Species Diversity and Distribution in Presence-Absence Matrices: Mathematical Relationships and Biological Implications

Héctor T. Arita,[1,*] J. Andrés Christen,[2,†] Pilar Rodríguez,[1,‡] and Jorge Soberón[3,§]

1. Instituto de Ecología, Universidad Nacional Autónoma de México, Apartado Postal 70-275, CP 04510 Mexico City, Mexico;
2. Centro de Investigación en Matemáticas, Apartado Postal 402, CP 36000 Guanajuato, Mexico;
3. Natural History Museum, University of Kansas, Lawrence, Kansas 66045

abstract: The diversity of sites and the distribution of species are fundamental pieces in the analysis of biogeographic and macroecological questions. A link between these two variables is the correlation between the species diversity of sites and the mean range size of species occurring there. Alternatively, one could correlate the range sizes of species and the mean species diversity within those ranges. Here we show that both approaches are mirror images of the same patterns, reflecting fundamental mathematical and biological relationships. We develop a theory and analyze data for North American mammals to interpret range-diversity plots in which the species diversity of sites and the geographic range of species can be depicted simultaneously. We show that such plots contain much more information than traditional correlative approaches do, and we demonstrate that the positions of points in the plots are determined to a large extent by the average, minimum, and maximum values of range and diversity but that the dispersion of points depends on the association among species and the similitude among sites. These generalizations can be applied to biogeographic studies of diversity and distribution and in the identification of hotspots of diversity and endemism.

Keywords: beta diversity, diversity, co-occurrence, distribution, presence-absence matrix, range size.

* Corresponding author; e-mail: arita@ecologia.unam.mx.

† E-mail: jac@cimat.mx.

‡ E-mail: prodriguez@ecologia.unam.mx.

§ E-mail: jsoberon@ku.edu.

The most conspicuous pattern in biogeography and macroecology is the heterogeneous distribution of species. From local communities to continental biotas, spatial patterns of diversity determine the way in which species are distributed (Rosenzweig 1995). Simpson (1964) assessed the diversity of North American mammals by measuring the overlap of their geographic ranges on quadrats of equal size. Since then, this basic method has been applied to test ecological and evolutionary hypotheses concerning the spatial distribution of diversity on continental and global scales (Ceballos et al. 2005; Graves and Rahbek 2005; Orme et al. 2005; Hawkins and Diniz-Filho 2006).

In all of these studies, the basic analytical tool is the presence-absence matrix, which summarizes the two fundamental units of biogeography: the distributional range of species (Brown et al. 1996; Gaston 2003) and species diversity (the number of species occurring in a given site; Rosenzweig 1995). In presence-absence matrices, rows represent species, columns are localities or samples, and the elements of the matrix are binary entries representing the presence (1) or absence (0) of a given species at a given site (Gotelli 2000).

Traditionally, data in presence-absence matrices are analyzed by rows (R-mode) or by columns (Q-mode) to summarize information by species or by sites, respectively (Simberloff and Connor 1979; Legendre and Legendre 1983; Bell 2003; fig. 1A, 1B). The sum of elements along a row (R-mode) is the occupancy of a species, that is, the number of areas in which that species occurs (McGeoch and Gaston 2002). In biogeographic studies, occupancy is used to measure the range size of species over continents (Anderson and Koopman 1981; Arita et al. 1997; Gaston 2003; Graves and Rahbek 2005; Hawkins and Diniz-Filho 2006; Orme et al. 2006). Patterns in the distribution of species have been also used to ask whether general assembly rules determine the composition of natural communities (Gotelli and McCabe 2002; Horner-Devine et al.
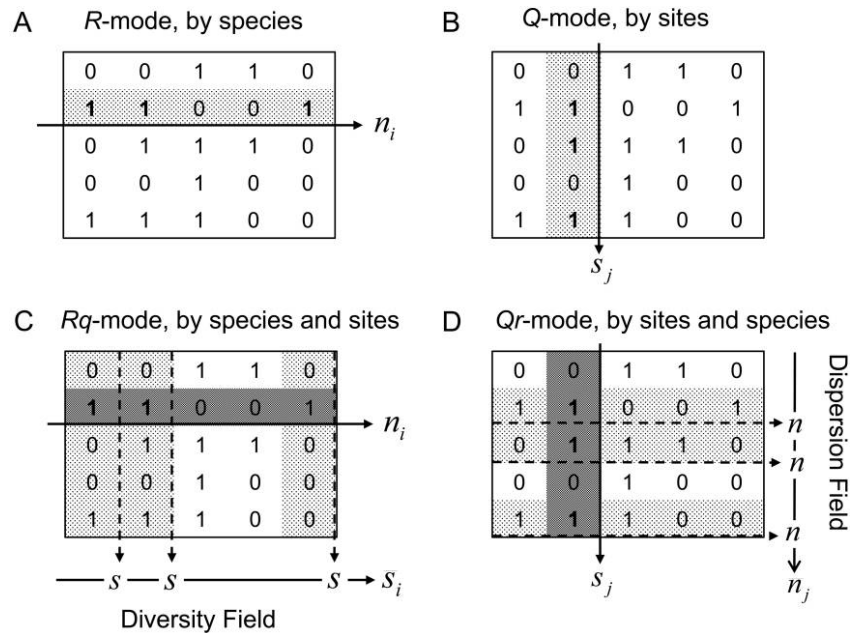
Figure 1: Diversity and distribution in presence-absence matrices. In traditional analyses (*A*, *B*), the range size of species (n_i) or the species diversity of sites (s_j) are computed by adding the elements along rows (*R*-mode) or columns (*Q*-mode), respectively. In the *Rq*-mode, the analysis by species is enhanced by incorporating the species diversity (column totals) of sites in which the focal species occurs to generate the diversity field. Then, the mean diversity of those sites can be calculated ($\bar{s}_i$ in *C*). In the *Qr*-mode, the analysis by sites is enriched by considering the range size (row totals) of species occurring in the focal site to generate the dispersion field. Then, the mean range size of those species can be calculated ($\bar{n}_j$ in *D*). Mathematical notation follows that outlined in table 1.

2007). Since Connor and Simberloff (1979) showed that many of the bird distribution patterns that Diamond (1975) had explained in terms of interspecific competition could be reproduced using Monte Carlo null models of presence-absence matrices, the idea of general assembly rules has been hotly debated. Modern tests have been based on the analysis of the co-occurrence of species using presence-absence data (Stone and Roberts 1990; Gotelli 2000; Gotelli and McCabe 2002).

The sum along a column (*Q*-mode) is the species diversity of a site, a parameter that has been used to test ecological and evolutionary hypotheses concerning the determinants of biological diversity (Hawkins et al. 2003; Hillebrand 2004; Kalmar and Currie 2007; Mittelbach et al. 2007). The count of species has also received considerable attention in identifying hotspots, that is, sites with unusual levels of species diversity (Myers et al. 2000; Roberts et al. 2002; Ceballos et al. 2005; Orme et al. 2005).

Since the rise of null models as analytical tools to extract information from presence-absence matrices, the importance of considering marginal totals of both the rows and the columns has been stressed (Connor and Simberloff 1979; Gotelli and Graves 1996; Gotelli 2000). Since then, several algorithms have been proposed to randomize the

elements of presence-absence matrices while conserving the row and column totals (review in Gotelli 2000). Despite this, most studies still focus on either the distribution of species or the diversity of sites, ignoring the interaction between rows and columns in presence-absence matrices.

One way of examining the whole structure of presence-absence matrices is by analyzing their degree of nestedness. Perfectly nested sets, in which species that form less diverse biotas also occur in more diverse ones (Patterson and Atmar 1986; Patterson 1987), generate triangular submatrices in presence-absence matrices. Although most assemblages of natural species seem to be highly nested, there is still much debate on how to measure nestedness and on the biological mechanisms that generate such striking patterns (Greve and Chown 2006; Ulrich and Gotelli 2007*b*).

Another way of considering the interaction of rows and columns is by simultaneously measuring distribution and diversity in analyses by sites or by species in what can be called the *Rq*-mode or the *Qr*-mode (fig. 1*C*, 1*D*). This is done by sites (*Qr*-mode), for example, when examining the continental variation in species' geographic range sizes (Hawkins and Diniz-Filho 2006). This pattern is generally assessed by calculating the mean or median range size of

Table 1: Parameters of range size and diversity used in this article.

Symbol	Parameter
Parameters of the region:	
S	Total species diversity in the region
N	Total number of sites in which the region is divided
β	Whittaker's β diversity: the ratio of S and the average diversity among sites in the region
f	The fill of the matrix, the total number of species occurrences
f^*	The proportional fill of the matrix
Parameters of species:	
n_i	The range size of species i, measured as the number of sites in which it occurs
$\bar{n}$	The average range size of the S species occurring in the region
n_i^*	The range size of species i as a proportion of N
$\bar{n}^*$	The average proportional range size of the S species occurring in the region
$\bar{s}_i^*$	The average proportional species diversity of sites in which species i occurs
ρ_i	The average covariance of species i with all species in the region
Parameters of sites:	
s_j	The species diversity of site j measured as the number of species occurring there
$\bar{s}$	The average species diversity of the N sites forming the region
s_j^*	The species diversity of site j as a proportion of S
$\bar{s}^*$	The average proportional species diversity of the N sites forming the region
$\bar{n}_j^*$	The average proportional range size of species occurring in site j
τ_j	The average covariance of site j with all sites

a species occurring at a given site (Lutz 1921; Rapoport 1975; Anderson and Koopman 1981; Orme et al. 2006). Graves and Rahbek (2005) have extended the concept by examining the traits of a "dispersion field," which is the set of geographic ranges of species occurring in a given site. The equivalent analysis by species (*Rq*-mode) quantifies the species diversity of all sites in which a particular species occurs, generating a "diversity field," which is analogous to Diamond's (1975) incidence functions.

Here we examine theoretical aspects of the interaction between rows and columns in presence-absence matrices. Our analyses are not based solely on row or column counts, as traditional approaches are (fig. 1*A*, 1*B*), but on more complex quantities that take into account information from rows and from columns: the per-site mean range size and the mean diversity within species ranges (fig. 1*C*, 1*D*). We introduce range-diversity plots as a way

to display these new variables along with the traditional parameters (species diversity and range size). We show that the intrinsic structure of presence-absence matrices determines mathematical and biological constraints that limit the dispersion of possible points in range-diversity plots, and we interpret such dispersion in terms of the degree of association among species and similarity among sites. Finally, we demonstrate that the dispersion of points is constrained, but not totally determined, by the fill of the presence-absence matrix.

Theoretical Framework

Species Diversity and Distribution in Presence-Absence Matrices

Imagine a region or continent divided into N sites of equal area and containing S species (see table 1 for mathematical

notation). The distribution of the S species among the N sites is represented by a presence-absence matrix whose elements are $\delta(i, j) = 1$ if species i is present in site j and are $\delta(i, j) = 0$ otherwise. The range size of species i is the sum of the elements of row i, $n_i = \sum_{j=1}^{N} \delta(i, j)$. The proportional range size is the fraction of the total number of sites in the region in which species i occurs ($n_i^* = n_i/N$). The species diversity of site j is the column sum $s_j = \sum_{i=1}^{S} \delta(i, j)$, and $s_j^* = s_j/S$ is the proportional species diversity of site j.

Let $\bar{s}$ be the average species diversity among all sites and $\bar{n}$ be the average range size among all species in the region:

$$\bar{s} = \frac{1}{N}\sum_{j=1}^{N} s_j = \frac{1}{N}\sum_{j=1}^{N}\sum_{i=1}^{S} \delta(i, j),$$

$$\bar{n} = \frac{1}{S}\sum_{i=1}^{S} n_i = \frac{1}{S}\sum_{i=1}^{S}\sum_{j=1}^{N} \delta(i, j). \quad (1)$$

Similarly, let $\bar{s}^* = \bar{s}/S$ be the average proportional species diversity and let $\bar{n}^* = \bar{n}/N$ be the average proportional range size. Let the fill of the matrix be the total number of all species occurrences ($f = \sum_{j=1}^{N} s_j = \sum_{i=1}^{S} n_i$) and let $f^* = f/(NS)$ be its proportional value relative to the dimension of the matrix. Then, from equation (1), $f^* = \bar{s}^* = \bar{n}^*$ and, consequently, $\bar{s}/\bar{n} = S/N$; that is, the ratio of the average species diversity among all sites to the average range size among all species is equal to the ratio of the species diversity of the whole region to the number of sites in which the region is subdivided.

Whittaker (1972) defined beta diversity as the ratio between the regional or gamma diversity and the average local or alpha diversity. With our notation, $\beta = S/\bar{s} = (\bar{s}^*)^{-1}$ or $\beta = (\bar{n}^*)^{-1} = N/\bar{n}$. Therefore, Whittaker's beta diversity can be interpreted in two complementary ways (Whittaker 1972; Routledge 1977): (1) as the inverse of the average species diversity measured as a proportion of the total diversity of the region, $(\bar{s}^*)^{-1}$, or (2) as the inverse of the average range of species, measured as a proportion of the total area of the region, $(\bar{n}^*)^{-1}$. Note also that $\beta = (f^*)^{-1}$.

Analyses Based on Species (Rq-Mode)

We define the "diversity field" of species i as the set of diversity values of sites in which species i occurs. Let D_i be the diversity field volume, that is, the summation of those species diversity values:

$$D_i = \sum_{j=1}^{N} \delta(i, j)s_j = \sum_{l=1}^{S}\sum_{j=1}^{N} \delta(i, j)\delta(l, j). \quad (2)$$

Thus, $d_i = D_i/n_i$ is the average species diversity in the sites in which species i occurs and $\bar{s}_i^* = d_i/S$ is the mean species diversity within the range of species i, as a proportion of the total number of species in the region.

The covariance between rows i and l of the presence-absence matrix is

$$c(i, l) = \frac{1}{N}\sum_{j=1}^{N} \delta(i, j)\delta(l, j) - \frac{n_i}{N}\frac{n_l}{N}. \quad (3)$$

Note that $\sum_{j=1}^{N} \delta(i, j)\delta(l, j)/N$ is the proportion of sites in which species i and l co-occur and that n_i/N and n_l/N are the proportion of sites in which species i and l occur, respectively. Note also that $n_i n_l/N$ is the expected number of sites in which species i and l would co-occur under the null hypothesis of independent distribution. Therefore, $c(i, l)$ measures the association between species i and l by contrasting the observed and expected proportions of sites in which those species co-occur under the assumption of no association (see Bell 2005 for an alternate way to derive the same expression).

Now let ρ_i be the average covariance of species i, that is,

$$\rho_i = \frac{1}{S}\sum_{l=1}^{S} c(i, l) = \frac{1}{SN}\sum_{l=1}^{S}\sum_{j=1}^{N} \delta(i, j)\delta(l, j) - \frac{n_i}{SN}\sum_{l=1}^{S}\frac{n_i}{N}. \quad (4)$$

It is clear from equations (2) and (3) that

$$\rho_i = \frac{D_i}{SN} - \frac{n_i}{SN}\frac{1}{N}\sum_{l=1}^{S}\sum_{j=1}^{N} \delta(l, j) = \frac{D_i}{SN} - \frac{n_i}{SN}\bar{s} = \frac{n_i(d_i - \bar{s})}{SN} \quad (5)$$

and

$$\rho_i = n_i^*(\bar{s}_i^* - \bar{s}^*) \quad \text{or} \quad n_i^* = \frac{\rho_i}{\bar{s}_i^* - \bar{s}^*}. \quad (6)$$

Thus, there is an inverse relationship between the proportional range of a species and the difference between the mean proportional diversity within its range and the average proportional diversity in the whole region. This relationship is determined by the average covariance of species i, that is, its average association with all species. Therefore, analyzing the diversity field within the range of a species is equivalent to studying its covariance with all the species.

In a scatterplot of n_i^* versus $\bar{s}_i^*$, which we call a range-diversity plot, points corresponding to species of equal average covariance fall along hyperbolic curves following

equation (6), and their positions depend on the average proportional species diversity in the region (fig. 2). In these plots, boundaries for the cloud of points can be predicted using a combination of mathematical and biological restrictions. In theory, the interval of possible proportional range sizes (n_i^*) goes from $1/N$ (a species found at only one site) to 1 (a species found at all sites). Similarly, the proportional mean diversity ($\bar{s}_i^*$) could take any value from $1/S$ (species i is the only one found within its own range) to 1 (all species are found in all sites where species i occurs). These four points define a rectangle within which all points of the scatterplot must be contained.

However, mathematical relationships and biological properties define more restrictive boundaries for the cloud of points. For example, if a species occurs in all sites ($n_i^* = 1$), then the proportional diversity within its range has to be, by definition, $\bar{s}^*$, the average proportional diversity among sites of the region. At the other extreme, if a species is found in only one site ($n_i^* = 1/N$), values for $\bar{s}_i^*$ could, in principle, take any value between $1/S$ and 1. These three points define a triangle within the original rectangular boundaries.

In real communities, ecological and historical processes determine minimum and maximum values for the proportional range of a species (n_{min}^* and n_{max}^*) and for the proportional number of species that coexist in a given site (s_{min}^* and s_{max}^*). These four parameters determine mathematical upper boundaries for the values of range size, imposed by the close relationship between range size and species diversity (fig. 3). For values of $s_{min}^* \leq \bar{s}_i^* \leq \bar{s}^*$, a left upper boundary is determined by $\bar{s}^*$ and s_{max}^*. For values of $s^* \leq \bar{s}_i^* \leq \bar{s}_{max}^*$, a right upper boundary is imposed by $\bar{s}^*$ and s_{min}^*. It is important to emphasize that, while s_i^* is a trait of species i, $\bar{s}^*$, s_{min}^*, and s_{max}^* are properties of the whole assemblage (table 1). Thus, possible values for the proportional mean diversity within the range of a given species are determined by the average, minimum, and maximum values of proportional species diversity among sites. These boundaries are illustrated in figure 3 and are fully developed in the appendix.

Analyses Based on Sites (Qr-Mode)

The derivations based on sites are mirror images of those based on species, and analyses that are analogous to those presented in the previous section can be applied directly to the analyses based on sites if the matrix is transposed. Let R_j be the addition of the range sizes of the species occurring in site j:

$$R_j = \sum_{i=1}^{S} \delta(i, j)n_i = \sum_{m=1}^{N} \sum_{i=1}^{S} \delta(i, j)\delta(i, m), \qquad (7)$$

which is equal to the dispersion field volume (Graves and Rahbek 2005). Thus, $r_j = R_j/s_j$ is the mean range of species
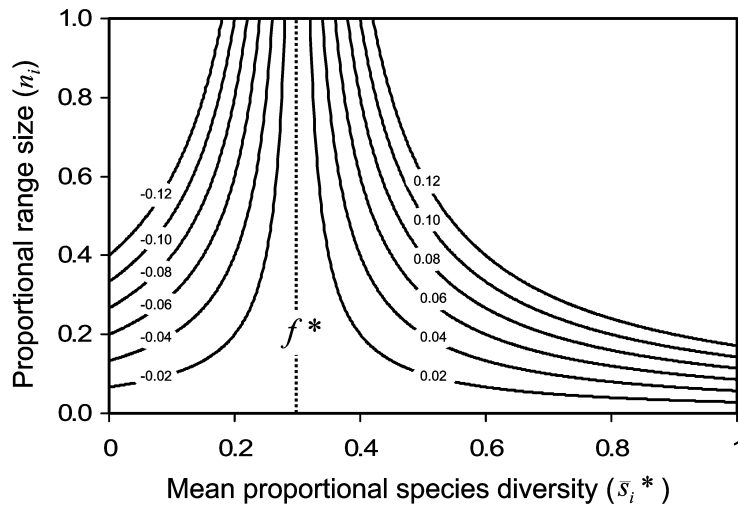

Figure 2: Isocovariance lines in range-diversity plots by species. When the proportional range size of a species versus the average proportional diversity within its range is plotted, the position of points depends on the average covariance of each species, that is, on its average association with all species. Points in the graph corresponding to species with the same average covariance arrange along isocovariance lines, such as those shown for different values of average covariance. The vertical dashed line is the zero-covariance isoline, corresponding to the average proportional species diversity among sites in the region. Identical lines can be drawn for plots by sites, in which points corresponding to sites with the same average covariance array along isocovariance lines.
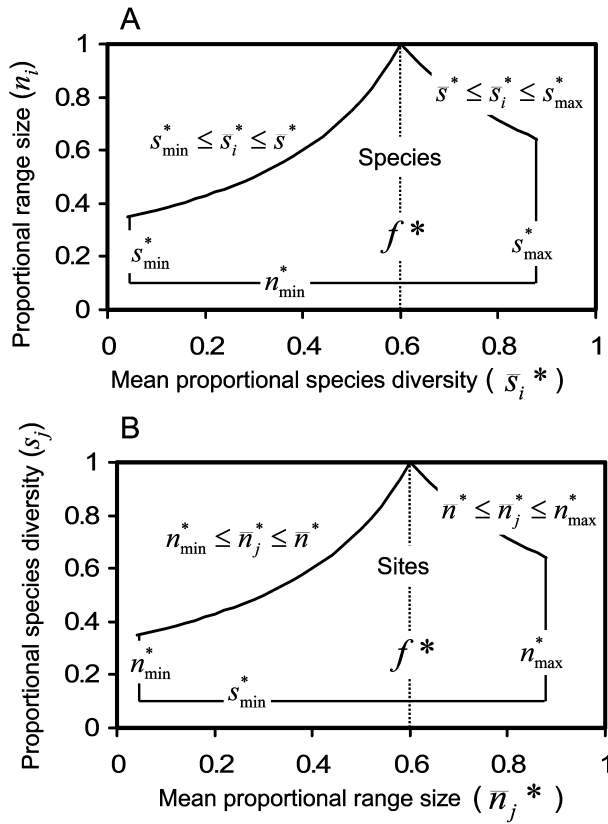
Figure 3: Mathematical and biological boundaries for the clouds of points in range-diversity plots. In plots by species (*A*), feasible points are determined by the minimum, average, and maximum values of proportional species diversity of sites. Mathematical boundaries demonstrated in the appendix are shown by the curved lines, which establish an upper limit for the cloud of points. Similar limits apply for sites (*B*). Mathematical notation follows that outlined in table 1.

occurring in site j and $\bar{n}_j^* = r_j/N$ is that average in proportion to the total number of sites of the region. Also, the covariance between columns j and m is

$$c(j, m) = \frac{1}{S} \sum_{i=1}^{S} \delta(i, j)\delta(i, m) - \frac{s_j}{S}\frac{s_m}{S}. \quad (8)$$

In this case, the covariance can be interpreted as an index of similarity between sites. Following the same argument as in the analysis by species, $c(j, m)$ is equal to the difference between the observed and the expected proportion of species that are found simultaneously in sites j and m.

Now let τ_j be the average covariance of site j with all the sites in the region. Equivalent to equation (6), we obtain

$$\tau_j = s_j^*(\bar{n}_j^* - \bar{n}^*) \text{ or } s_j^* = \frac{\tau_j}{\bar{n}_j^* - \bar{n}^*}. \quad (9)$$

Thus, the proportional diversity of a site is inversely proportional to the difference between its proportional mean range and the average proportional range of all species in the region.

A second type of range-diversity plot can be created for sites by plotting their proportional species diversities versus the proportional mean range sizes (s_j^* vs. $\bar{n}_j^*$). These plots have biological and mathematical boundaries that correspond directly to those described for species. In this case, the range of possible values of proportional species diversity for a given site is determined by the average, minimum, and maximum values of proportional range size among species in the assemblage (fig. 3; appendix). Points corresponding to sites of equal covariance form hyperbolic lines following equation (9). Therefore, comparing the range size of species occurring in a site is equivalent to studying the covariance of that site with all the areas in the region.

Range-Diversity Plots

Range-diversity plots capture a great deal of the information contained in presence-absence matrices, going well beyond the simple *R*-mode or *Q*-mode approaches that are used to describe patterns of range size or species diversity alone. Some authors, knowingly or not, have used a *Qr*-mode approach to examine the range size of species occurring at a given site but have not realized the implications of the close link between diversity and distribution (Rosenzweig 1975; Anderson and Koopman 1981; Graves and Rahbek 2005; Orme et al. 2006). Analyses based on the *Rq*-mode approach examining the pattern of species diversity within geographic ranges have not been published. Both types of relationships are mirror representations of the information contained in the structure of the presence-absence matrix that characterizes the fauna or flora of a region. Together they provide a powerful new analytical tool that allows for the visualization of much more information than traditional approaches do (fig. 1).

From the theoretical development, some generalizations can be advanced for range-diversity plots. First, although the average proportional range size of species and the average proportional species diversity of sites must have the same value ($\bar{n}^* = \bar{s}^* = f^*$), the shape of the statistical frequency distributions of the two variables can be independent of each other. Second, the value of f^* determines, in both types of plots, the position of the vertical line of zero covariance and thus the general location of the cloud of points in the plots. When no statistical association exists,

points follow these vertical lines. Additionally, points have to be distributed within the biological and mathematical limits shown in figure 3. Moreover, within those boundaries, along the ordinates the plots show the original frequency distribution of range sizes (in plots by species) and of species diversity (in plots by sites). Finally, the dispersion of points along abscissas is determined by the degree of association between species (patterns of co-occurrence) or between sites (similitude) as measured by the covariance, which follows hyperbolic lines (fig. 2).

These generalizations enable interpretation of range-diversity plots from a perspective that goes beyond the simple correlation between the variables. Rosenzweig (1975) reported a negative correlation between mean range size and the number of co-occurring North American bat and turtle species. He speculated that, in more diverse sites, species should segregate by habitat use and thus would tend to have smaller ranges. Rosenzweig's plot, however, shows a nonlinear relationship between the variables that seems to follow our isocovariance lines. Although not reported by Rosenzweig (1975), it can be inferred from the plot that the average range of all bat species considered in the study should be approximately three million mi^2. In our range-diversity plots, points for sites would be to the right of this mean value (above in Rosenzweig's [1975] inverted plot), showing that most sites show positive covariance (association) with all the sites.

For larger databases of birds, range-diversity plots by sites show considerable complexity, even on logarithmic scales. Graves and Rahbek (2005) reported a weak positive correlation between median range size and bird species diversity for 1° quadrats in South America, whereas Orme et al. (2006) documented a negative correlation for the global fauna of birds and weak negative or positive correlations for individual continents. In both cases, plots showed considerable scattering that could not be interpreted with simple correlation analyses. In fact, even on a logarithmic scale, plots show patterns that seem to follow the position of isocovariance lines (fig. 2) and mathematical boundaries (fig. 3). The two cases are examples of large faunas distributed over large continental masses. At this scale, most groups show a highly right-skewed frequency distribution of range sizes, with a majority of species having small geographic ranges and very few covering a substantial part of the continent (Gaston 2003). However, at smaller scales and for different taxa, other kinds of frequency distributions are possible, responding to a variety of environmental and biological factors (McGeoch and Gaston 2002; Mora and Robertson 2005). In fact, this variation in frequency distributions determines different patterns in the scaling of species diversity (Arita and Rodriguez 2002). Therefore, if both diversity and distribution vary with scale, we should expect different forms of range-diversity plots for different combinations of taxa and scale. This variety of forms can be explored by examining two extreme cases, as follows.

High and Low Beta Diversity Regions

Taking advantage of the relationship $\beta = (f^*)^{-1}$, regions for which the fill of the presence-absence matrix is low can be characterized as high beta diversity areas. For these regions, we would expect range-diversity plots that are similar to those pictured in figure 4*A* and 4*C*. Because of the small f^* value, points corresponding to species with large ranges could be accommodated only in the left sector of the plot, so sites within the range of widespread species would tend to have comparatively low levels of species diversity. Similarly, sites with high species diversity would have low values of average proportional range size. In contrast, restricted species would show a wide variation in mean diversity, and species-poor sites would show a similar variation in average range size. As a consequence, a negative correlation between range size and species diversity is expected in this case, both for species and for sites.

However, even if such correlations exist, it does not necessarily imply a complete congruence between hotspots of diversity and hotspots of rare (restricted) taxa. Species occurring in sites that have extremely high species diversity (diversity hotspots) would have only an average range size, although sites of high (but not the highest) diversity would indeed tend to contain restricted species (fig. 4*B*). Species with range sizes below average (restricted or rare taxa) can have, in theory, any value of proportional range diversity (fig. 4*A*). Therefore, it is possible to have a low level of congruence between criteria of prioritization of conservation of areas and species, even with a significant correlation between the two variables.

When most species of a region are widespread, f^* is high and beta diversity across the region is low. Under these circumstances, very restricted species can acquire, in principle, any average value of diversity within its range (fig. 4*B*), but very widespread species are constrained to have high values of mean diversity within their ranges, that is, to occur in sites of comparatively high diversity. Diversity hotspots are constrained in this case to contain, on average, species with very wide ranges (fig. 4*D*). Sites with low species diversity, in contrast, can acquire any value of mean proportional range size within the permissible limits. In this case, a general positive correlation between range size and species diversity can be predicted, both among species and among sites. Therefore, little congruence could be expected for priority sites defined on the criteria of diversity and rarity.

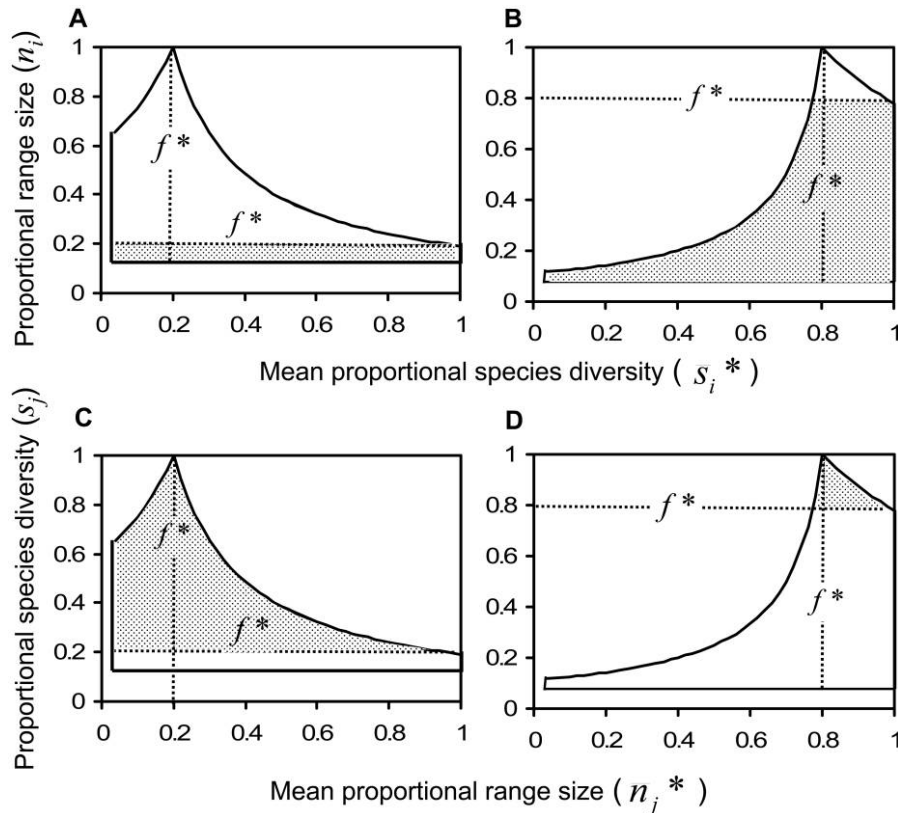Patterns of low beta diversity can be expected for vagile

Figure 4: Range-diversity plots by species (*A*, *B*) and by sites (*C*, *D*) when the mean proportional range size and the mean proportional species diversity are low (0.2; *A*, *C*) and high (0.8; *B*, *D*). *Shaded areas* = the possible dispersion of points corresponding to species with a range size that is lower than average or to sites with diversity that is higher than average. Notice that, in all cases, $f^* = \bar{n}^* = \bar{s}^* = \beta^{-1}$.

species in small continents or regions. Therefore, two general predictions can be posited: (1) for the same kind of organisms, the correlation between range size and species diversity should be stronger (more negative) the larger the area under study and (2) for the same continent or region, the correlation should be stronger for amphibians, reptiles, and small mammals than for birds, bats, and large mammals. As a consequence, the congruence between hotspots of diversity and rarity should be lower for the more vagile taxa in small continents and regions. For birds, Orme et al. (2006) reported a negative correlation between species diversity and mean range size at the global scale but positive correlations in three out of eight continents. However, despite the negative correlation at the global scale, there is little congruence between hotspots of diversity and hotspots of rare species (Orme et al. 2005).

Diversity, Rarity, and Nested Subsets

Range-diversity plots help to explain empirical analyses that have dissected the relative contribution of common

(widespread), rare (restricted), and endemic (restricted to a single site) species to spatial patterns of species diversity (Lennon et al. 2004; Greve and Chown 2006). Lennon et al. (2004) found a higher coincidence of diversity patterns when considering widespread species than when selecting only the restricted ones. Our results show that average diversity within the range of widespread species is more likely to be close to the average diversity of the whole region than average diversity would be within the range of restricted species. Lennon et al. (2004) showed that artificial communities can be constructed with a high coincidence of species diversity for the whole region and for the range of restricted species. Our analysis shows that this pattern is feasible when there is a high degree of association among species (co-occurrence) and among sites (similitude). In fact, the assemblage presented by Lennon et al. (2004) is a highly nested one. Perfectly nested subsets arrange in range-diversity plots along straight lines, with slope = −2, that cross the point ($\bar{s}^*$, 1) in plots by species or the point ($\bar{n}^*$, 1) in plots by sites (fig. 5). The position of points in the right sector of plots is evidence of the high
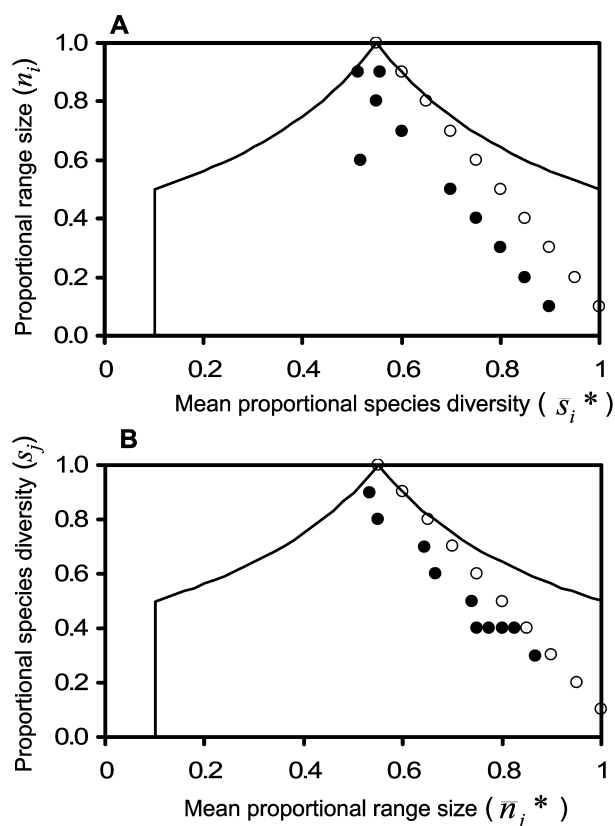
Figure 5: Range-diversity plots for species (*A*) and for sites (*B*) of a perfectly nested assemblage of 10 species in 10 sites (*open circles*). Species 1 occurs in site 1, species 2 occurs in sites 1 and 2, and so on, with species 10 occurring in all sites. Points would arrange along straight lines. *Solid circles*, the points corresponding to the example in Lennon et al. (2004), discussed in the main text.

level of association among species and among sites in perfectly nested sets.

Greve and Chown (2006) documented the effect of adding endemic species to perfectly or partially nested sets of species. In our example in figure 5, the addition of a new species to a single site would have totally different effects depending on the identity of the site. In all cases, by definition, the range size of the new species would be $n_{new} = 1$ and $n_{new}^* = 1/10$, so its point would be located at the bottom of figure 5A. If the new species is added to the richest site, then its proportional range diversity would be $\bar{s}_{new}^* = 1$ and its point would be at the lower right corner of the plot, thus contributing to the general nested pattern. If the species is added to the poorest site, however, $\bar{s}_{new}^* = 2/11$ and its point would lay at the lower left corner of the plot, totally disrupting the nested pattern. In the plot by sites, the first case would generate a point located exactly at the top of the plot because the richest site would

still contain all species. In the second case, in contrast, the poorest site would also have a very low mean range size, so its point would lay in the lower left part of the plot.

From a more general perspective, it can be shown that the mean species diversity within the range of a species is also a measure of nestedness; $R_{i,l} = \sum_{j=1}^{N} \delta(i, j)\delta(l, j)$ is the number of sites shared between species i and l (see also equation [3]), and $R_{i,l}/n_i$ is the proportion of sites occupied by species i inside the range of species l, which would be equal to 1 if the range of species i is fully nested within the range of l, and would equal 0 if there is no overlap of the ranges. Thus, the proportion can be considered to be a measure of nestedness of i within l. Therefore, $1/S\sum_{l=1}^{S} R_{i,l}/n_i$, the average proportion of overlap of species i with all species, is also the average nestedness of i.

The diversity field volume of species i, which is the summation of species diversity values of sites within its range, can also be computed as $D_i = \sum_{l=1}^{S} R_{i,l}$ because this last figure is the summation of overlaps of all species with species i. Therefore,

$$\frac{1}{S}\sum_{l=1}^{S}\frac{R_{i,l}}{n_i} = \frac{D_i}{Sn_i} = \frac{d_i}{S} = \bar{s}_i^*. \tag{10}$$

In other words, the average proportional species diversity within the range of species i equals the mean overlap of the range of i with the range all species and thus can be seen also as a measure of average nestedness of the range of i within the range of all species.

Moreover, because $\bar{s}^* = f^*$ (the proportional fill of the matrix), the degree of nestedness (or nonnestedness) of species i may be observed by the departure of $\bar{s}_i$ from $f^* = \bar{s}^*$. This is because f^* is the expected value of $\bar{s}_i^*$ under the null hypothesis of no association between species i and the rest of the species. Exactly the same mathematical reasoning can be applied by sites. Therefore, range-diversity plots, by species or by sites, that form clouds of points to the right of f^* will show some degree of nestedness, and plots with points arranged along the extreme right of the graph (such as those in fig. 5) will be evidence of highly nested assemblages.

Because range-diversity plots display information based both on distribution by species and on diversity by sites, they can show a more complete picture of patterns of nestedness than can measures based on single parameters, such as C, which measures the number of cases in which the occurrence of a species at a given site correctly predicts its presence at a richer site (Wright and Reeves 1992), or the temperature T, which measures the degree of disorder in the matrix (Atmar and Patterson 1993). Additionally, our mathematical derivation shows that nestedness implies association among species and similitude among sites, but the contrary is not necessarily true. A similar conclusion,

that nestedness and co-occurrence of species do not necessarily correlate in species assemblages, has recently been reached using numerical simulations (Ulrich and Gotelli 2007*a*).

Diversity and Distribution in North American Mammals

We explored the implications of the mathematical relationships developed in the previous sections with data for the 744 species of terrestrial North American mammals, from Alaska and Canada to Panama (see Arita et al. 2005 for details of the database). A grid of equal-area quadrats (2,500 km^2 each) was overlaid on distributional maps for every species, generating a 744 species × 13,195 site presence-absence matrix.

For the mammalian fauna of North America, the mean proportional range size and mean proportional species diversity are $f^* = 0.053$, which implies a very high beta diversity ($\beta = 18.87$). Thus, range-diversity plots are similar to those in figure 4*A* and 4*C* (fig. 6). Proportional range sizes for species varied from 7.58×10^{-5} (a species occurring in only one quadrat) to 0.845 (a species whose range covers 84.5% of the continent). Proportional species diversity for sites varied from 0.0013 to 0.259, meaning that a given site harbors at most 25.9% of the total diversity of the continent. These maxima and minima determined the dispersion of points in range-diversity plots as expected from boundaries shown in figure 3.

In the plot by species (fig. 6*A*), there was a significant negative linear correlation between range size and mean species diversity among species (-0.413). A closer look, however, reveals that the dispersion of points can be better explained in terms of the mathematical constraints rather than with correlations. Points located to the extreme left and the extreme right follow the lines of covariance -0.01 and 0.01, respectively. Because the value of f^* is very low (0.053), the absolute boundary for mean diversity to the left (0.0013) is closer to the mean than the limit to the right (0.259) is, producing an asymmetric arrangement of points with a long right-hand tail. Points were concentrated on the right side of the plot, revealing an overall positive association among species. Because of the mathematical constraints, species with positive covariance arranged in such a way that they formed a long tail along the *X*-axis. Species here are taxa with small ranges that occur at high-diversity sites, mostly tropical small mammals. Widespread species are arranged according to the predictions of the mathematical model, following the vertical line where $\bar{s}_i^* = f^*$, so the average diversity inside their range is close to the average for the whole continent. Species with $\bar{s}_i^* < f^*$ are mostly Nearctic taxa with medium-sized or small ranges and that occur in low-diversity areas of North America.

The pattern in figure 6*A* is generated by the combined effect of the latitudinal gradient in species diversity, in which tropical areas support higher levels of diversity than temperate zones do (Willig et al. 2003; Mittelbach et al. 2007), and Rapoport's rule, whereby species occurring at lower latitudes tend to have smaller range sizes (Rapoport 1975; Stevens 1989). Both effects are very strong for the
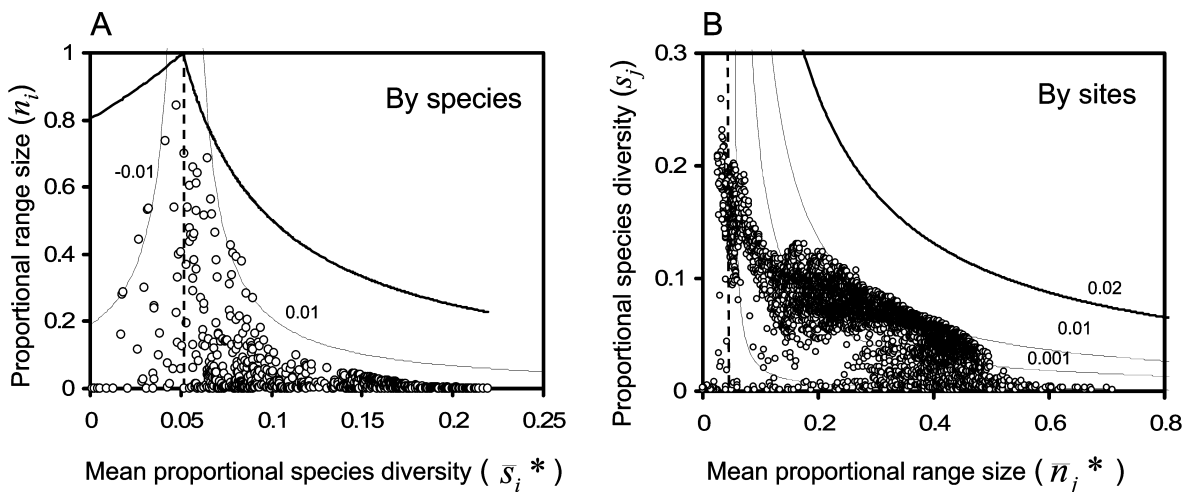

Figure 6: Range-diversity plots for North American mammalian fauna, by species (*A*) and by sites (*B*). *Thick lines* = absolute boundaries depicted in figure 3. *Thin lines* = isocovariance curves as shown in figure 2. *Vertical dashed lines* = average values of proportional range size and species diversity.

mammals of North America (Arita et al. 2005; Ruggiero and Werenkraut 2007).

The same effects are evident in the plot by sites (fig. 6*B*). In this case, the maximum value for range size is high ($n_{max} = 0.845$), which, coupled with the low f^* value, generates a highly asymmetrical dispersion of points. Because of the latitudinal gradient of diversity, tropical sites tend to arrange in the top sector of the graph and temperate sites are at the bottom. As a consequence of Rapoport's effect, low-diversity sites tend to support species with larger ranges, so they lay in the right side of the plot. Because of mathematical constraints, high-diversity sites tend to accommodate along or close to the line where $\bar{n}_j^* = f^*$. All these patterns generate an overall negative correlation between species diversity and per-site range size (-0.791).

Range-diversity plots for North American mammals also provide information for conservation-oriented analyses. Rare (restricted) species show a wide variation in the number of species occurring inside their ranges (fig. 6*A*). However, most of the restricted North American mammal species occur in high-diversity sites, as shown by the points along the right-hand tail in figure 6*A*. This means that, in general, sites designated as protected areas because of their high level of diversity will also contain rare species. This fact is corroborated in the plot by sites, where high-diversity sites tend to support species with small ranges (fig. 6*B*). Therefore, hotspots of diversity, in general, coincide with hotspots of rare (restricted) species (Arita et al. 1997).

Conclusions

In recent years, the old idea of studying geographic patterns of diversity by overlapping the ranges of species (Simpson 1964) has been considerably enhanced by the availability of very large data sets and by the development of new analytical tools that take advantage of modern computer power (Ceballos et al. 2005; Graves and Rahbek 2005; Orme et al. 2005, 2006). In this article, we stress the fact that presence-absence matrices, which for a long time have been essential tools in the description of local-scale ecological communities (Pielou 1984; Legendre and Legendre 1983) or discrete species assemblages in archipelagos (Connor and Simberloff 1979; Gotelli 2000), also provide a natural ground for understanding the relations among a variety of indicators of biological diversity at the continental scale.

Because $\beta = (f^*)^{-1}$, Whittaker's beta diversity depends directly on the total number of occurrences of species among all sites. We have shown that f^*, and therefore β, determines the general shape and position of the permissible clouds of points in range-diversity plots. We have

demonstrated, however, that Whittaker's beta diversity alone in fact contains very little information about the internal structure of the matrix because its value is invariant to permutations of presences and absences that maintain constant the dimension and the fill of a matrix (Lira-Noriega et al. 2007). For example, Whittaker's beta diversity does not change if the ranges of species are displaced, deformed, or even fragmented as long as the number of species and the summation of range sizes do not change. The fact that such drastic transformations in the structure of presence-absence matrices do not affect Whittaker's beta diversity is an arithmetic corollary of its definition, but it also suggests the need for more sensitive descriptors of the structure of the matrix.

Our *Rq*-mode and *Qr*-mode analyses, whereby the covariances of sites in terms of composition and of species in terms of range are incorporated, capture far more information regarding the species assemblages, opening new possibilities for richer interpretations of biological diversity across regions and taxa. However, these analyses also reveal that f^*, and consequently Whittaker's beta diversity, plays a central role in setting the position of the mathematical boundaries in range-diversity plots, even affecting the likelihood of positive or negative correlations between range and diversity and the possible congruence of hotspots based on diversity and rarity. This is a previously unexplored feature of Whittaker's beta diversity.

Because beta Whittaker's diversity and the fill of the matrix depend on the number of sites that are used to describe the distribution of species, they are sensitive to changes in the scale of analysis (Arita and Rodriguez 2002). However, the mathematical constraints that we have deduced are universal to any presence-absence matrix. Moreover, using the proportional values for species diversity, range size, and the fill of the matrix enables the comparison of range-diversity plots for biological systems across scales, taxonomic groups, and geographic locations.

One important conclusion derives from the proofs: parameters such as species diversity, Whittaker's beta diversity, the dispersion field volume, the average covariance among sites in terms of species composition, and nestedness are all interrelated and are subject to constraints inherently determined by their mathematical definitions. This probably means that patterns in biological diversity can be described using only a few interrelated parameters, making the proliferation of independent "indices" of diversity unnecessary. A possible unification of the varied concepts of biodiversity might be in sight.

Finally, although the approach presented here is mostly descriptive and static, it establishes the constraints under which ecological and evolutionary hypotheses on biodiversity patterns can be tested. For example, hypotheses based on nonrandom distribution of species can be tested

against a null model in which the distribution of each species is independent. Under these models, range-diversity plots by species would show points (species) arranged along a vertical line determined by the value of f^*. In contrast, assemblages determined by a positive association between species would produce plots with clouds of points to the right of that line, and sets of species whose distribution is determined by negative associations (produced, e.g., by competition) would generate plots with clouds of points to the left of the vertical f^* line.

In a similar fashion, hypotheses based on the diversity of sites can be tested with range-diversity plots. For example, if energy determines the species diversity of sites, then we would expect a high level of similitude between sites with similar levels of available energy. Under such a scenario, range-diversity plots by sites would generate clouds of points in the right side of the plot. However, if biogeographic history is incorporated, that simple pattern can be broken and the general similitude between sites might be lower. All those patterns can be contrasted with a null model of no association among sites using the range-diversity plots.

The theoretical exploration of ecological processes might benefit from a wider understanding of the mathematical constraints that limit the evolutionary paths of diversity and distribution. In applied studies, the identification of priority sites for conservation, an endeavor based mostly on the analysis of patterns, might benefit from better and more efficient strategies and algorithms that take into account the structural interrelations of the different indices of diversity and endemism that have been used to define conservation targets.

APPENDIX

Proof of the Boundaries in Figure 3

Let $D_i' = \sum_{i=1}^{S} \sum_{j=1}^{N} [1 - \delta(i, j)] \delta(l, j)$ be the addition of diversities in each of the sites in which species i does not

occur and let $d_i' = D_i'/(N - n_i)$ be the average diversity of the area not occupied by species i. It is clear that $D_i' + D_i = N\bar{s}$ and, remembering that $n_i^* = n_i/N$ ($0 < n_i^* \leq 1$),

$$\bar{s} = \frac{D_i'}{N} + \frac{D_i}{N} = \frac{N - n_i}{N} \frac{D_i'}{N - n_i} + \frac{n_i}{N} \frac{D_i}{n_i} \quad (A1)$$

or

$$\bar{s} = \left(1 - \frac{n_i}{N}\right) d_i' + \frac{n_i}{N} d_i = (1 - n_i^*) d_i' + n_i^* d_i. \quad (A2)$$

Defining $s_{\min}$ and $s_{\max}$ as the minimum and maximum values for species diversity in all sites, respectively, and considering that $\bar{s} \geq n_i^* d_i + (1 - n_i^*) s_{\min}$ and $\bar{s} \leq n_i^* d_i + (1 - n_i^*) s_{\max}$:

$$n_i^* \leq \frac{\bar{s} - s_{\min}}{d_i - s_{\min}},$$

$$n_i^* \leq \frac{s_{\max} - \bar{s}}{s_{\max} - d_i} \quad (A3)$$

In the first inequality above, given that $n_i^* \leq 1$, we require that the numerator be equal to or less than the denominator for the bound to be of any use, and so therefore $d_i \geq \bar{s}$. Equivalently, in the second inequality, $d_i \leq \bar{s}$. So, defining $s_{\min}^* = s_{\min}/S$, $s_{\max}^* = s_{\max}/S$, $\bar{s}^* = \bar{s}/S$, and $\bar{s}_i^* = d_i/S$, for any species in the set,

$$0 < n_i^* \leq \frac{\bar{s}^* - s_{\min}^*}{\bar{s}_i^* - s_{\min}^*}$$

if $\bar{s}^* \leq \bar{s}_i^* \leq s_{\max}^*$ and

$$0 < n_i^* \leq \frac{s_{\max}^* - \bar{s}^*}{s_{\max}^* - \bar{s}_i^*}$$

if $s_{\min}^* \leq \bar{s}_i^* \leq \bar{s}^*$.

With the same reasoning, it can be shown that the boundaries for a plot of s_j^* (the proportional species diversity in site j) and $\bar{n}_j^*$ (the average proportional range size of species occurring in site j) are

$$0 < s_j^* \leq \frac{\bar{n}^* - n_{\min}^*}{\bar{n}_j^* - n_{\min}^*}$$

if $\bar{n}^* \leq \bar{n}_j^* \leq n_{\max}^*$ and

$$0 < s_j^* \leq \frac{n_{\max}^* - \bar{n}^*}{n_{\max}^* - \bar{n}_j^*}$$

if $n^*_{\min} \le \bar{n}^*_j \le \bar{n}^*$, where $n^*_{\min}$ and $n^*_{\max}$ are the minimum and maximum range sizes, respectively, measured as proportions of the total number of sites in the region, and $\bar{n}^*$ is the average proportional range size among all species in the region.

Literature Cited

Anderson, S., and K. F. Koopman. 1981. Does interspecific competition limit the sizes of ranges of species? American Museum Novitates 2716:1–10.

Arita, H. T., and P. Rodriguez. 2002. Geographic range, turnover rate and the scaling of species diversity. Ecography 25:541–550.

Arita, H. T., F. Figueroa, A. Frisch, P. Rodriguez, and K. Santos del Prado. 1997. Geographical range size and the conservation of Mexican mammals. Conservation Biology 11:92–100.

Arita, H. T., P. Rodríguez, and E. Vázquez-Domínguez. 2005. Continental and regional ranges of North American mammals: Rapoport's rule in real and null worlds. Journal of Biogeography 32:961–971.

Atmar, W., and B. D. Patterson. 1993. The measure of order and disorder in the distribution of species in fragmented habitats. Oecologia (Berlin) 96:373–382.

Bell, G. 2003. The interpretation of biological surveys. Proceedings of the Royal Society B: Biological Sciences 270:2531–2542.

———. 2005. The co-distribution of species in relation to the neutral theory of community ecology. Ecology 86:1757–1770.

Brown, J. H., G. C. Stevens, and D. M. Kaufman. 1996. The geographic range: size, shape, boundaries, and internal structure. Annual Review of Ecology and Systematics 27:597–623.

Ceballos, G., P. R. Ehrlich, J. Soberón, I. Salazar, and J. P. Fay. 2005. Global mammal conservation: what must we manage? Science 309:603–607.

Connor, E. F., and S. Simberloff. 1979. The assembly of species communities: chance or competition? Ecology 60:1132–1140.

Diamond, J. M. 1975. Assembly of species communities. Pages 342–444 *in* M. L. Cody and J. M. Diamond, eds. Ecology and evolution of communities. Belknap, Cambridge, MA.

Gaston, K. J. 2003. Structure and dynamics of geographic ranges. Oxford University Press, Oxford.

Gotelli, N. J. 2000. Null model analysis of species co-occurrence patterns. Ecology 81:2606–2621.

Gotelli, N. J., and G. R. Graves. 1996. Null models in ecology. Smithsonian Institution, Washington, DC.

Gotelli, N. J., and D. J. McCabe. 2002. Species co-occurrence: a meta-analysis of J. M. Diamond's assembly rules model. Ecology 83:2091–2096.

Graves, G. R., and C. Rahbek. 2005. Source pool geometry and the assembly of continental avifaunas. Proceedings of the National Academy of Sciences of the USA 102:7871–7876.

Greve, M., and S. L. Chown. 2006. Endemicity biases nestedness metrics: a demonstration, explanation and solution. Ecography 29:347–356.

Hawkins, B. A., and J. A. F. Diniz-Filho. 2006. Beyond Rapoport's rule: evaluating range size patterns of New World birds in a two-dimensional framework. Global Ecology and Biogeography 15:461–469.

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

Hillebrand, H. 2004. On the generality of the latitudinal diversity gradient. American Naturalist 163:192–211.

Horner-Devine, M. C., J. M. Silver, M. A. Leibold, B. J. M. Bohannan, R. K. Colwell, J. A. Fuhrman, J. L. Green, et al. 2007. A comparison of taxon co-occurrence patterns for macro- and microorganisms. Ecology 88:1345–1353.

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

Legendre, L., and P. Legendre. 1983. Numerical ecology: developments in environmental modelling. Vol. 3. Elsevier Scientific, Amsterdam.

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

Lira-Noriega, A., J. Soberón, A. Navarro-Sigüenza, Y. Nakazawa, and A. T. Peterson. 2007. Scale dependency of diversity components estimated from primary biodiversity data and distribution maps. Diversity and Distributions 13:185–195.

Lutz, F. E. 1921. Geographic average, a suggested method for the study of distribution. American Museum Novitates 5:1–7.

McGeoch, M. A., and K. J. Gaston. 2002. Occupancy frequency distributions: patterns, artefacts, and mechanisms. Biological Reviews 77:311–331.

Mittelbach, G. G., D. W. Schemske, H. V. Cornell, A. P. Allen, J. M. Brown, M. B. Bush, S. P. Harrison, et al. 2007. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. Ecology Letters 10:315–331.

Mora, C., and D. R. Robertson. 2005. Factors shaping the range-size frequency distribution of the endemic fish fauna of the tropical Eastern Pacific. Journal of Biogeography 32:277–286.

Myers, N., R. A. Mittermeier, G. C. Mittermeier, G. A. B. da Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. Nature 403:853–858.

Orme, C. D. L., R. G. Davies, M. Burgess, F. Eigenbrod, N. Pickup, V. A. Olson, A. J. Webster, et al. 2005. Global hotspots of species richness are not congruent with endemism or threat. Nature 436:1016–1019.

Orme, C. D. L., R. G. Davies, V. A. Olson, G. H. Thomas, T.-S. Ding, P. C. Rasmussen, R. S. Ridgely, et al. 2006. Global patterns of geographic range size in birds. PLoS Biology 4:1276–1283.

Patterson, B. D. 1987. The principle of nested subsets and its implications for biological conservation. Conservation Biology 1:323–334.

Patterson, B. D., and W. Atmar. 1986. Nested subsets and the structure of insular mammalian faunas and archipelagos. Biological Journal of the Linnean Society 28:65–82.

Pielou, E. C. 1984. The interpretation of ecological data. Wiley-Interscience, New York.

Rapoport, E. H. 1975. Areografía: estrategias geográficas de las especies. Fondo de Cultura Económica, Mexico City.

Roberts, C. M., C. J. McClean, J. E. Veron, J. P. Hawkins, G. R. Allen, D. E. McAllister, C. G. Mittermeier, et al. 2002. Marine biodiversity hotspots and conservation priorities for tropical reefs. Science 295:1280–1284.

Rosenzweig, M. L. 1975. On continental steady states of species diversity. Pages 121–140 *in* M. L. Cody and J. M. Diamond, eds. Ecology and evolution of communities. Belknap, Cambridge, MA.

———. 1995. Species diversity in space and time. Cambridge University Press, Cambridge.

Routledge, R. D. 1977. On Whittaker's components of diversity. Ecology 58:1120–1127.

Ruggiero, A., and V. Werenkraut. 2007. One-dimensional analyses of Rapoport's rule reviewed through meta-analysis. Global Ecology and Biogeography 16:401–414.

Simberloff, S., and E. F. Connor. 1979. *Q*-mode and *R*-mode analysis of biogeographic distributions: null hypotheses based on random colonization. Pages 123–138 *in* G. P. Patil and M. L. Rosenzweig, eds. Contemporary quantitative ecology and related ecometrics. International Cooperative, Fairland, MD.

Simpson, G. G. 1964. Species density of North American recent mammals. Systematic Zoology 12:57–73.

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

Stone, L., and A. Roberts. 1990. The checkerboard score and species distributions. Oecologia (Berlin) 85:74–79.

Ulrich, W., and N. J. Gotelli. 2007*a*. Disentangling community patterns of nestedness and species co-occurrence. Oikos 116:2053–2061.

———. 2007*b*. Null model analysis of species nestedness patterns. Ecology 88:1824–1831.

Whittaker, R. H. 1972. Evolution and measurements of species diversity. Taxon 21:213–251.

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

Wright, D. H., and J. H. Reeves. 1992. On the meaning and measurement of nestedness of species assemblages. Oecologia (Berlin) 92:416–428.