Ecol. 70, 715–730
40 Pielou, E.C. Ecology and Biodiversity, John Wiley and Sons
48 Rohde, K. et al. (1991) Rapoport’s rule does not apply to marine teleosts and cannot explain latitudinal gradients in species richness. Am. Nat. 142, 1–18

Current trends
– articles of ecological or evolutionary interest in recent issues of other Trends magazines
• Lateral genomics, W. Ford Docility Trends in Genetics 15, M5–M8 (December 1999)
• Constructing an RNA world, David P. Bartel and Peter J. Urnau Trends in Genetics 15, M9–M13 (December 1999)
• Human evolution, Shvanê Pâbê Trends in Genetics 15, M13–M16 (December 1999)
• Genetic instability and darwinian selection in tumours, Daniel O. Cahill et al. Trends in Genetics 15, M57–M60 (December 1999)
• Plants, animals and the logic of development, Elliott M. Meyenowitz Trends in Cell Biology 9, M65–M68 (December 1999)
• The origin and evolution of segmentation, Gregory K. Davis and Nipam H. Patel Trends in Cell Biology 9, M68–M72 (December 1999)