Partitioning Species Diversity across Landscapes and Regions: A Hierarchical Analysis of α, β, and γ Diversity

Thomas O. Crist,[*] Joseph A. Veech,[†] Jon C. Gering,[‡] and Keith S. Summerville[§]

Department of Zoology, Miami University, Oxford, Ohio 45056

Submitted February 1, 2002; Accepted May 28, 2003;
Electronically published December 19, 2003

abstract: Species diversity may be additively partitioned within and among samples (α and β diversity) from hierarchically scaled studies to assess the proportion of the total diversity (γ) found in different habitats, landscapes, or regions. We developed a statistical approach for testing null hypotheses that observed partitions of species richness or diversity indices differed from those expected by chance, and we illustrate these tests using data from a hierarchical study of forest-canopy beetles. Two null hypotheses were implemented using individual- and sample-based randomization tests to generate null distributions for α and β components of diversity at multiple sampling scales. The two tests differed in their null distributions and power to detect statistically significant diversity components. Individual-based randomization was more powerful at all hierarchical levels and was sensitive to departures between observed and null partitions due to intraspecific aggregation of individuals. Sample-based randomization had less power but still may be useful for determining whether different habitats show a higher degree of differentiation in species diversity compared with random samples from the landscape. Null hypothesis tests provide a basis for inferences on partitions of species richness or diversity indices at multiple sampling levels, thereby increasing our understanding of how α and β diversity change across spatial scales.

Keywords: null hypothesis, randomization tests, spatial scale, species richness, statistical tests.

[*] Corresponding author; e-mail: cristto@muohio.edu.

[†] Present address: Department of Fishery and Wildlife Biology, Colorado State University, Fort Collins, Colorado 80523; e-mail: javeech@cnr.colostate.edu.

[‡] Present address: Division of Science, Truman State University, Kirksville, Missouri 63501; e-mail: jgering@truman.edu.

[§] Present address: Department of Environmental Science and Policy, Drake University, Des Moines, Iowa 50311; e-mail: ksummerville@drake.edu.

Am. Nat. 2003. Vol. 162, pp. 734–743.

Spatial patterns of species diversity change over multiple spatial scales. The pattern observed within a local community might be very different from those found over broader areas such as landscapes or regions. The analysis of scale-dependent patterns of species richness has primarily focused on changes in the species-area relationship over a continuous range of spatial scales (Palmer and White 1994; Crawley and Harral 2001; Willis and Whittaker 2002). The additive partitioning of species diversity is another promising approach for analyzing patterns of diversity sampled from hierarchically scaled studies (Lande 1996; Loreau 2000; Godfray and Lawton 2001). For over 30 yr, however, diversity partitioning has been more conceptual than operational, dating back to Whittaker's (1960) concepts of α, β, and γ diversity.

Historically, additive partitioning of diversity was not linked to Whittaker's concepts (Veech et al. 2002) but instead developed from niche theory (MacArthur et al. 1966; Levins 1968). Allan (1975*a*, 1975*b*) later used additive partitions of the Shannon index to estimate diversity found within and among habitats, and Rao (1982) derived additive partitions of species diversity from statistical theory. Recent interest in diversity partitioning largely stems from Lande's (1996) application of Whittaker's terms to additive partitions of total species richness or diversity (γ) into components found within α samples and among β samples. Additive partitions have now been used to analyze hierarchical patterns of species diversity in agricultural landscapes (Wagner et al. 2000; Fournier and Loreau 2001), tropical forests (DeVries et al. 1997, 1999; DeVries and Walla 2001), and temperate forests (Gering et al. 2003; Summerville et al. 2003). Diversity partitioning has several advantages over other quantitative techniques, yet there is still a need to make diversity partitioning more operational and statistically rigorous.

Diversity partitioning differs from other analyses of species diversity in several ways. Analyses of α diversity often involve comparisons of mean species richness found in two or more sets of samples using analysis of variance. Alternatively, the species richness of samples can be compared using species accumulation curves and rarefaction

(Colwell and Coddington 1994; Gotelli and Colwell 2001). In contrast, analyses of β diversity usually involve dimensionless metrics based on species dissimilarity or turnover between pairs of samples, which lead to different definitions of β diversity (Magurran 1988; Vellend 2001). Additive partitions of diversity, however, decompose γ diversity into α and β components that are expressed in the same units. The total species richness (γ) found in a collection of samples at any spatial scale can be partitioned into the average number of species that occur within a sample (α) and the average number of species absent from a sample (β; Veech et al. 2002). Because β is expressed in units of species richness, the contributions of α and β to total species richness can be compared across spatial or temporal sampling scales (Wagner et al. 2000; DeVries and Walla 2001; Fournier and Loreau 2001; Gering et al. 2003; Summerville et al. 2003).

Despite a growing empirical interest in diversity partitioning, however, its use is still descriptive with little theoretical basis for interpreting the observed patterns of α and β diversity (Rosenzweig 1995; Loreau 2000) or statistical methods for testing null hypotheses on observed diversity partitions. Null models and associated hypothesis tests have likewise played a central role in the theoretical development of several other concepts in community ecology, including patterns of species cooccurrence (Gotelli and Graves 1996; Gotelli 2000), nested species subsets (Brualdi and Sanderson 1999; Summerville et al. 2002), and the relationships between spatial distribution and abundance (Blackburn and Gaston 2000; Bell 2001).

Here we develop a statistical approach for testing two different null hypotheses that observed partitions of species diversity do not differ from those expected by chance. We illustrate this technique with a hierarchical study of forest-canopy beetles to test whether observed partitions are significantly different from the null hypothesis across several sampling scales. The statistical power of each null hypothesis is also assessed by the differences in the means and overlap in the observed and null distributions of diversity components. Finally, we examine whether the observed diversity partitions deviate from the null hypothesis because of nonrandom patterns of intraspecific aggregation or the spatial differentiation of diversity among samples at different scales.

Methods

Diversity Partitioning

Let γ be the total species diversity (as measured by species richness or a diversity index) found in a collection of samples. The total diversity is partitioned into the average diversity within samples (α) and among samples (β) so that $\gamma = \alpha + \beta$, and β diversity can be estimated by $\beta = \gamma - \alpha$ (Wagner et al. 2000). To extend across multiple scales, suppose we have a hierarchical design with $i = 1, 2, 3 \ldots m$ levels of sampling. Samples in the lowest hierarchical level $i = 1$ represent the smallest sampling unit nested within samples at $i = 2$, and samples at $i = 2 \ldots m$ are formed by pooling together the appropriate groups of nested samples from each level $i - 1$. Let α_i represent the average diversity found within samples at each level i. The diversity components are calculated as $\beta_m = \gamma - \alpha_m$ at the highest sampling level and $\beta_i = \alpha_{i+1} - \alpha_i$ for each lower sampling level. Then, the additive partition of diversity is

$$\gamma = \alpha_1 + \sum_{i=1}^{m} \beta_i. \tag{1}$$

The total diversity can therefore be expressed as the proportional contributions of diversity due to each level in the hierarchical sampling design.

Any measure of species diversity can be additively partitioned if $\gamma \geq \alpha$, a condition that ensures that β is nonnegative (Lande 1996). Additive partitioning can be conducted on the most widely used diversity metrics: species richness, the Shannon index, and the Simpson index (Lande 1996). The partitions among hierarchical levels are expected to differ for the three diversity metrics since they quantify different aspects of community structure (Magurran 1988). All three, however, can be partitioned in the same way. Let D_{ij} be the diversity metric recorded in each sample $j = 1, 2, 3 \ldots n_i$, where n_i is the number of samples taken at level i of the hierarchical sampling design. The average diversity (α_i) for samples at each hierarchical level i is

$$\alpha_i = \sum_{j=1}^{n_i} D_{ij} q_{ij}, \tag{2}$$

where q_{ij} are the sample weights determined by the proportion of the total number of individuals found in each sample j. The β diversity at each level i is then obtained by the above formulas. To partition species richness, D_{ij} is the number of species in sample j. The Shannon index of diversity as a measure of D_{ij} is obtained by $D_{ij} = -\sum_k p_{ijk} \ln p_{ijk}$, where p_{ijk} is the proportional abundance of species k in each sample j. Finally, the complement of the Simpson index of diversity (the Gini coefficient) is used as a measure of D_{ij} and is obtained by $D_{ij} = 1 - \sum_k p_{ijk}^2$. The Gini coefficient (hereafter "Simpson index") expresses the probability that two individuals drawn at random from sample j belong to different species (Magurran 1988; Lande 1996).

Null Hypotheses and Randomization Tests

We developed two null hypotheses that provide quite different tests on the observed partitions of diversity. The first (H_1) tests whether the observed partition of diversity could have been explained by a random distribution of individuals. It may useful for testing alternative explanations of species diversity that are based on the nonrandom distribution of individuals such as intraspecific aggregation, resource partitioning, and community saturation (Shorrocks and Sevenster 1995). In contrast, the second (H_2) is appropriate for testing the significance of an observed diversity component as it relates to the observed distribution of diversity among samples at lower levels in a nested design. The primary difference between the two hypotheses is whether individuals (H_1) or samples (H_2) are randomized (table 1).

Our first hypothesis assumes that the observed numbers of individuals of each species are randomly placed among samples at the lowest level ($i = 1$), and samples containing randomly placed individuals are then grouped together into progressively larger samples at each higher level $i = 2 \ldots m$. Thus our first null hypothesis (H_1) is the observed components of diversity (α_1 and β_i) could have been obtained by the random distribution of individuals among samples at all hierarchical levels.

To test this hypothesis, we conduct a simulation procedure that we refer to as individual-based randomization because the numbers of individuals and species in samples are determined by the random placement of individuals into samples at the lowest level. The randomization procedure uses a reshuffling algorithm to place individuals in samples randomly while preserving the original species-abundance and sample-size distributions. Samples at higher levels are obtained by pooling the appropriate samples at lower levels, as when the actual data are partitioned. This type of randomization has been used in other theoretical explorations of species diversity and composition and is used when the random placement of individuals is the null hypothesis of interest (Coleman et al. 1982; Solow 1993; Dixon 1994; Phillipi et al. 1998).

Each randomized data set is then partitioned into α and β components at each hierarchical level for the three diversity measures as previously described. The randomization process is repeated 1,000–10,000 times to obtain null distributions of the α and β diversity estimates at each hierarchical level. The probability (P value) that a diversity component greater than the observed could be obtained by chance is obtained from the proportion of null values that are greater than the observed value; alternatively, the proportion of null values less than the observed is the probability of obtaining a diversity component less than the observed value (Manly 1991; Gotelli and Graves 1996).

An important property of individual-based randomization is that it allows one to compare the observed diversity components at each hierarchical level with the mean values of the randomizations because γ diversity is the same for the observed and randomized data. Another important feature is that each randomization produces a diversity partition that is additive as in the observed partition. Because of random placement and additivity, this null hypothesis is useful in identifying how intra- and interspecific patterns of spatial aggregation may lead to different kinds of diversity partitions. Intra- and interspecific aggregation have a key role in explanations of species coexistence and diversity (Ives 1991; Shorrocks and Sevenster 1995; He and Legendre 2002).

The second null hypothesis (H_2) tests whether the observed partition of diversity could have been obtained by the random allocation of lower-level samples among higher-level samples. In other words, could the observed partition merely be a consequence of the sampling design? Because H_2 preserves the patterns of intraspecific aggregation in the observed data, it is most useful in testing explanations of species diversity that are based on nonrandom species assemblages. The second null hypothesis is motivated by hypothesis testing in nested experimental designs where the effect of each sampling level is tested using the next lowest level as an error term. We emphasize, however, that hypothesis tests and variance components in nested ANOVA only consider α diversity and therefore differ analytically from diversity partitioning (cf. Fournier and Loreau 2001). The second hypothesis (H_2) is that the

Table 1: Summary of the randomization procedures used to test two different null hypotheses on the observed additive partitions of species diversity

Description	Hypothesis 1	Hypothesis 2
Units of randomization	Individuals	Samples
Randomization procedure	Units assigned to any sample at level $i = 1$	Units at $i - 1$ assigned to any sample at level i nested within the same sample at $i + 1$
Separate randomization required for each level i?	No	Yes
Additive?	Yes	No

observed component of diversity at each hierarchical level i (α_1 and β_i) could have been obtained by a random distribution of samples at level $i - 1$ among the level i samples that form an $i + 1$ sample.

To test the significance of α_1 and β_i, samples at level $i - 1$ are randomly allocated only to those samples at level i that belong to the same sample unit at $i + 1$. We therefore refer to this procedure as sample-based randomization because samples are randomized rather than individuals. Because samples differ at each hierarchical level, a separate randomization test is required to test for each level i. For example, to test the significance of β_2, samples at level $i = 1$ are randomly assigned to those level $i = 2$ samples that belong to the same sample unit at level $i = 3$, and so on for each β_i. Null values of β_i obtained from 1,000–10,000 randomizations are used to obtain a P value for the observed β_i at each hierarchical level. Because the statistical significance of each diversity component is tested using a separate set of randomizations, the expected values of α_1 and β_i are not additive to the total diversity (table 1).

Our randomization procedures can be applied when the sample size varies (i.e., the number of individuals recorded in each sample) because the samples are weighted by abundance (q_{jk}) when α diversity is calculated, as in partitioning of the observed data. Larger samples contribute more to the diversity estimate than do small samples, and these differences in sample size are retained in the randomized data. Therefore, randomization procedures are robust to sample-size variation, and there is no need for rarefaction (Solow 1993).

We developed a computer program, PARTITION, to partition species richness, Shannon diversity, and Simpson diversity on balanced or unbalanced designs with up to six hierarchical levels. The program also does the two types of randomization and significance testing as described above.

To illustrate our approach, we used data from a study of forest-canopy beetles involving four sampling scales from individual trees to regions (Gering et al. 2003). The broadest sampling scale included two regions of the eastern deciduous forest in Ohio and Indiana, which differ in their biogeography history. In each region, we selected three study sites that were parks and preserves with large areas (>200 ha) of mature forest. Within each site, four forest stands were sampled, with two stands located in xeric uplands and two in mesic lowlands, for a total of 24 stands. Beetles were sampled by fogging tree crowns in each stand with pyrethrin insecticide. Individual trees therefore comprised the lowest level of sampling. Within each stand, four trees within a 1-ha area were fogged based on the dominant trees within each stand. Thus the example data set is based on samples from a total of 96 trees.

Statistical Power

We estimated the power of the individual- and sample-based randomization methods by comparing the sampling distribution of each beetle diversity component with its corresponding null distribution. The α and β components of diversity derived from diversity partitioning are means (Veech et al. 2002), and therefore a standard error can be calculated for each diversity component based on the sample distributions. For each diversity component, the standard error of the samples was then used to obtain the normal probability distribution, which we assumed to be representative of the true distribution of each diversity component. From the null distributions, we obtained critical values that were either the ninety-fifth (upper critical value) or fifth percentile (lower critical value) depending on whether the observed component of diversity was significantly larger or smaller than the mean of null distribution. We then used the critical values and the sampling distributions to determine the probability of a Type II error failing to reject a false null hypothesis. This probability was either the proportion of the sampling distribution to the left of the upper critical value (for a significantly large observed component) or the proportion to the right of the lower critical value (for a significantly small component). Power was then estimated as 1 minus the probability of a Type II error.

Intraspecific Aggregation

Since H_1 is based on the random placement of individuals, we tested whether differences between the observed and expected components of diversity could be explained by the nonrandom spatial aggregation of conspecifics. As a measure of aggregation, we used the Morisita index because it is relatively insensitive to differences in sample size (Krebs 2000). Values of Morisita's index can be interpreted as the probability that two individuals randomly drawn from the observed population belong to the same sample compared with the expected probability from a population with a random dispersion. Thus it is expressed as a likelihood ratio where values >1 indicate spatial aggregation of individuals and values <1 indicate a regular dispersion pattern (Hurlbert 1990). We first calculated the mean of the values of the index for the nonsingleton species (i.e., species represented by >1 individual) of beetles observed at each sampling level. Next, we calculated the mean of the values at each sampling level obtained from each of the 10,000 randomizations of the data. The observed mean values at each hierarchical level were then compared with the range of means obtained by randomization.

Results

Diversity Partitioning

The example data set had 8,662 beetles of 467 species from canopy-fogging samples of 96 trees (Gering et al. 2003). Less than 20% of the total species richness was due to within- and among-tree components (α_1 and β_1; fig. 1*A*). The components due to stand, sites, and regions were roughly equal and comprised most of the total species richness (fig. 1*A*). The expected components of species richness from null hypothesis H_1 (i.e., the means of the null distributions) are also additive to the total richness, and their proportional contributions can be compared directly with the observed values (fig. 1*A*). The species richness at the lower sampling levels comprised a greater proportion of the total than in the observed partition.

Additive partitions of Shannon index showed that α diversity comprised 62% of the total beetle diversity (fig. 1*B*). In contrast to species richness, only 13% was found among trees (β_1), 11% among stands (β_2), and <10% each at the site (β_3) and regional levels (β_4; fig. 1*B*). The differences in the two types of partitions can be explained primarily by the influence of widespread dominant species (Gering et al. 2003). Even though a greater portion of the Shannon diversity is found locally within and among trees, the observed levels of β diversity were greater than the values expected from individual-based randomization (fig. 1*B*).

The two null hypothesis tests on the observed partitions of species diversity resulted in very different patterns of statistical significance. In tests of H_1, species richness was significantly lower than expected for sampling levels up to and including β_2 for stands (fig. 2*A*) and was greater than expected for sites and regions. In contrast, tests of H_2 showed that only the β_2 component at the stand level was significantly greater than expected (fig. 2*B*). The α_1 for the Shannon index was large in comparison to species richness but still significantly less than that expected from H_1 (fig. 2*C*). The remaining β components of diversity all were significantly higher than expected even though they decreased across hierarchical levels (fig. 2*D*). Tests with H_2 generally showed very similar observed and expected components for the Shannon index; only the observed diversity for sites (β_3) was significantly greater than expected (fig. 2*D*).

The observed partition of the Simpson index (not shown) was qualitatively similar to that from the Shannon index but showed an even greater importance of α diversity with 89% of the total beetle diversity within trees. Only 1%–4% of the total was found within each of the components at higher sampling levels. Randomization tests with H_1 produced the same patterns of statistical significance as those found with Shannon diversity,

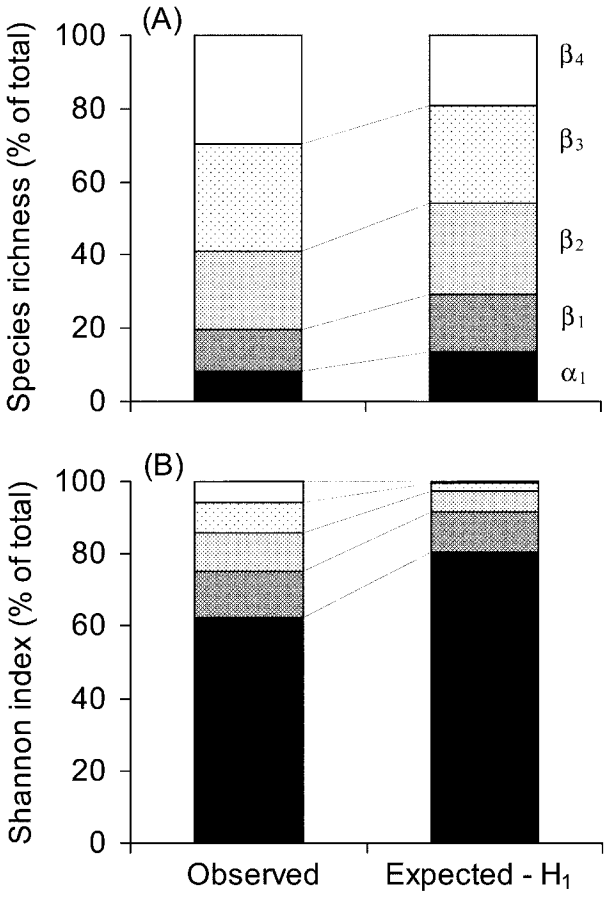

Figure 1: Example of the additive partition of species richness and the Shannon index of diversity across four sampling scales. Values are expressed as the percent of the total diversity of forest-canopy beetles explained by each hierarchical level. The observed partitions are compared to expected values from individual-based randomization (H_1). α_1, β_1 = trees, β_2 = stands, β_3 = sites, β_4 = regions.

whereas tests with H_2 showed no significant differences between observed and expected diversity components.

Power estimates for hypothesis tests were generally greater for individual-based (H_1) than for sample-based (H_2) randomization tests (table 2) because the latter produced wider null distributions and expected values that were closer to the observed components (fig. 2). Statistical power in tests of H_1 on species richness declined slightly with decreased sample size at higher sampling levels but was generally high for both richness and the Shannon index. Power in tests of H_2 on species richness was highest at the stand level (β_2), which is where a statistically significant departure was detected between observed and null values. The Shannon index had slightly higher power than species richness in tests of H_2 (table 2). Power estimates from the Simpson index (not shown) were very similar to

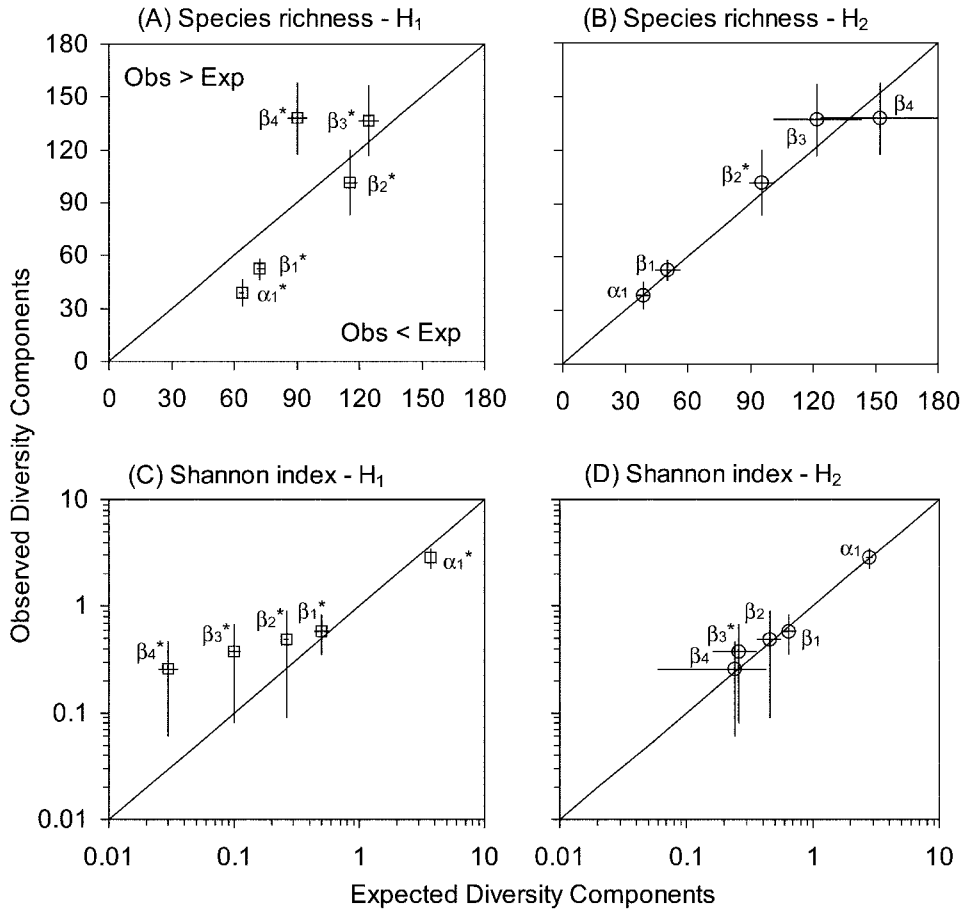

Figure 2: Comparisons of observed and expected diversity from two different null hypotheses and randomization tests (H_1 and H_2) on species richness (*A*, *B*) and the Shannon index of diversity (*C*, *D*). Observed components of diversity are plotted against means of the null distributions (expected values). Statistically significant ($P < .05$) departures from equal observed and expected components (45° line) are indicated by an asterisk. Vertical error bars are 2 SE of the mean of the observed sample distribution, and horizontal error bars are the critical upper and lower values that encompass 95% of the null distribution obtained from 10,000 randomizations.

those obtained from the Shannon index in both types of hypothesis tests.

Intraspecific Aggregation

The observed means of the Morisita indices showed that beetle species were aggregated across all spatial scales but especially at the tree and stand levels of sampling (fig. 3). Thus a high degree of intraspecific aggregation on individual trees or within forest stands corresponded to significantly low levels of α and β diversity for trees and stands in tests of species-richness components using H_1, which is based on the same null hypothesis of a random distribution of individuals among samples as in the Morisita index.

Discussion

The additive partitioning of diversity unifies the disparate approaches to studying species diversity and composition. Traditionally, sample (α) and differentiation (β) diversity have been analyzed using separate hypothesis tests and analyses. We have shown how different null hypotheses and randomization tests can be used to analyze the relative contributions of sample and differentation diversity across multiple sampling scales. The α and β components of diversity at each sampling scale are means: α is the average diversity found within a sample, and β is the average diversity that is absent from a randomly chosen sample (Veech et al. 2002). These means also have sample variances based on species that are present and absent from each sample at a given hierarchical level. Here we used

Table 2: Estimated power to detect significant differences between the observed and expected components of diversity[a] shown in figure 2

Diversity component	Sample size	Species richness H_1	Species richness H_2	Shannon index H_1	Shannon index H_2
Sites (β_3)	6	.77	.26	.90	.52
Stands (β_2)	24	.89	.50	.84	.39
Among trees (β_1)	96	1.00	.09	.76	.48
Within trees (α_1)	96	1.00	.26	.99	.43

[a] Power for β_4 was not estimated because there were only two samples at the regional level. Power is estimated from the fraction of the observed sampling distribution that is above the upper critical value of the null distribution if the observed diversity component is greater than the expected or the fraction of the observed distribution that is below the lower critical value of the null distribution if the observed component is less than the expected.

this sampling distribution to estimate the statistical power of the null hypothesis tests in a hierarchical sampling design, but the contribution of individual samples to the overall α and β diversity components has other applications as well. For example, Gering and Crist (2002) showed how the richness components of individual samples can be used in a regression analysis to determine how the relative importance of α and β components change in local-regional relationships of species richness, which provided a quantitative assessment of this concept (Loreau 2000; Godfray and Lawton 2001).

Statistics and Hypothesis Tests on Diversity Partitions

Until now, a major impediment to interpreting diversity partitions has been the lack of a statistical framework for comparing observed partitions to expected values or testing them against a null hypothesis. Here we provide direct statistical tests of the diversity partitions using all of the species abundance data present in the original samples. Although the use of randomization to test for differences in species diversity and composition is not new (Solow 1993; Dixon 1994; Phillipi et al. 1998), the application to diversity partitioning provides a new basis for inferences and interpretation of observed diversity partitions against null hypotheses.

The statistical power to detect differences between observed and null partitions of diversity depends on the hypothesis test and the observed variability in α and β diversity among samples. For a given diversity component, the power of the individual- and sample-based randomization tests is determined from the width and location of the sampling distribution relative to the null distribution. If the entire sampling distribution lies to the right of the upper critical value or to the left of the lower critical value, then the power of the test equals 1.0 (e.g., α_1, β_1, β_2 for H_1). If the standard error of β_1 is increased from 2.9 up to 52.4 (equal to the mean), the sampling distribution becomes wider and power gradually decreases, but it still remains high (power = 0.64, with SE = 52.4). A decrease in power also occurs by lessening the effect size (difference between the observed mean and the upper or lower critical value of the null distribution) with the standard error a constant proportion of the mean.

Power estimates for individual-based randomization were always greater than those for sample-based randomization because the latter produced wider null distributions with means that were also closer to those of the sampling distributions (table 2; fig. 2). For any given diversity component, the power of the randomization test will depend on those factors that determine the distance between the sampling and null distributions by affecting either the widths of the distributions or their locations. These factors include the number of samples at each hierarchical level and the size of each sample unit as it affects the numbers

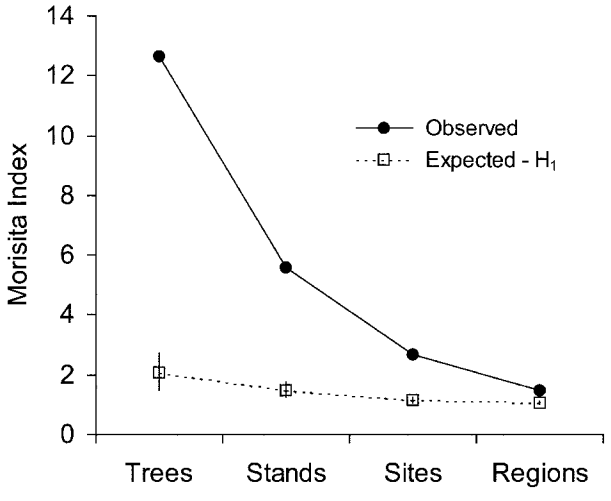

Figure 3: Intraspecific aggregation of beetle species across sampling levels. The mean values of the Morisita index of aggregation for the observed data (*solid lines*) are compared with the expected values under H_1 (*dashed lines*) derived from the mean of 10,000 randomizations. Both observed and expected values use only species that are represented by two or more individuals. Error bars represent 95% of the upper and lower values from randomizations.

of species and individuals. Although a detailed analysis of each of these factors is beyond the scope of this article, we have partitioned more than 75 data sets and found that individual-based randomization is usually sufficiently powerful for most data sets containing at least a few hundred individuals. The number of species and samples is not as crucial to the power and performance of the randomization tests. Indeed, we have found significantly high β-diversity in data sets consisting of <300 individuals representing 30 or fewer species distributed among only three samples. If, however, H_2 is the hypothesis of interest, then statistical power is more of an issue because the expected values may be more similar to the observed and the widths of the null distributions from sample-based randomization are substantially larger. For example, the low power of H_2 for α_1 and β_1 can be attributed to the small difference between observed and expected components, whereas limited power for β_2 and β_3 was due to the large variability in the null distributions (fig. 2).

Which Null Hypothesis?

The differences in statistical significance and power raise the issue of which hypothesis test is most appropriate, a problem that is not unique to our null hypotheses tests on diversity partitioning. Indeed, there is a long and controversial history over the choice of null models and randomization tests in community ecology (Gotelli and Graves 1996). This is especially true for tests involving species co-occurrence patterns, where the merits of several randomization algorithms are still debated (Gotelli and Entsminger 2001; Manly and Sanderson 2002).

Several different null hypotheses can also be specified for diversity partitions. We have considered other alternative hypotheses and randomization procedures, but we focused on these two because they have clear precedents in ecological theory or statistical analysis. The choice of the null hypothesis (H_1 or H_2) to test against observed diversity partitions will depend on the objectives of the investigator. Individual-based randomization (H_1) will be useful in determining how observed diversity partitions may depart from expected values as a result of intraspecific aggregation (e.g., fig. 3). Intraspecific aggregation has a key role in explanations of species coexistence and diversity (Ives 1991; Shorrocks and Sevenster 1995; He and Legendre 2002). For example, if individual species are aggregated in different resource patches or habitats, then within-habitat diversity would be lower and among-habitat diversity would be greater than expected by chance (Shorrocks and Sevenster 1995). Consistent with this prediction were the high levels of intraspecific aggregation at fine sampling scales (fig. 3); the significantly low levels of species richness at α_1, β_1, and β_2; and the high levels of richness at β_3 and β_4 (fig. 2).

Sample-based randomization (H_2) maintains the integrity of samples at each hierarchical level. Thus it is most useful when one is primarily interested in testing whether nonrandom groups of species explain the observed patterns of diversity. For example, this hypothesis tests whether samples from different habitats (e.g., forest stands) show a higher (or lower) degree of differentiation in species richness compared with random samples from the landscape. Similarly, one could ask whether sites differ to a greater (or lesser) extent in their species evenness or dominance (Shannon or Simpson index) compared with random samples from the region. In our example, tests of H_2 generally showed these components were more similar to expected values compared to H_1 (fig. 2) because species aggregation was retained in randomization of sample units. This explains why the β_2 component of species richness was significantly less than expected from H_1 and significantly greater than expected from H_2. When the randomization test retains species aggregation within stands, as in H_2, the beetle species richness was more highly differentiated among forest stands within a site than would be expected by chance because of differences in beetle species composition and relative abundance among mesic and xeric stands. At broader scales, both hypothesis tests showed greater levels of Shannon diversity among sites (β_3) than expected by random placement of sites within regions because sites show spatial differentiation of diversity within regions.

Our comparisons of these two null hypotheses are not meant to be prescriptive. Rather, they illustrate how alternative null hypotheses can provide different insights into factors that influence the spatial structure of species diversity (cf. Gotelli 2000). Moreover, they provide a starting point for future hypothesis testing on diversity partitions. A natural extension of H_2, for example, would be to use sample-based randomization or bootstrapping to determine whether there are statistically significant differences in the observed levels of β diversity obtained from separate partitions of two or more habitats, landscapes, or regions (Harrison 1999; Fournier and Loreau 2001; Summerville and Crist 2001; Summerville et al. 2003).

Comparison of Species Richness and the Shannon Index

The contrasting partitions of species richness and the Shannon index of diversity can be explained by patterns of species dominance or rarity. In our example, about half of the total species richness of beetles was due to the β diversity of sites and region. In contrast, most of the total Shannon index was explained by α diversity within trees or stands. The most abundant species were also widespread

so that the same common species comprise most of the α diversity of trees or stands (Gering et al. 2003). Conversely, most rare species were found on a single tree or within a stand, resulting in a small contribution to α diversity for species dominance in stands. We therefore expect that additive partitions of species richness will have a high β diversity and partitions of Shannon or Simpson diversity will have a high α diversity due to the generality of the distribution-abundance relationship (Brown 1995; Blackburn and Gaston 2000). Wagner et al. (2000) also found that diversity partitions of plant species richness and dominance showed striking differences in the relative importance of α and β diversity. Over half of the total plant species richness was due to β diversity found among agricultural fields, whereas most of the Simpson diversity was explained by α diversity within habitats (Wagner et al. 2000). The use of null hypothesis tests on partitions of richness or diversity should facilitate comparisons of diversity partitions because effect sizes from null distributions can be standardized across studies and taxa (Gotelli and McCabe 2002).

Conclusions

The additive partitioning of species diversity provides an operational method for analyzing species diversity across multiple spatial scales, but its use has been limited because it lacked an analytical and theoretical framework for testing hypotheses on patterns of species diversity. Here we provided two different hypothesis tests on diversity partitions from hierarchical studies, estimates of statistical power for the two hypothesis tests at different sampling scales, and explanations for departures from null hypotheses based on patterns of species aggregation or the spatial structure of diversity across scales. A logical extension would be to develop hypothesis tests to compare the β diversities found in different landscapes and regions or to analyze diversity partitions across continuous spatial and temporal scales.

Acknowledgments

Discussions with R. Lande sharpened our thinking on hypothesis testing of diversity partitions. S. Harrison, H. Wagner, and several anonymous reviewers provided several insightful comments and criticisms on earlier versions of the manuscript. This research was funded by the Nature Conservancy Ecosystem Research Program and the National Geographic Society.

Literature Cited

Allan, J. D. 1975*a*. Components of diversity. Oecologia (Berlin) 18:359–367.

———. 1975*b*. The distributional ecology and diversity of benthic insects in Cement Creek, Colorado. Ecology 56:1040–1053.

Bell, G. 2001. Neutral macroecology. Science 293:2413–2418.

Blackburn, T. M., and K. J. Gaston. 2000. Pattern and process in macroecology. Blackwell Science, Oxford.

Brown, J. H. 1995. Macroecology. University of Chicago Press, Chicago.

Brualdi, R. A., and J. G. Sanderson. 1999. Nested species subsets, gaps, and discrepancy. Oecologia (Berlin) 119: 256–264.

Coleman, B. D., M. A. Mares, M. R. Willig, and Y.-H. Hsieh. 1982. Randomness, area, and species richness. Ecology 63:1121–1133.

Colwell, R. K., and J. A. Coddington. 1994. Estimating the extent of terrestrial biodiversity through extrapolation. Philosophical Transactions of the Royal Society of London B, Biological Sciences 345:101–118.

Crawley, M. J., and J. E. Harral. 2001. Scale dependence in plant biodiversity. Science 291:864–868.

DeVries, P. J., and T. R. Walla. 2001. Species diversity and community structure in Neotropical fruit-feeding butterflies. Biological Journal of the Linnean Society 74:1–15.

DeVries, P. J., D. Murray, and R. Lande. 1997. Species diversity in vertical, horizontal, and temporal dimensions of a fruit-feeding butterfly community in an Ecuadorian forest. Biological Journal of the Linnean Society 62:343–364.

DeVries, P. J., T. R. Walla, and H. F. Greeney. 1999. Species diversity in spatial and temporal dimensions of fruit-feeding butterflies from two Ecuadorian forests. Biological Journal of the Linnean Society 68:333–353.

Dixon, P. M. 1994. Resampling techniques: size hierarchies and similarity indices. Pages 290–318 *in* S. Scheiner and J. Gurevitch, eds. Design and analysis of ecological experiments. Chapman & Hall, London.

Fournier, E., and M. Loreau. 2001. Respective roles of hedges and forest patch remnants in the maintenance of ground-beetle (Coleoptera: Carabidae) diversity in an agricultural landscape. Landscape Ecology 16:17–32.

Gering, J. C., and T. O. Crist. 2002. The alpha-beta-regional relationship: providing new insights into local-regional patterns of species richness and scale dependence of diversity components. Ecology Letters 5:433–444.

Gering, J. C., T. O. Crist, and J. A. Veech. 2003. Additive partitioning of species diversity across multiple spatial scales: implications for regional conservation of biodiversity. Conservation Biology 17:488–499.

Godfray, H. C. J., and J. H. Lawton. 2001. Scale and species numbers. Trends in Ecology & Evolution 16:400–404.

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

Gotelli, N. J., and R. K. Colwell. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecology Letters 4: 379–391.

Gotelli, N. J., and G. L. Entsminger. 2001. Swap and fill algorithms in null model analysis: rethinking the knight's tour. Oecologia (Berlin) 129:281–291.

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

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.

Harrison, S. 1999. Local and regional diversity in a patchy landscape: native, alien, and endemic herbs on serpentine. Ecology 80:70–80.

He, F. L., and P. Legendre. 2002. Species diversity patterns derived from species-area models. Ecology 83:1185–1198.

Hurlbert, S. H. 1990. Spatial distribution of the montane unicorn. Oikos 58:257–271.

Ives, A. R. 1991. Aggregation and coexistence in a carrion fly community. Ecological Monographs 61:75–94.

Krebs, C. J. 2000. Ecological methodology. 2d ed. Benjamin Cummins, Menlo Park, Calif.

Lande, R. 1996. Statistics and partitioning of species diversity, and similarity among multiple communities. Oikos 76:5–13.

Levins, R. 1968. Evolution in changing environments: some theoretical explorations. Princeton University Press, Princeton, N.J.

Loreau, M. 2000. Are communities saturated? on the relationship between α, β and γ diversity. Ecology Letters 3:73–76.

MacArthur, R. H., H. Recher, and M. Cody. 1966. On the relation between habitat selection and species diversity. American Naturalist 100:319–332.

Magurran, A. E. 1988. Ecological diversity and its measurement. Princeton University Press, Princeton, N.J.

Manly, B. F. J. 1991. Randomization and Monte Carlo methods in biology. Chapman & Hall, London.

Manly, B. F. J., and J. G. Sanderson. 2002. A note on null models: justifying the methodology. Ecology 83:580–582.

Palmer, M. W., and P. S. White. 1994. Scale dependence and the species-area relationship. American Naturalist 144:717–740.

Phillipi, T. E., P. M. Dixon, and B. E. Taylor. 1998. Detecting trends in species composition. Ecological Applications 8:300–308.

Rao, C. R. 1982. Diversity: its measurement, decomposition, apportionment and analysis. Sankhya A 44:1–22.

Rosenzweig, M. L. 1995. Species diversity in space and time. Cambridge University Press, Cambridge.

Shorrocks, B., and J. G. Sevenster. 1995. Explaining local species diversity. Proceedings of the Royal Society of London B, Biological Sciences 260:305–309.

Solow, A. R. 1993. A simple test for change in community structure. Journal of Animal Ecology 62:191–193.

Summerville, K. S., and T. O. Crist. 2001. Diversity of Lepidoptera in Ohio forests at local and regional scales: how heterogeneous is the fauna? Annals of the Entomological Society of America 94:583–591.

Summerville, K. S., J. A. Veech, and T. O. Crist. 2002. Does variation in patch use among butterfly species contribute to nestedness at fine spatial scales? Oikos 97:195–204.

Summerville, K. S., M. Boulware, J. A. Veech, and T. O. Crist. 2003. Spatial variation in species diversity and composition of forest Lepidoptera: patterns and implications for conservation. Conservation Biology 17: 1045–1057.

Veech, J. A., K. S. Summerville, T. O. Crist, and J. C. Gering. 2002. The additive partitioning of diversity: recent revival of an old idea. Oikos 99:3–9.

Vellend, M. 2001. Do commonly used indices of β-diversity measure species turnover? Journal of Vegetation Science 12:545–552.

Wagner, H. H., O. Wildi, and K. C. Ewald. 2000. Additive partitioning of plant species diversity in an agricultural mosaic landscape. Landscape Ecology 15:219–227.

Whittaker, R. H. 1960. Vegetation of the Siskiyou Mountains, Oregon and California. Ecological Monographs 30:279–338.

Willis, K. J., and R. J. Whittaker. 2002. Species diversity: scale matters. Science 295:1245–1248.

Associate Editor: William J. Boecklen