

The river domain: why are there more species halfway up the river?

Robert R. Dunn, Robert K. Colwell and Christer Nilsson

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

Biologists have long noted higher levels of species diversity in the longitudinal middle-courses of river systems and have proposed many explanations. As a new explanation for this widespread pattern, we suggest that many middle-course peaks in richness may be, at least in part, a consequence of geometric constraints on the location of species' ranges along river courses, considering river headwaters and mouths as boundaries for the taxa considered. We demonstrate this extension of the mid-domain effect (MDE) to river systems for riparian plants along two rivers in Sweden, where a previous study found a middle-course peak in richness of natural (non-ruderal) species. We compare patterns of empirical richness of these species to null model predictions of species richness along the two river systems and to spatial patterns for six environmental variables (channel width, substrate fineness, substrate heterogeneity, ice scour, bank height, and bank area). In addition, we examine the independent prediction of mid-domain effects models that species with large ranges, because the location of their ranges is more constrained, are more likely to produce a mid-domain peak in richness than are species with small ranges. Species richness patterns of riparian plants were best predicted by models including both null model predictions and environmental variables. When species were divided into large-ranged and small-ranged groups, the mid-domain effect was more prominent and the null model predictions were a better fit to the empirical richness patterns of large-ranged species than those of small-ranged species. Our results suggest that the peak in riparian plant species richness in the middle courses of the rivers studied can be explained by an underlying mid-domain effect (driven by geometric constraints on large-ranged species), together with environmental effects on richness patterns (particularly on small-ranged species). We suggest that the mid-domain effect may help to explain similar middle-course richness peaks along other rivers.

R. R. Dunn (Rob_Dunn@ncsu.edu), Dept of Zoology, North Carolina State Univ., Raleigh, NC 27695-7617, USA. – R. K. Colwell, Dept of Ecology and Evolutionary Biology, Univ. of Connecticut, Storrs, CT 06269-3043, USA. – C. Nilsson, Landscape Ecology Group, Dept of Ecology and Environmental Science, Umeå Univ., SE-901 87 Umeå, Sweden.

Biologists have long been interested in patterns of species composition and diversity along rivers (Bates 1863). As early as 1912 ecologists noted that the number of species of some taxa vary in seemingly predictable ways from headwaters to mouth (references in Hawkes 1975). One of the more frequently discussed patterns is a peak in species richness in the longitudinal middle-courses of rivers, half-way between headwaters and mouth (we use

the term longitude here in its limnological sense of the length dimension of a river, not to indicate east-west geographical coordinates). Middle-course peaks have been reported in groups as varied as aquatic arthropods (Minshall et al. 1985, Naiman et al. 1987), mussels (Hughes and Parmalee 1999), and riparian vegetation (Nilsson et al. 1989, 1991, Tabacchi et al. 1996). Reports of middle-course richness peaks are especially common

Accepted 19 December 2005

Copyright © ECOGRAPHY 2006
ISSN 0906-7590

when the entire lengths of river systems have been considered (Tabacchi et al. 1996, Ward 1998). It is too soon to call the peak in species richness in the middle courses of rivers a general pattern, but not too soon to wonder about its causes when it does occur.

Historically, ecologists have often explained these peaks or humps in terms of environmental or biological variables (Vannote et al. 1980), disturbance regimes (Ward 1998), or combinations of the two (Nilsson et al. 1989). Typically, different explanations are invoked for different taxa (Tabacchi et al. 1996, Ward 1998). For example, high temperature variability in middle courses of rivers is often cited as the explanation for high arthropod diversity (Minshall et al. 1985, Naiman et al. 1987), but patterns of riparian plant diversity are explained in terms of river-edge heterogeneity (Tabacchi et al. 1996) or substrate transitions related to glacial history (Nilsson et al. 1991). Just as for latitudinal, longitudinal, and elevational transects (reviewed by Colwell et al. 2004, 2005), the implicit null model most biologists have historically assumed when studying river courses is that species richness would be constant along the entire length of a river in the absence of the influence on richness of environmental or biological gradients along the river course (Vannote et al. 1980, Statzner and Higler 1985, Tabacchi et al. 1996).

We suggest that the appropriate null model for species richness patterns along river courses is unlikely to be a uniform richness distribution, equal throughout the length of the river. As a starting point for the consideration of null models along entire river lengths, we hypothesize that, for riverine species, the biological boundaries imposed by non-river habitats at the river headwaters and by the sea at the river mouth may tend to shape the expected pattern of species distributions along rivers toward a middle-course peak in richness, independent of any additional influence of environmental or other gradients. Such a model approximates the pattern expected if species' ranges were to evolve at random along the domain and if, collectively, species were no more likely to evolve adaptations to any one particular point on the river than to any other. This pattern is known as the "mid-domain effect" (MDE), for the mid-domain hump that null models of species distributions tend to produce when the geographical domain is bounded (Colwell and Lees 2000, Colwell et al. 2004, 2005). MDE models for other types of domains, including elevational (e.g. Grytnes and Vetaas 2002, Cardelús et al. 2006, Watkins et al. 2006), latitudinal (e.g. Lees et al. 1999, Jetz and Rahbek 2001, McCain 2003), depth gradients (Pineda and Caswell 1998), phenological patterns (Morales et al. 2004), and microhabitat distributions (Lusk et al. 2006) have shown that random placement of ranges within a bounded

domain, according to any of a variety of rules, produces a peak in richness in the middle of the domain.

Based on MDE studies in other spatial domains we can make an additional prediction for "river domains". Larger species' ranges or extents along the river are expected to produce a more pronounced mid-domain peak in richness than smaller species' ranges along the same river, because the midpoints of larger ranges are more constrained with regard to their placement within the domain. Conversely, small-ranged taxa, which are less constrained in midpoint location by the river's mouth and headwaters, are less likely to produce or contribute to a middle-course peak in richness (reviewed by Colwell et al. 2004).

Nilsson et al. (1989) documented a middle-course peak in natural riparian plant species richness along two Swedish rivers, but were unable to adequately explain the pattern. Using data from Nilsson et al.'s study (1989) we conduct a first test of the explanatory potential of MDE models for species richness along rivers using a novel, discrete model, in which the ranges of riparian plant species are randomized among sampled sites along the rivers. We compare empirical richness patterns to richness predicted by these models, which preserve the recorded size of species' ranges along the river as well as their recorded degree of patchiness (distributional gaps). Using multiple regression, we assess the explanatory power of MDE model predictions, channel width, substrate fineness and heterogeneity, ice scour, bank height, and bank area along the rivers, with recorded richness of natural riparian plants as the response variable. We also examine the independent prediction of MDE models that species with large ranges along the domain are more likely to display a mid-domain peak in richness than are species with small ranges. We use the analysis of these datasets to show how mid-domain null models can be extended to river systems, and we examine the extent to which the MDE approach offers new insights and asks new questions about river systems.

Methods

Although rivers have width, depth, tributaries, catchment area, discharge rate, and many other measurable features, our models focus on the distributions of organisms over the longitudinal dimension (the length) of rivers, and thus are effectively one-dimensional. The river domains considered are consequently analogous to other domains that can be reasonably treated as if they were one-dimensional, such as coastlines (Pielou 1977), elevational and depth gradients, and latitudinal gradients (reviewed by Colwell et al. 2004, 2005). Geometric influences orthogonal to a river-course domain (e.g. bank height, bank area, channel width, channel depth)

can be treated statistically as explanatory variables that are spatially indexed by distance along the river domain.

We analyzed the data of Nilsson et al. (1989) for “natural” (non-ruderal) riparian plants of the Torne and Kalix Rivers in northern Sweden, two of the longest free-flowing rivers in Europe (510 and 450 km, respectively) (Appendices 1–3). Both rivers were sampled from the Bothnian Bay to the respective river headwaters by Nilsson et al. (1989) at sites spaced every ten kilometers on each of the two rivers, yielding 46 sample sites on the Torne River and 37 on the Kalix. At each study site, they sampled a 200-m long section of the riverbank between the high water level and the lowest summer water level, recording all vascular plants within each 200 m transect.

To date, mid-domain analyses for one-dimensional domains have generally treated ranges as if they were continuous, even when analyzing data that were taken at discrete sampling sites (Cardelús et al. 2006, Watkins et al. 2006). In these studies, gaps in observed distributions (in the simplest case, a species is present at sites n and $n+2$, but not at site $n+1$) were filled by “interpolation” between the most distant sampling points at which each species was recorded along the domain. This is the one-dimensional equivalent of drawing continuous polygons around known occurrences for two-dimensional ranges (e.g. Jetz and Rahbek 2002). Observed gaps may represent genuine breaks in a species’ distribution (at the grain size of the study). Alternatively, gaps may represent undersampling error since, in most field inventories, only a proportion of the species that can often be found at a given site are actually present or are successfully detected at that site in a given sampling event.

Whatever the reason for observed distributional gaps, Grytnes and Vetaas (2002) demonstrated through simulations that ignoring gaps through the use of interpolation tends to inflate diversity estimates in the middle of the domain more than at the limits of the domain, which might spuriously amplify or even create mid-domain peaks in richness. (Zapata et al. [2003] also emphasized this issue.) The problem can be addressed statistically (e.g. Cardelús et al. 2006, Watkins et al. 2006), if gaps are infrequent, but the Swedish river data of Nilsson et al. (1989), like many other datasets, have numerous gaps. For this reason, inspired by the hypothetical “sampling” models of Grytnes and Vetaas (2002), we developed for this study a novel, discrete null model based on Colwell and Hurr’s (1994) continuous Model 2. (This new model is implemented in RangeModel software, ver. 5, Colwell 2006.) In this model, sampling points along the domain are treated as ordered, evenly spaced, discrete bins. A species’ range is measured as the number of bins between the most upriver site at which a species occurs and the most downriver site at which a species occurs, including those extreme sites, regardless of how many occupied sites or unoccupied sites (gaps) occur within the range.

Occupancy is measured as the total number of sites at which a species occurs, including the extreme sites that define its range. As an example, if a species occurs at sampling sites 1, 3, and 4, but not at site 2, its range is 4 and its occupancy is 3.

The objective of the null model is to generate the pattern of species richness over the ordered bins (the domain) that would be expected if observed species’ ranges were placed at random within the domain, maintaining their observed range sizes and occupancies, under the geometric constraint that no range may extend beyond domain limits. The ordered bins that represent the domain are indexed by the integers $1..Q$. Species’ ranges are systematically selected, one at a time (without replacement), then placed independently and at random on the discrete domain. This process begins by choosing a (uniform) random digit between 1 and Q to represent the potential midpoint of the chosen range. If the range lies fully within the domain, given its size and this random midpoint, the range is kept in that position. Initially, a successfully placed range occupies all intervening bins between its range endpoints, but occupancy is immediately adjusted to the recorded occupancy for that species by deleting, at random, as many occurrences as necessary from the interior of the range (endpoints are never deleted). The algorithm then repeats with the next range.

If, on the other hand, the range extends beyond the domain limit, given the initial midpoint assigned, that midpoint is discarded and another midpoint is chosen from the uniform random distribution of integers $1..Q$. The process of choosing a midpoint is repeated as many times as necessary until the range lies fully within the domain, followed by stochastic adjustment of occupancy as detailed above. The range placement algorithm repeats with each empirical range until all have been placed on the domain and each range has been adjusted to its observed level of occupancy. Modeled species richness for each bin (representing each sampling site along a river) is simply the total number of species present in that bin at the end of this process. The entire process is repeated N times to establish expected mean richness and its 95% confidence interval over the domain. One practical detail of the algorithm bears mention. Because range sizes are measured by integers, the concept of “midpoint” is ambiguous for range sizes represented by even integers. For example, a range of size 4 could have its “midpoint” at either sampling site 2 or site 3 of its range. For these “even” ranges, the midpoint used for testing fit within the domain is chosen at random from the two candidate integers.

We carried out simulations for the full dataset for each of the two rivers. Then we divided the dataset for each river into the 50% of species with the largest ranges and the 50% of species with the smallest ranges and repeated the simulations for both range-size subsets for each river.

For the full dataset and for the two range-size subsets for each river, we repeated the resampling 5000 times. Based on the data for these 5000 runs, we estimated 95% confidence intervals for the pattern of species richness along each river for the full dataset and each range-size subset. This model approximates the pattern expected in a scenario in which the position of species' ranges over along the river is collectively random with respect to any biological or environmental gradient along the river, while the distribution of possible range sizes remains characteristic of the biology of the taxon in relation to the characteristics of the river (Colwell et al. 2004, 2005). Empirical richness patterns might also be compared to patterns based on theoretical range size frequency distributions (RSFDs), such as the "fully stochastic" ("binomial") model (Colwell and Lees 2000, Laurie and Silander 2002). However, as explained in detail by Colwell et al. (2004) and demonstrated by McCain (2003), using a theoretical RSFD (instead of the empirical one) confounds any comparison of predicted versus observed richness patterns with a comparison of theoretical versus empirical RSFDs.

We examined the explanatory potential of model predictions by comparing the observed species richness based on raw species counts to the mean richness and 95% confidence intervals for the modeled richness as a function of distance along the river domain (Appendices 1 and 2). In addition, to assess the role of MDE in the context of other candidate influences on pattern of species richness, we used OLS (Ordinary Least Squares) multiple regression, relying on Akaike Information Criterion (AIC), as implemented in SAM (Rangel et al. 2005), to select the best model. In these regressions, observed species richness of natural riparian plant species was the response variable, with modeled species richness and several environmental factors as explanatory variables.

In the initial stage of our model selection algorithm, we carried out OLS regression for all single-variable models and selected the model with the minimum AIC. In the second stage, we then examined all two-variable models that included the variable chosen in the first step and chose the model with the minimum AIC. We then repeated the procedure for all three-variable models that included the two already selected, and so on, until AIC could not be further reduced. At each stage we saved the residuals from the model regression and examined spatial autocorrelation in the residuals using Moran's I correlograms. Because spatial data are usually not independent, OLS regression is not justified for spatial analyses unless spatial autocorrelation in model residuals is insignificant (Diniz-Filho et al. 2003).

The environmental variables we chose to examine were those considered by Nilsson et al. (1989) as potentially important. The variables considered (explained in detail in Nilsson et al. 1989) were channel width, substrate

fineness and heterogeneity, ice scour (percentage of riverbank showing evidence of winter ice action), bank (flood) height, and bank area (Appendices 1 and 2). We examined the distribution of each variable and applied transformations as necessary before running the regressions. We repeated the model selection process for all natural riparian plant species and for large-ranged and small-ranged subsets of natural riparian plant species.

To test whether riparian plants' ranges along river domains are consistent between the two Swedish rivers, we compared the range sizes between the datasets for these rivers using simple correlations. If riparian plants have consistent ranges along rivers that reflect distinct upstream and downstream boundaries for each species, the expectation is that range sizes should be correlated among river systems in the same region, even if those river systems differ in topography and other factors. On the other hand, if range cohesion is weak and the factors that govern distribution of species along rivers are less tied to distance along river than to other factors independent of river length, we expect little or no correlation between the distributions of the same species on different rivers.

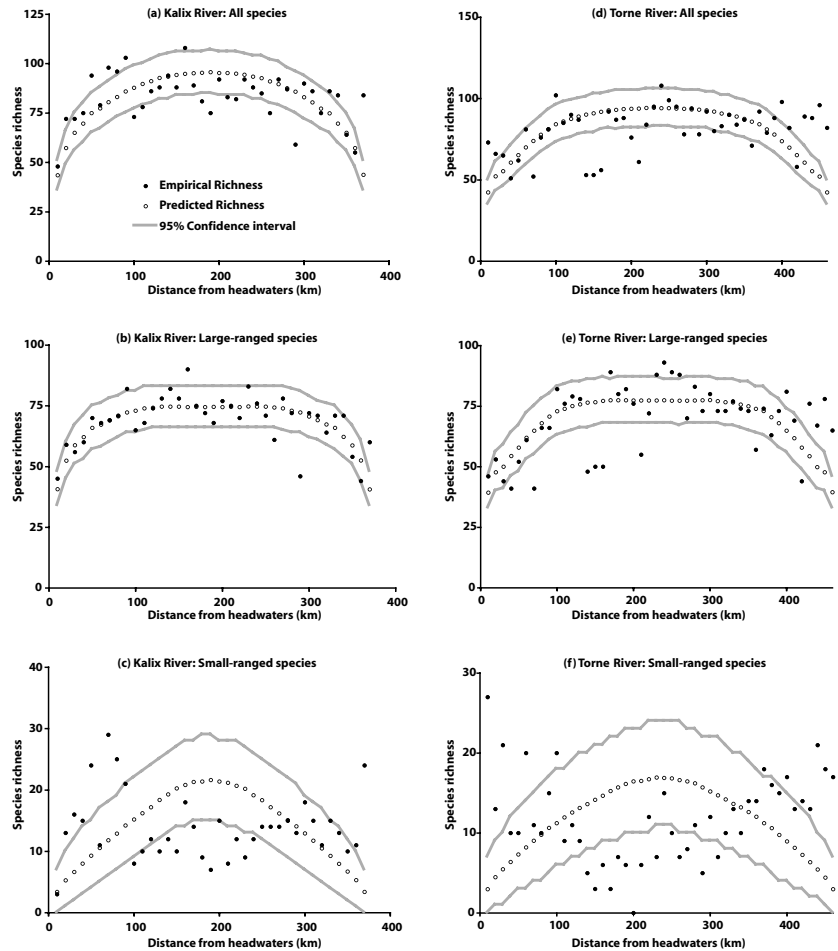
Results

Mid-domain effect and environmental correlates of richness

For the Kalix and Torne Rivers, natural plant species showed a mid-domain peak in empirical richness (as earlier reported by Nilsson et al. 1989), but with substantial variation around the overall pattern (Fig. 1). Judging from the distribution of empirical richness values (solid points) in relation to the 95% confidence envelope for the model predictions (open points), the discrete MDE model revealed, as predicted, a considerably better fit for large-ranged species (Fig. 1b and e) than for small-ranged species (Fig. 1c and f).

Results of OLS multiple regressions of observed richness as a function of MDE predictions and environmental variables appear in Table 1. For all regressions, we log-transformed channel width, bank area, and bank height to normalize the distributions and reduce the effect of outliers. Moran's I correlograms for all six final (best) models in Table 1 revealed no significant spatial autocorrelation in the OLS regression residuals for any individual distance class (we used 10 distance classes) and no overall trends in autocorrelation, justifying the use of OLS regression for these models. (This result is an excellent example of the phenomenon discussed by Diniz-Filho et al. 2003.) There was inconsistent evidence for limited residual autocorrelation in a few of the intermediate models, particularly for small ranges, but none in models that included MDE.

Fig. 1. Empirical and predicted (null) species richness of riparian plant species as a function of distance from headwaters, for the Kalix River (a) when all natural species are considered, (b) for the larger-ranged half of the species, and (c) for the smaller-ranged half; and for the Torne River (d) when all species are considered, (e) for the larger-ranged half of species, and (f) for the smaller-ranged half. Points run from headwaters at left to river mouth at right. The 95% confidence intervals are based on 5000 randomizations of range placement for the set of empirical range sizes, taking distributional gaps into account.



When plant species of all range sizes (“All species” in Table 1) were analyzed together, MDE was the first explanatory variable to enter the regression model for Kalix River plants (explaining 29% of the variation in richness), with a smaller explanatory contribution (an additional 18%) from substrate heterogeneity and a minimal contribution from ice scour and bank height (Table 1) in the final (best) model. On the Torne River, bank area entered the model first (indicating sample size effects). MDE was the second (and final) variable to enter the regression, but increased the variance explained by the regression model by only 7% (Table 1).

The regression models confirmed that, when species were divided into large-ranged and small-ranged species, the mid-domain effect was considerably more explanatory for large-ranged than for small-ranged species for both rivers (Table 1). For large-ranged species, MDE explained about a quarter of the variation in richness along the Torne River (relative to 7% for all species) and almost half of the variation in richness along the Kalix river (Table 1). Similar to the results when all species were considered, bank area explained additional variation in richness of large-ranged species along the Torne

River and substrate heterogeneity and bank height explained additional variation along the Kalix River.

Small-ranged species were richer near the mouths and headwaters of both rivers than predicted by MDE (Fig. 1c, f) and their richness was best explained by different variables than for all species or for large-ranged species (Table 1). Not only was the MDE model a poor predictor of richness for small-ranged species, but for the Torne River, the MDE predictions entered the model, but with a negative coefficient, such that where the MDE predicted high diversity, diversity was low and vice versa, as evident from Fig. 1c and f. River bank height also entered the model for small-ranged species along the Torne River, with sampling sites that had higher banks hosting more species. Heterogeneity and substrate fineness also entered the final model for small-ranged species along the Torne River, but were negatively associated with richness. For the Kalix River, a small amount of variation was explained by ice scour, substrate heterogeneity, and bank height. Ice scour was negatively associated with richness whereas heterogeneity and bank height were positively associated with richness (Table 1).

Table 1. Model selection for candidate explanatory variables for vascular plant species richness along two Swedish rivers. Variables considered for inclusion in each model were: channel width (Width), substrate fineness (Fineness), substrate heterogeneity (Heterogeneity), percentage of riverbank showing evidence of winter ice action (Ice scour), bank (flood) height (Bank height), the total surface area of the sampling site (Bank area), and species richness as predicted by the mid-domain model (MDE). Variables that entered each model are numbered in order of entry (see text for model selection algorithm). Slope is the standardized partial regression slope of the variable in the final (best) model. For the numbered variables, R^2 is the increase in total variance explained as each variable entered the model. For the final (best) models, R^2 is the total explained variance for each final model. AIC is the Akaike Information Criterion for each model.

Dataset	River	Model	Slope	R^2	AIC
All species	Torne	1. Log (Bank area)	0.531	0.272	237.63
		2. MDE	0.273	0.074	235.09
		Final model		0.346	235.09
	Kalix	1. MDE	0.516	0.293	181.27
		2. Heterogeneity	0.614	0.184	176.28
		3. Ice scour	-0.484	0.047	175.80
4. Log (Bank height)		0.294	0.054	174.69	
Final model		0.524	174.69		
Large ranges	Torne	1. MDE	0.507	0.232	238.98
		2. Log (Bank area)	0.477	0.227	225.30
		Final model		0.459	225.30
	Kalix	1. MDE	0.446	0.477	156.74
		2. Heterogeneity	0.361	0.098	151.63
		3. Log (Bank height)	0.227	0.038	150.79
Final model			0.613	150.79	
Small ranges	Torne	1. MDE	-0.677	0.476	133.80
		2. Log (Bank height)	0.374	0.054	131.17
		3. Heterogeneity	-0.211	0.029	130.83
		4. Fineness	-0.180	0.025	130.73
	Final model		0.584	130.73	
	Kalix	1. Ice scour	-0.743	0.058	128.83
		2. Heterogeneity	0.577	0.201	125.26
		3. Log (Bank height)	0.318	0.271	124.55
		Final model		0.530	124.55

In summary, as predicted by MDE theory, the richness pattern among small-ranged species is poorly accounted for by MDE predictions (a negative relation for the Torne River, no relation for the Kalix River), whereas MDE contributes more substantially to the richness pattern among large-ranged species (26% of explained variation for the Torne River, 48% for the Kalix River). Not surprisingly, when large and small ranges are considered together, contributions from MDE are intermediate (7 and 29%, respectively).

Between-river correlations in range size

The correlation between range size of individual species along the two rivers was quite strong ($R^2 = 0.67$, $n = 210$, $p < 0001$). However, range sizes were on average smaller relative to river length on the Kalix than on the Torne River (Fig. 2).

Discussion

The riparian plant datasets for both rivers showed a mid-domain (middle-course) peak in the richness of "natural" (non-ruderal) species richness. Overall, the mid-domain effect accounted for a significant propor-

tion of the pattern of species richness along both rivers. Although quantitative patterns of richness along river gradients have not been comprehensively reviewed in the

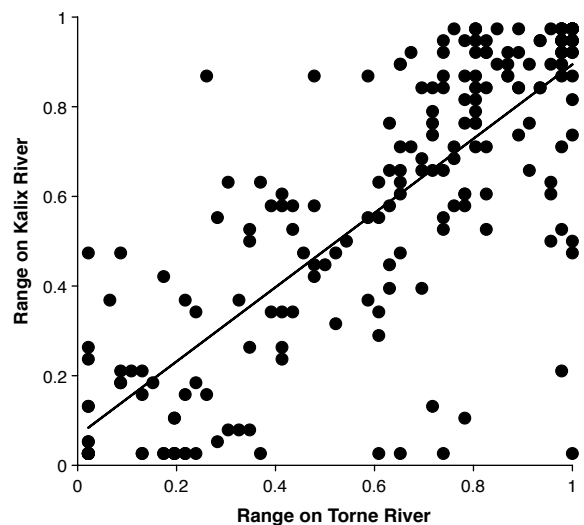


Fig. 2. Relationship between range size of riparian plant species along the Kalix and Torne Rivers. Each point represents a single species. Range sizes are measured as a proportion of the total length of the river. The figure includes all species that occurred along both rivers ($N = 210$).

literature (nor will we attempt a comprehensive review here), such a peak, long noted in studies of river ecology (Hawkes 1975, Tabacchi et al. 1996, Ward 1998), is common, at least for riparian plants, and apparently for other taxa. Nonetheless, well-documented cases also exist for other patterns (Naiman et al. 1987), including a monotonic increase in richness from headwaters to the sea (e.g. Gould and Walker 1997). Indeed, Nilsson et al. (1989) found such a monotonic pattern for ruderal plants along the Kalix and Torne rivers, in contrast to the patterns of "natural" plants treated in this study.

There exists no unifying explanation for patterns of richness along rivers, particularly for middle-course peaks in richness, that would apply equally well to aquatic organisms, such as fish, and to riverside organisms, such as riparian plants. Minshall et al. (1985) found a peak in richness of rock-dwelling insects in middle-order streams (within part of an entire river system) and attributed it to heterogeneity in thermal environment. Similarly, in a recent conceptual review of riverine richness patterns of invertebrates, Ward (1998) echoed earlier authors (e.g. Vannote et al. 1980) in highlighting the importance of the environmental heterogeneity to the middle-course peak in diversity. Peaks in plant species richness have been found in the middle courses of a variety of river systems (Tabacchi et al. 1996) and have been ascribed to changes in soil substrata (Naiman et al. 1987, Nilsson et al. 1989, Tabacchi et al. 1996), intermediate disturbance, glacial history, and at least a dozen other case-specific explanations (Naiman et al. 1987, Tabacchi et al. 1996, Ward 1998). Rarely has species richness actually been compared statistically to the factors hypothesized to be driving the patterns, particularly along the courses of entire river systems (Nilsson et al. 1989 is an exception). Several researchers have fit curves to the hump-shaped richness patterns, but it is not clear why a quadratic (Nilsson et al. 1989) or log-normal (Minshall et al. 1985) function might be expected.

Here we suggest that the null expectation for richness of any taxon limited by the river's longitudinal domain may be a peak in richness in the middle courses of the river, independent of the effects any other biological, environmental, or geomorphological gradients on the location of species along the river course. Whereas the empirical datasets examined here were for riparian plants, the geometric constraints that the headwater and mouths of rivers impose on distributions of longitudinal ranges may well be similar for other groups of aquatic organisms, as well. Such peaks can emerge even when, as we have shown, species distributions' along the river are patchy (see Connolly [2005] for theoretical support for this result). We have introduced a discrete MDE model that incorporates patchiness in modeling the ranges of species. For one-dimensional domains with (at least approximately) uniformly spaced, discrete

sampling points or sites, we believe this model is a step forward in making biogeographic null models more realistic. Of course, rivers are not unique in having species with patchy ranges, so this discrete model has potential applicability to a wide variety of domains and additional dimensions. Reproducing empirically recorded levels and patterns of patchiness in two-dimensional MDE models, on a stochastic basis, would be more complex, but could be achieved by algorithms similar to the one we used.

When all species were considered for riparian plants along the Torne and Kalix Rivers, MDE models explained a significant, but modest to moderate proportion of the longitudinal variation in species richness. Along the Torne River, bank area also contributed significantly to explaining site richness, but this factor is best viewed as a sampling effect: at sites where the bank (the area between low and high water) was larger in area, it also supported more individuals (Nilsson unpubl.), hence more species were likely to be detected (Gotelli and Colwell 2001). Along the Kalix River, on the other hand, substrate heterogeneity explained additional variation, once MDE had been accounted for as the most explanatory factor.

Our null models for species distributions along rivers are not entirely without precedent in river research. In its original form (Vannote et al. 1980), the River Continuum Concept described the peak in richness in mid-domains of rivers, taxonomically, as the area of overlap between freshwater and saltwater clades, particularly between insects (freshwater) and crustaceans (primarily saltwater). Although Vannote et al. (1980) suggested that this overlap between marine and freshwater taxa could contribute to the middle-course peak in overall richness, there is no a priori reason to expect local richness to increase in portions of domains corresponding to biotic transitions, unless those transitions happen to lie near mid-domain. On the other side of the coin, biotic mixing in middle courses is not in any way inconsistent with the mid-domain effect (see Grytnes 2003b, Colwell et al. 2004).

The prediction that mid-domain models are likely to be more explanatory for species with large rather than small ranges (reviewed by Colwell et al. 2004, 2005) is confirmed by visual inspection of the relationship between the null model predictions and correlation coefficients in our examples (Fig. 1) and by the regression models (Table 1). Small-ranged species may best show the impacts of local history and biology in shaping large-scale patterns of species richness (Jetz and Rahbek 2002, Jetz et al. 2004). Unless small-ranged species are explicitly examined separately, however, the pattern in their richness is likely to be obscured by the more widespread species, which occur in many more individual sites and thus have a disproportionate and often misleading effect on the overall pattern of species

richness (Lees et al. 1999, Jetz and Rahbek 2002, Lennon et al. 2004). The richness patterns of small-ranged species are simply more easily arranged than large-ranged species, whether by evolution or by random-placement algorithms in the models, in ways that do not generate a mid-domain peak.

Like other linear domains, including coastlines (Pielou 1977, Ellison 2002) and elevational gradients (Kessler 2001, Sanders 2002, Grytnes and Vetaas 2002), the most interesting patterns along rivers, at least for riparian plants and likely for other groups, may be those that deviate from the expected mid-domain peak in richness. In our study, not only did small-ranged species fail to show a mid-domain peak in richness, they actually showed a mid-domain trough in richness. Importantly, when Nilsson et al. (1989) considered all natural species, this trough was masked by the richness patterns of large-ranged species and not considered. We offer no conclusive explanation for this trough, but suggest that it is perhaps the most interesting pattern in richness along the two rivers considered and the most worthy of explanation.

An important, unanswered question regarding null models along rivers is whether river systems are analogous to elevational gradients or they are one and the same. For any river or river system, elevation necessarily decreases from headwaters to mouth such that the ranges of species are actually two-dimensional, with both a minimum and maximum elevation and a minimum and maximum distance from the headwaters or mouth. Because mid-domain effects can be strong along elevational gradients (Grytnes 2003a, Cardelús et al. 2006, Watkins et al. 2006), middle-course peaks in richness might result simply from the overlap in the elevational ranges of species. However, whereas a river and the elevational relief of the land it drains share a lower bound, the sea, their upper bounds must generally be quite different since no river starts on a mountaintop and the distance from river headwaters to mountain top varies among rivers. Thus if elevation and not river longitude were behind the mid-domain peaks we observed, then we would expect such peaks only when the river headwaters approach the peaks of the mountains they descend from. Nilsson et al. (1991) considered patterns along the Öre and Sävar rivers, which lie in the same region as the Torne and Kalix, but are both substantially shorter than the Torne and Kalix and consequently span less variation in elevation. In these shorter rivers, richness nonetheless peaked in mid-longitude, which at least suggests that for this study system river longitude is more important than elevation per se.

An important and persistent question in applying mid-domain null models is how to define domain limits (Colwell et al. 2004). In this, the first and admittedly preliminary application of mid-domain null models to

rivers, we have assumed that the headwaters and mouth are boundaries for riparian plants. The upper limit of the river system is likely to be a relatively complete or "hard" boundary (in the sense of Colwell and Hurtt 1994) for obligately riparian plants, but other plants recorded along riverbanks may be found in other habitats. Likewise, at the lower bound – the ocean – some principally riparian species may extend their ranges to coastal habitats. Nonetheless, as long as a substantial proportion of wide-ranged riparian plants are largely restricted to riverbank habitats, a mid-domain peak in overall richness is expected. Similar considerations apply to other riverine taxa.

Mid-domain models predict that the peak in richness for taxa bounded by the headwaters and mouth of rivers should occur in the rivers' middle courses, regardless of river length. In contrast, most other hypotheses intended to explain richness patterns along rivers tend toward predictions that the peak in richness should occur some absolute distance from the mouth or headwaters (Statzner and Higler 1985, Nilsson et al. 1991). The only study we know of that has explicitly compared rivers of different length that are otherwise similar found that peaks in richness are in the middle-courses of rivers, regardless of length (Nilsson et al. 1991), concordant with the predictions of MDE models. Many other rivers of different lengths might also be compared, with results bearing directly on the mechanisms that pattern species richness along rivers.

To sum up, we have shown that the richness patterns of riparian plants along two biogeographically linked rivers are substantially influenced by the mid-domain effect. We view this as a preliminary but important step toward a more unified understanding of the influence of MDE on richness patterns along river courses. We predict that heretofore undocumented mid-domain peaks along river courses may well exist for many taxa. Aquatic insects, riverine birds, and amphibians, all of which have lower bounds at the river mouth or further upstream, seem likely to display mid-domain peaks in rivers that lack strong, countervailing environmental drivers of species richness. Environmental factors are clearly important to the distribution of individual species along rivers, but their influence on longitudinal patterns of richness along rivers may often be modified or even dominated by MDE. River domains are in some ways more amenable to exploring ecological theory than entire continental domains. There are few continents on which to test theories of diversity gradients, but many thousands of rivers.

Acknowledgements – We are grateful to (alphabetically) Robin L. Chazdon, Jon-Arvid Grytnes, Andrew Latimer, Jennifer Martin, Juan Manuel Morales, Thiago Rangel, Tom S. Romdal, and Monica Sanchez for valuable help and discussions about this project. Thanks also to the generous soul, whose identity RKC has disgracefully forgotten, who first pointed out middle-course peaks in richness along rivers to him.

Robert Dunn was funded during this project by an NSF predoctoral fellowship, and Robert Colwell by NSFDEB-0072702.

References

- Bates, H. W. 1863. The naturalist on the River Amazons, a record of adventures, habits of animals, sketches of Brazilian and Indian life and aspects of nature under the Equator during eleven years of travel. – J. Murray.
- Cardelús, C., Colwell, R. K. and Watkins, J. E. Jr 2006. Vascular epiphyte distribution patterns: explaining the mid-elevation richness peak. – *J. Ecol.* 94: 144–156.
- Colwell, R. K. 2006. RangeModel: a Monte Carlo simulation tool for assessing geometric constraints on species richness. Version 5. – User's guide and application published at <<http://viceroy.eeb.uconn.edu/RangeModel>>.
- Colwell, R. K. and Hurtt, G. C. 1994. Nonbiological gradients in species richness and a spurious Rapoport effect. – *Am. Nat.* 144: 570–595.
- Colwell, R. K. and Lees, D. C. 2000. The mid-domain effect: geometric constraints on the geography of species richness. – *Trends Ecol. Evol.* 15: 70–76.
- Colwell, R. K., Rahbek, C. and Gotelli, N. J. 2004. The mid-domain effect and species richness patterns: what have we learned so far? – *Am. Nat.* 163: E1–E23.
- Colwell, R. K., Rahbek, C. and Gotelli, N. J. 2005. The mid-domain effect: there's a baby in the bathwater. – *Am. Nat.* 166: E149–E154.
- Connolly, S. R. 2005. Process-based models of species distributions and the mid-domain effect. – *Am. Nat.* 166: 1–11.
- Diniz-Filho, J. A. F., Bini, L. M. and Hawkins, B. A. 2003. Spatial autocorrelation and red herrings in geographical ecology. – *Global Ecol. Biogeogr.* 12: 53–64.
- Ellison, A. M. 2002. Macroecology of mangroves: large-scale patterns and processes in tropical coastal forests. – *Trees – Structure and Function* 16: 181–194.
- Gotelli, N. J. and Colwell, R. K. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. – *Ecol. Lett.* 4: 379–391.
- Gould, W. and Walker, M. 1997. Landscape-scale patterns in plant species richness along an Arctic river. – *Can. J. Bot.* 75: 1748–1765.
- Grytnes, J. A. 2003a. Species-richness patterns of vascular plants along several altitudinal transects in Norway. – *Ecography* 26: 291–300.
- Grytnes, J. A. 2003b. Ecological interpretations of the mid-domain effect. – *Ecol. Lett.* 6: 883–888.
- Grytnes, J. A. and Vetaas, O. R. 2002. Species richness and altitude: a comparison between null models and interpolated plant species richness along the Himalayan altitudinal gradient, Nepal. – *Am. Nat.* 159: 294–304.
- Hawkes, H. A. 1975. River zonation and classification. – In: Whitton, B. E. (ed.), *River ecology*. Blackwell, pp. 312–374.
- Hughes, M. H. and Parmalee, P. W. 1999. Prehistoric and modern freshwater mussel Mollusca: Bivalvia: Unioniodes faunas of the Tennessee River: Alabama, Kentucky and Tennessee. – *Regulated Rivers – Research and Management* 15: 25–42.
- Jetz, W. and Rahbek, C. 2001. Geometric constraints explain much of the species richness pattern in African birds. – *Proc. Nat. Acad. Sci. USA* 98: 5661–5666.
- Jetz, W. and Rahbek, C. 2002. Geographic range size and determinants of avian species richness. – *Science* 297: 1548–1551.
- Jetz, W., Rahbek, C. and Colwell, R. K. 2004. The coincidence of rarity and richness and the potential signature of history in centres of endemism. – *Ecol. Lett.* 7: 1180–1191.
- Kessler, M. 2001. Patterns of diversity and range size of selected plant groups along an elevational transect in the Bolivian Andes. – *Biodiv. Conserv.* 10: 1897–1921.
- Laurie, H. and Silander, J. A. 2002. Geometric constraints and spatial pattern of species richness: critique of range-based null models. – *Div. Distrib.* 8: 351–364.
- Lees, D. C., Kremen, C. and Andriamampianina, L. 1999. A null model for species richness gradients: bounded range overlap of butterflies and other rainforest endemics in Madagascar. – *Biol. J. Linn. Soc.* 67: 529–584.
- Lennon, J. J. et al. 2004. Contribution of rarity and commonness to patterns of species richness. – *Ecol. Lett.* 7: 81–87.
- Lusk, C. H., Chazdon, R. L. and Hofmann, G. 2006. A bounded null model explains juvenile tree community structure along light availability gradients in a temperate rain forest. – *Oikos* 112: 131–137.
- McCain, C. M. 2003. North American desert rodents: a test of the mid-domain effect in species richness. – *J. Mammal.* 84: 967–980.
- Minshall, G. W., Petersen, R. C. Jr and Nimz, C. F. 1985. Species richness in streams of different size from the same drainage basin. – *Am. Nat.* 125: 16–38.
- Morales, M. A., Dodge, G. J. and Inouye, D. W. 2004. A phenological mid-domain effect in flowering diversity. – *Oecologia* 142: 83–89.
- Naiman, R. J. et al. 1987. Longitudinal patterns of ecosystem processes and community structure in a subarctic river continuum. – *Ecology* 68: 1139–1156.
- Nilsson, C. et al. 1989. Patterns of plant species richness along riverbanks. – *Ecology* 70: 77–84.
- Nilsson, C. et al. 1991. Small rivers behave like large rivers: effects of postglacial history on plant species richness along riverbanks. – *J. Biogeogr.* 18: 533–541.
- Pielou, E. C. 1977. The latitudinal spans of seaweed species and their patterns of overlap. – *J. Biogeogr.* 4: 299–311.
- Pineda, J. and Caswell, H. 1998. Bathymetric species-diversity patterns and boundary constraints on vertical range distributions. – *Deep-Sea Research Part II – Topical Studies in Oceanography* 45: 83–101.
- Rangel, T. F. L. V. B., Diniz-Filho, J. A. F. and Bini, L. M. 2005. SAM version 1.0: Spatial Analysis in Macroecology (software and user's guide). – <www.ecoevol.ufg.br/sam>.
- Sanders, N. J. 2002. Elevational gradients in ant species richness: area, geometry, and Rapoport's rule. – *Ecography* 25: 25–32.
- Statzner, B. and Higler, B. 1985. Questions and comments on the river continuum concept. – *Can. J. Fish. Aquat. Sci.* 42: 1038–1044.
- Tabacchi, E. et al. 1996. Landscape structure and diversity in riparian plant communities: a longitudinal comparative study. – *Regulated Rivers – Research and Management* 12: 367–390.
- Vannote, R.L. et al. 1980. The River Continuum Concept. – *Can. J. Fish. Aquat. Sci.* 37: 130–137.
- Ward, J. V. 1998. Riverine landscapes: biodiversity patterns, disturbance regimes, and aquatic conservation. – *Biol. Conserv.* 83: 269–278.
- Watkins, J. E. Jr et al. 2006. Diversity and distribution of ferns along an elevational gradient in Costa Rica. – *Am. J. Bot.* 93: 73–83.
- Zapata, F. A., Gaston, K. J. and Chown, S. L. 2003. Mid-domain models of species richness gradients: assumptions, methods and evidence. – *J. Anim. Ecol.* 72: 677–690.

Download the appendix as file E4259 from
<www.oikos.ekol.lu.se/appendix>.

Subject Editor: Carsten Rahbek.