Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness

Nicholas J. Gotelli[1]
and Robert K. Colwell[2]

[1]*Department of Biology,
University of Vermont,
Burlington,
Vermont 05405, U.S.A.
E-mail: ngotelli@zoo.uvm.edu*
[2]*Department of Ecology
and Evolutionary Biology,
U-43, University of Connecticut,
Storrs,
Connecticut 06269, U.S.A.*

Abstract

Species richness is a fundamental measurement of community and regional diversity, and it underlies many ecological models and conservation strategies. In spite of its importance, ecologists have not always appreciated the effects of abundance and sampling effort on richness measures and comparisons. We survey a series of common pitfalls in quantifying and comparing taxon richness. These pitfalls can be largely avoided by using accumulation and rarefaction curves, which may be based on either individuals or samples. These taxon sampling curves contain the basic information for valid richness comparisons, including category–subcategory ratios (species-to-genus and species-to-individual ratios). Rarefaction methods – both sample-based and individual-based – allow for meaningful standardization and comparison of datasets. Standardizing data sets by area or sampling effort may produce very different results compared to standardizing by number of individuals collected, and it is not always clear which measure of diversity is more appropriate. Asymptotic richness estimators provide lower-bound estimates for taxon-rich groups such as tropical arthropods, in which observed richness rarely reaches an asymptote, despite intensive sampling. Recent examples of diversity studies of tropical trees, stream invertebrates, and herbaceous plants emphasize the importance of carefully quantifying species richness using taxon sampling curves.

Keywords

Species richness, species density, taxon sampling, taxonomic ratios, biodiversity, rarefaction, accumulation curves, asymptotic richness, richness estimation, category–subcategory ratios.

Ecology Letters (2001) 4: 379–391

Species richness is the simplest way to describe community and regional diversity (Magurran 1988), and this variable – number of species – forms the basis of many ecological models of community structure (MacArthur & Wilson 1967; Connell 1978; Stevens 1989). Quantifying species richness is important, not only for basic comparisons among sites, but also for addressing the saturation of local communities colonized from regional source pools (Cornell 1999). Maximizing species richness is often an explicit or implicit goal of conservation studies (May 1988), and current and background rates of species extinction are calibrated against patterns of species richness (Simberloff 1986). Therefore, it is important to examine how ecologists have quantified this fundamental measure of biodiversity and to highlight some recurrent pitfalls. Even the most recent reviews of biodiversity assessment (Lawton *et al.* 1998; Gaston 2000; Purvis & Hector 2000) have not discussed the sampling issues we address in this review in relation to the measurement and comparison of species richness. In contrast, the uses and abuses of species diversity indices, which, by design, combine richness with relative abundance, enjoy a substantial and venerable literature (e.g. Washington 1984), and are thus beyond the scope of this review. We begin by placing several concepts of diverse origin in a common conceptual framework.

TAXON SAMPLING CURVES

Although species richness is a natural measure of biodiversity, it is an elusive quantity to measure properly (May 1988). The problem is that, for diverse taxa, as more individuals are sampled, more species will be recorded (Bunge & Fitzpatrick 1993). The same, of course, is true for higher taxa, such as genera or families. This sampling curve rises relatively rapidly at first, then much more slowly in later samples as increasingly rare taxa are added.

In principle, for a survey of some well-defined spatial scope, an asymptote will eventually be reached and no further taxa will be added.

We distinguish four kinds of *taxon sampling curves*, based on two dichotomies (Fig. 1). Although we will present these curves in terms of species richness, they apply just as well to richness of higher taxa.

The first dichotomy concerns the sampling protocol used to assess species richness. Suppose one wishes to compare the number of tree species in two contrasting 10-ha forest plots. One approach is to examine some number of individual trees at random within each plot, recording sequentially the species identity of one tree after another. We refer to such an assessment protocol as *individual-based* (Fig. 1). Alternatively, one could establish a series of quadrats in each plot, record the number and identity of all the trees within each, and accumulate the total number of species as additional quadrats are censused (e.g. Cannon *et al.* 1998; Chazdon *et al.* 1998; Hubbell *et al.* 1999; Vandermeer *et al.* 2000). This is an example of a *sample-based* assessment
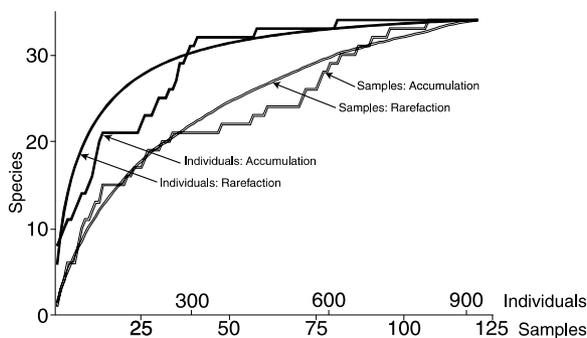
Figure 1 Sample- and individual-based rarefaction and accumulation curves. Accumulation curves (jagged curves) represent a single ordering of individuals (solid-line, jagged curve) or samples (open-line, jagged curve), as they are successively pooled. Rarefaction curves (smooth curves) represent the means of repeated re-sampling of all pooled individuals (solid-line, smooth curve) or all pooled samples (open-line, smooth curve). The smoothed rarefaction curves thus represent the statistical expectation for the corresponding accumulation curves. The sample-based curves lie below the individual-based curves because of the spatial aggregation of species. All four curves are based on the benchmark seedbank dataset of Butler & Chazdon (1998), analysed by Colwell & Coddington (1994) and available online with *EstimateS* (Colwell 2000a). The individual-based accumulation curve shows one particular random ordering of all individuals pooled. The individual-based rarefaction curve was computed by *EstimateS* using the Coleman method (Coleman 1981). The sample-based accumulation curve shows one particular random ordering of all samples in the dataset. The sample-based rarefaction curve was computed by repeated re-sampling, using *EstimateS*. For both sample-based curves, the patchiness parameter in *EstimateS* set to 0.8, to emphasize the effect of spatial aggregation.

protocol (Fig. 1). The relative merit of these approaches for estimating species richness of trees is not the point here. Rather, we emphasize that species richness censuses can be validly based on datasets consisting either of individuals or of replicated, multi-individual samples. The key distinction is the unit of replication: the individual vs. a sample of individuals – a distinction that turns out to be far from trivial.

Examples of individual-based protocols include birders' "life lists" (e.g. Howard & Moore 1984), Christmas bird counts (e.g. Robbins *et al.* 1989), time-based "collector's curves" (e.g. Clench 1979; Lamas *et al.* 1991), and taxon-richness counts (often families or genera) from palaeontological sites (e.g. Raup 1979). In addition, when an unreplicated mass sample (such as a deep-sea dredge sample, e.g. Sanders 1968) is treated as if it were set of randomly captured individuals from the source habitat, an individual-based taxon-sampling curve can be produced for the sample. Examples of sample-based protocols using sampling units other than quadrats include replicated mist-net samples for birds (Melhop & Lynch 1986) and replicated trap data for arthropods (e.g. Stork 1991; Longino & Colwell 1997; Gotelli & Arnett 2000).

A "hybrid" between individual-based and sample-based taxon sampling curves is produced by the "*m*-species list" method, in current use by some ornithologists (e.g. Poulsen *et al.* 1997). A list is kept of the first *m* (usually 20) species observed (disregarding abundances) in a sampling area – an individual-based list. Then, additional "samples", each based on a new list of *m* species from the same area, are successively pooled. The cumulative number of species observed is plotted as a function of the number of *m*-species lists pooled to produce a curve that reaches an asymptote when all species have been observed.

Sample- and individual-based data sets are sometimes treated interchangeably in statistical analyses. For example, depending upon the scale of interest or the focus of a hypothesis (in the sense of Scheiner *et al.* 2000), a group of individual-based datasets or mass samples can be analysed as if they were replicate samples from the same statistical universe (e.g. Grassle & Maciolek 1992). Likewise, a set of replicated samples can usually be pooled and treated as a single, individual-based dataset, for some purposes (Engstrom & James 1981). (This is not possible with *m*-species-list curves, since abundances are not recorded.)

The second dichotomy distinguishes *accumulation curves* from *rarefaction curves*. A species (or higher taxon) *accumulation curve* records the total number of species revealed, during the process of data collection, as additional individuals or sample units are added to the *pool* of all previously observed or collected individuals or samples (Fig. 1). Accumulation curves may be either individual-based (e.g. Clench 1979; Robbins *et al.* 1989) or sample-based (e.g. Novotny & Basset 2000).

In contrast, a *rarefaction curve* is produced by repeatedly re-sampling the pool of N individuals or N samples, at random, plotting the average number of species represented by 1, 2,…N individuals or samples (Fig. 1). Sampling is generally done without replacement, within each re-sampling. Thus, rarefaction generates the expected number of species in a small collection of n individuals (or n samples) drawn at random from the large pool of N individuals (or N samples; Simberloff 1978).

These two dichotomies jointly define four kinds of taxon sampling curves, as shown in Fig. 1. Accumulation curves, in effect, move from left to right, as they are further extended by additional sampling. In contrast, rarefaction curves move from right to left, as the full dataset is increasingly "rarefied". Because the entire rarefaction curve depends upon every individual or sample in the pool at the accumulation curve's right-hand end, each individual or sample is equally likely to be included in the mean richness value for any level of re-sampling along the rarefaction curve. The corresponding rarefaction and accumulation curves are closely related to one another. Indeed, a rarefaction curve, whether based on individuals or on samples, can be viewed as the statistical expectation of the corresponding accumulation curve, over different reorderings of the individuals or samples.

In Fig. 1, note that the two sample-based curves lie below the two individual-based curves. The reason for this nearly universal pattern is that sample-based protocols aggregate individuals, within each sample, that are nearby in space or consecutive in time. Any spatial or temporal autocorrelation (patchiness or heterogeneity) in taxon occurrence will cause taxa to occur nonrandomly among samples. Consequently, when a group of samples is pooled, fewer species will be represented by those individuals than by an equal number of individuals censused randomly and independently in the same habitat.

Although the four kinds of taxon sampling curves in Fig. 1 provide a unifying framework for measuring species richness, they do not fully conform to current terminology. Sanders (1968) first used individual-based rarefaction to compare species richness *among* benthic marine mass collections. Noting that collections differed not only in number of species but also in number of individuals, Sanders suggested "rarefying" each collection to a common number of individuals, to match the size of the smallest collection.

Following Sanders, the term *rarefaction* has historically referred to *individual-based* taxon re-sampling curves. Although *sample-based* taxon re-sampling curves are precisely analogous, they have usually been referred to, instead, as "randomized", or "smoothed" species accumulation curves (e.g. Colwell & Coddington 1994) – an equally accurate characterization, which we do not oppose. The randomized

sample accumulation curve of Pielou's (1966, 1975) "pooled quadrat method" is effectively the same method, although originally intended to be used in the estimation of diversity indices.

COMPARING ASSEMBLAGES USING TAXON SAMPLING CURVES

Comparing species or higher-taxon richness without reference to a taxon sampling curve is problematic at best. Communities may differ in measured species richness because of differences in underlying species richness, differences in the shape of the relative abundance distribution, or because of differences in the number of individuals counted or collected (Denslow 1995). Differences in numbers of individuals counted may themselves reflect biologically meaningful patterns of resource availability or growth conditions. However, differences in abundance may also reflect differences in sampling effort or conditions for collection or observation. Comparing raw taxon counts for two or more assemblages will quite generally produce misleading results.

Raw species richness counts or higher taxon counts can be validly compared only when taxon accumulation curves have reached a clear asymptote. For invertebrate and microbial assemblages everywhere and for many taxa in tropical habitats, such asymptotes may never be reached (e.g. Stork 1991; Wolda *et al.* 1998; Fisher 1999; Anderson & Ashe 2000; Novotny & Basset 2000). Fortunately, if one or more accumulation curves fail to reach an asymptote, the curves themselves may often be compared, after appropriate scaling.

For individual-based datasets, it is not always possible to construct an accumulation curve as in Fig. 1. The order of identification of individuals within each sample may not have been recorded, or the collection may consist of mass captures. In such cases, rarefaction produces the only appropriate curves for dataset comparisons. Even when the order of individual identification is known (as in time-series data), rarefaction produces smooth curves that facilitate comparison. Likewise, in the case of sample-based datasets, sample order is often unavailable or arbitrary. Repeated, averaged sample-based rarefaction produces smooth curves for comparison, allowing standardization of sampling effort.

Whether to use individual-based or sample-based rarefaction to compare richness depends upon the data available. If the data are inherently individual-based, there is no alternative to using individual-based rarefaction to compare assemblages. If sample-based data are available, however, either sample-based or individual-based rarefaction could be used, but it is generally preferable to use the sample-based approach, to account for natural levels of sample heterogeneity (patchiness) in the data. For patchy

distributions, individual-based rarefaction inevitably overestimates the number of species (or higher taxa) that would have been found with less effort. In fact, the difference between the sample-based and individual-based rarefaction curves can be used as a measure of patchiness (Colwell & Coddington 1994).

Regardless of which approach is used, it is the individual that carries taxonomic information. When sample-based rarefaction curves are used to compare taxon richness at comparable levels of sampling effort, the number of taxa should be plotted as a function of the accumulated number of *individuals*, not accumulated number of *samples*, because datasets may differ systematically in the mean number of individuals per sample. (Here, we are assuming that taxon richness is the question, not taxon density; see below.)

An example makes this pitfall clear. Suppose you wish to know whether tropical old-growth forest or nearby tropical second-growth forest is richer in tree species. You identify all individual stems in *n* 10 × 10 m randomly placed quadrats in each forest type. The sample rarefaction curve for second-growth forest, plotted as a function of *samples*, lies above the corresponding curve for old-growth forest, but neither has reached an asymptote (Fig. 2a). The mean number of stems per quadrat is considerably greater in the second-growth forest, as would be expected. Are there really more species in the second-growth forest? Not even an approximate answer can be given to this question without re-scaling the *x*-axis to number of individuals (based on the average number of individuals per sample). Once re-scaled, the second-growth forest curve will drop relative to the old-growth forest curve; it may (still) lie above it, coincide, or fall below it (Fig. 2b). (Cannon *et al*. 1998 demonstrated this pitfall for logged vs. unlogged forests, which differ in stem density and in quadrat-based richness, but have similar species richness when re-scaled to individuals.) This example illustrates the importance of using taxon sampling curves to compare species richness, even when the comparisons are based on standardized methods and identical sampling protocols.
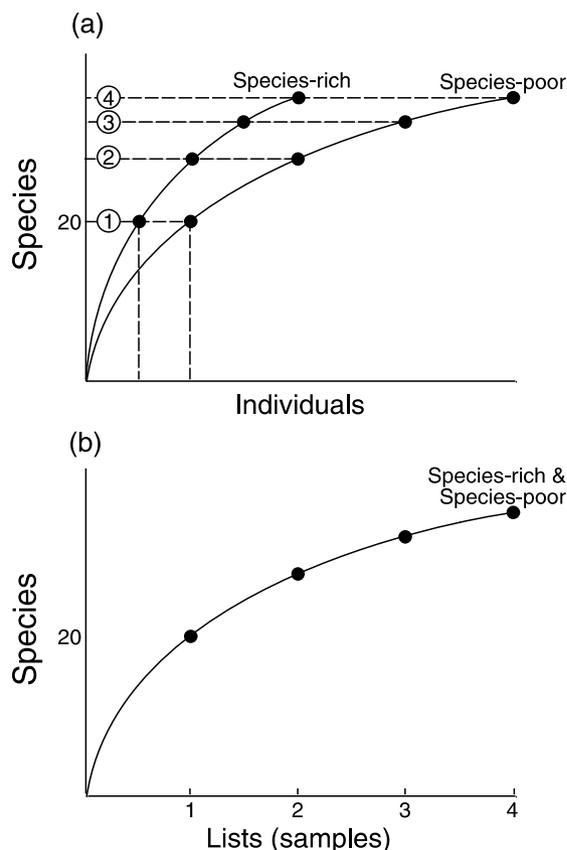
The *m*-species-list method (Poulsen *et al*. 1997) suffers from a related pitfall. Suppose two communities are sampled with this method, one more species-rich than the other, using 20-species lists. In the poorer community, for each 20-species sample, more individuals will need to be observed than in the richer community to reach 20 species. Thus, as samples accumulate, the poorer community will be increasingly better sampled than the richer one because more individuals will have been sampled. In fact, this bias may be strong enough that the cumulative number of species revealed in the poor community equals or exceeds that of the rich community, for the same number of 20-species samples, as long as both curves are increasing – as would often be the case for a rapid-assessment survey (Fig. 3). Of
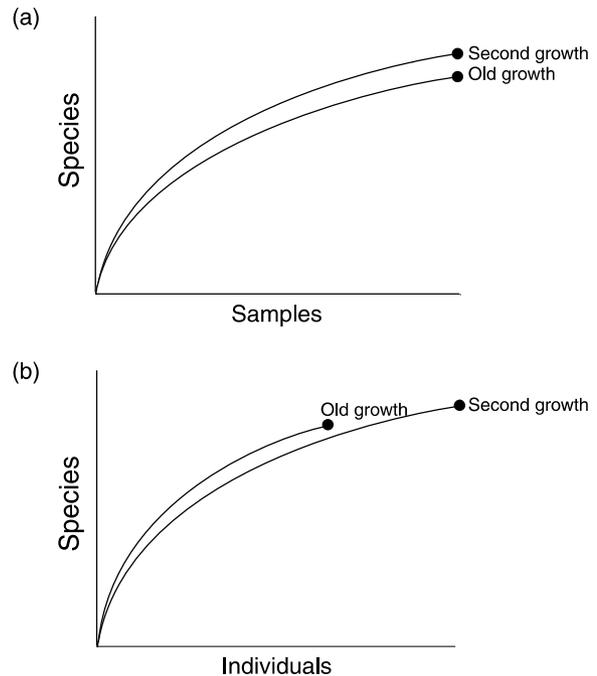
Figure 2 The effect on species richness of re-scaling the *x*-axis of sample-based rarefaction curves (randomized species accumulation curves) from samples to individuals, when individual densities vary. In this hypothetical example, species richness appears to be higher for a second-growth forest stand than for an old growth stand (a, based on corresponding numbers of accumulated samples. However, stem density is higher in the second-growth stand (with smaller trees) than for the old-growth stands (with larger trees). When the *x*-axis is re-scaled to individuals, the result is reversed (b).

course, eventually, the 20-species-sample accumulation curve for both communities will reach their asymptotes (the species-poor community first) and the curves will diverge, but the wrong inference can easily be made if both curves are still rising when sampling is stopped (Fig. 3). In short, it is perilous or impossible to make a valid comparison between two species accumulation curves that are based on the *m*-species-list method, unless both curves have reached an asymptote.

Other pitfalls to watch out for apply to individual-based rarefaction as well as sample-based rarefaction. A valid individual-based rarefaction analysis assumes not only that the spatial distribution of individuals in the environment is random (Kobayashi 1982), as discussed above, but that sample sizes are sufficient, and that assemblages being compared have been sampled in the same way (Abele & Walters 1979). If sample sizes are not sufficient, rarefaction will not distinguish between different richness patterns, because all rarefaction curves tend to converge at low abundances (Tipper 1979). If the assemblages are

Figure 3 A pitfall of the "*m*-species list" method of comparing species richness. In this method (Poulsen *et al.* 1997), lists of the first 20 (or other constant) species observed in repeated samples are accumulated, without regard to the number of individuals actually examined to reach 20 species. As this hypothetical example shows, in a species-poor community, more individuals will inevitably have to be examined to reach each successive set of 20 species than in a species-rich community (a). Nevertheless, as samples 1, 2, 3, 4… are pooled, in this example an identical cumulative number of species is reached as species are plotted against number of lists (1, 2, 3, 4…) on the *x*-axis (as is standard for the *m*-species list method) (b). In fact, the individual-based accumulation curves could be arranged to achieve a variety of misleading results, when cumulative species are plotted against number of lists (samples).

taxonomically very different, the sampling may not adequately characterize each taxon (Simberloff 1978). If the sampling methods are not identical, different kinds of species may be over- or under-represented in different samples, because no sampling method is completely random and unbiased (Boulinier *et al.* 1998). In addition, the shape of individual-based rarefaction curves depends upon relative abundance – the greater the evenness of the relative abundance distribution, the steeper the rarefaction curve (Gotelli & Graves 1996). For this reason, rarefaction curves

for two communities with different patterns of relative abundance may cross once or even twice. Likewise, sample-based rarefaction can cross, if based on communities that differ sufficiently in patchiness. Thus, the sample size to which one rarefies can potentially change the rank order of estimated richness among communities.

COMPUTING RAREFACTION CURVES

Individual-based rarefaction

For individual-based rarefaction curves, a precise mathematical expression based on combinatoric theory can be computed for expected richness, given *n* individuals, instead of actually re-sampling to randomize. Sanders (1968) provided what was intended as an individual-based rarefaction formula for calculating the expected number of species in a random subsample of individuals from a single, large collection. Although the principle of rarefaction was sound, Sanders derived the rarefaction formula incorrectly (Hurlbert 1971). The correct derivation is based on a hypergeometric sampling distribution, in which individuals are sampled randomly and without replacement (Heck *et al.* 1975).

From this model, both the expected number of species and its variance can be derived. A mathematically distinct but computationally much faster way to produce individual-based rarefaction curves is to compute the corresponding "random placement" curve of Coleman (1981; Coleman *et al.* 1982), which has been shown to very closely approximate the hypergeometric rarefaction curve (Brewer & Williamson 1994; Colwell & Coddington 1994).

Some theoretical progress has been made in modifying the rarefaction curve for cases of known spatial distributions, such as the negative binomial (Kobayashi 1982, 1983; Smith *et al.* 1985). However, these analyses still assume that individuals have been sampled randomly. In reality, ecologists rarely sample individuals randomly. Instead, quadrats or sampling devices are implemented randomly (or in stratified random design), and all of the individuals in a small collection are sorted, yielding datasets appropriate for sample-based rarefaction.

Sample-based rarefaction

Because the sample-based rarefaction curve depends on the spatial distribution of individuals as well as the size and placement of samples (Hurlbert 1990), it cannot be derived theoretically. Thus, computations require Monte Carlo re-sampling, in which samples are randomly accumulated in many iterations. Free software is available (Colwell 2000a) to compute sample-based rarefaction curves as well as the corresponding individual-based Coleman curves. Mean

number of accumulated individuals is also computed, to allow re-scaling of sample-based rarefaction curves. Free software is also available for the construction of individual-based rarefaction curves and confidence intervals for species richness and other diversity indices (Gotelli & Entsminger 2001).

CATEGORY-SUBCATEGORY RATIOS AND THEIR PITFALLS

Individuals and species

To introduce the concept, and the perils, of what we call *category–subcategory ratios*, let us return to the example (above) of assessing tree species richness in old-growth vs. second-growth forest. Recall that the problem with comparing sample-based rarefaction curves scaled by number of samples was that second-growth quadrats each had more stems than equal-sized old-growth quadrats, on average. Why not simply compare average species per stem, among quadrats, for each forest type, to remove the effect of stem density? This index is the species-per-individual ratio, a particular class of category-subcategory ratios.

Figure 4 illustrates the hazards of using the species-per-individual ratio to compare samples. Each panel in Fig. 4 shows hypothetical, sample-based rarefaction curves for contrasting forest habitats. Each curve is based on the same number of quadrats, but each is re-scaled to the number of individuals on the *x*-axis. The solid dots indicate total richness for the pooled quadrats in each forest habitat. The slopes of the lines connecting these points to the origin equal the ratio of species to individuals for the dots. In Fig. 4(a), old-growth and second-growth forest have identical species richness (at least as far as the curves extend), yet the number of species per individual is much lower for the second-growth forest. In Fig. 4(b), species richness is higher in forest gaps than in non-gaps (forest matrix), yet the number of species per individual is identical for total richness in gaps and non-gaps.

An example from the recent literature illustrates the perils of "normalizing" richness by dividing the number of species by the number of individuals. In support of their inference that tree species richness does not differ between gaps and non-gaps, Hubbell *et al.* (1999) showed that number of species divided by number of stems did not differ for saplings in gaps vs. non-gaps in a Panamanian forest. Using Hubbell's reported stem densities and richness values for saplings in 20 × 20-m quadrats, Chazdon *et al.* (1999) showed that true sapling species richness might in fact fit curves such as those in Fig. 4(b) (see also Kobe 1999; Vandermeer *et al.* 2000), with greater total richness in gaps. In his reply, Hubbell (1999) failed to provide the individual-based species accumulation curves to disprove Chazdon's
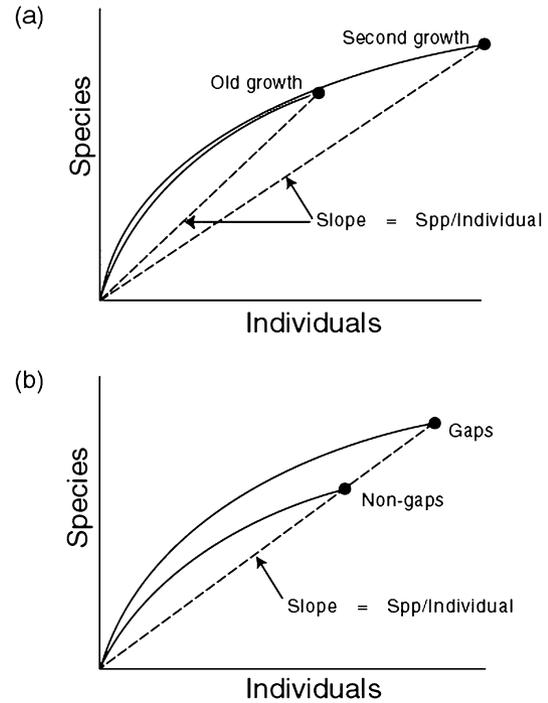
Figure 4 Pitfalls of using species/individual ratios to compare datasets. In (a), an old-growth and a second-growth forest stand are compared. The 2 stands have identical individual-scaled rarefaction curves, and thus do not differ in species richness. The second growth curve extends farther simply because stem density is greater, so that more individuals have been examined for the same number of samples. However, when the ratio of species/individual is computed for each, the ratio is much higher for the old-growth stand. In (b), species richness in treefall gap quadrats is compared with richness in non-gap (forest matrix) quadrats. In this case, species/individual ratios are identical, yet the true species richness is higher in gaps.
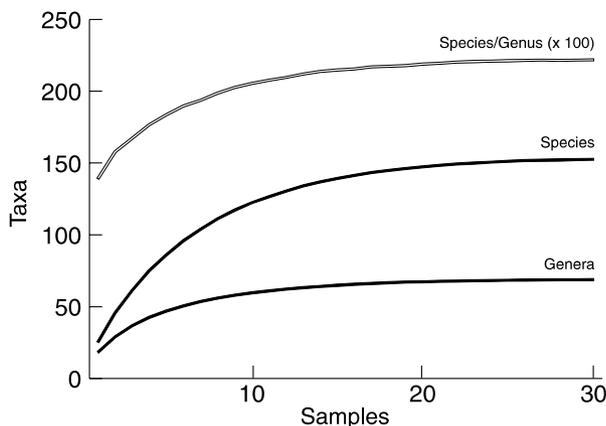
conjecture for the sapling dataset at issue. Instead, Hubbell *et al.* (1999) provided individual-based accumulation curves for a quite different dataset (no size class specified) and cited the fact that *area*-based accumulation curves do not differ for gaps and non-gaps, leaving the debate unresolved. Our point here is simply that, had individual-based accumulation curves been published for the sapling dataset at issue in the first place, the ambiguity that instigated the debate would never have arisen.

Using the species-per-individual ratio to correct for unequal numbers of individuals is invalid because it assumes that richness increases linearly with abundance – true only for the idealized case of extreme unevenness, in which one species is maximally dominant (Gotelli & Graves 1996). Because abundances are rarely this extreme, the species-per-abundance ratio will distort patterns of species richness.

The same problem affects the inverse ratio, individuals per species (e.g. Irwin 1997; his table 4.1), which Coddington (Coddington et al. 1991, 1996; Silva & Coddington 1996) suggested as a measure of "sampling intensity." For communities that are more or less equivalent in total richness and relative abundance patterns, the number of individuals per species (total number of individuals divided by total number of species) is indeed a decent rule of thumb for relative completeness of inventories. However, computing this ratio can be a misleading way to "standardize" sampling effort when comparing communities differing considerably in richness, or for which comparative richness is unknown. For example, in Fig. 4(b), the number of individuals per species is the same for both datasets, yet sampling is obviously more complete for the gap habitats.

Species and genera

In biogeography, category–subcategory taxonomic ratios have been repeatedly used and abused. The best known of these is the species–genus ratio, but family–order ratios or any other lower-taxon–higher-taxon ratios are subject to the same pitfalls. Figure 5 depicts sample-based species and genus rarefaction curves for a hummingbird dataset (Colwell 2000b). As the number of samples increases, the number of genera reaches an asymptote sooner than the number of species. This pattern is inevitable for any two taxonomic ranks (except in the unrealistic case of 100% monobasic taxa) since the higher rank (the genus, in this example) inevitably has fewer members than the lower rank (species). Because of this relationship, a plot of the number of subtaxa per taxon (species per genus, in this example – the upper curve in Fig. 5) always has a positive slope, as seen in Fig. 5.

The species-to-genus ratio has long been used to describe community patterns and to infer levels of competitive interactions among species within genera (reviews in Simberloff 1970; Järvinen 1982). Similar reasoning has been applied to the interpretation of species–family and other taxonomic ratios (e.g. MacArthur & Wilson 1967; Cook 1969). A low species-to-genus ratio was interpreted as a product of strong intrageneric competition (Elton 1946), which might limit congeneric coexistence (Darwin 1859). Consistent with this hypothesis was the widespread observation that species-to-genus ratios were usually smaller for island than mainland communities (Elton 1946). However, subtaxon–taxon ratios are an increasing function of sample size, and would be expected to decrease in small communities, regardless of the level of competition (Williams 1947, 1964; Simberloff 1970, 1972). Figure 6 shows this effect graphically by re-plotting the data of Fig. 5 as number of species as a function of number of genera. The slope of the diagonal broken lines shows that, in a small random sample (few species), there are fewer species per genus than in a large random sample (more species).

Sample-size dependence in taxonomic ratios was first demonstrated for plant communities by Maillefer (1929), who used draws of species from a deck of shuffled cards to calculate the expected generic richness in small
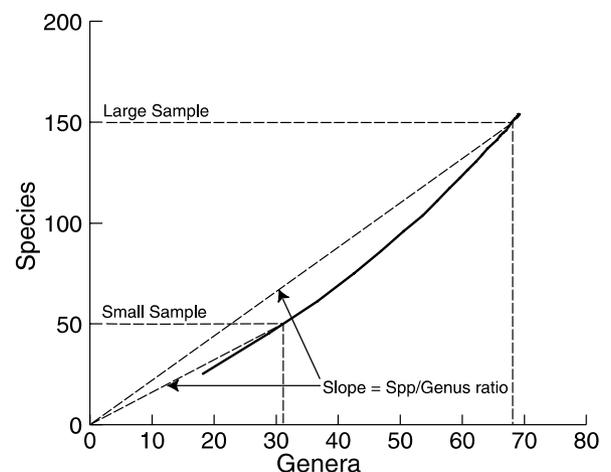

Figure 5 Taxon sampling curves for species and for the genera to which they belong, with the species–genus ratio. Note that the curve for genera reaches its asymptote at a smaller number of samples than the species curve. For this reason, the ratio of species to genera is nonlinear. This patterns is inevitable for any case of category-subcategory sampling curves. The curves are based on a sample of hummingbird specimens (Colwell 2000b; appendix B, pooled, distributed into "samples" at random, and then repeatedly re-sampled using *EstimateS* (Colwell 2000a).

Figure 6 The species-per-genus pitfall. The solid-line curve plots number of species as a function of number of genera for the hummingbird data of Fig. 5. Because the relationship is nonlinear with an increasing slope, the species–genus ratio (the slope of the broken lines) is greater for a larger sample of species than for a smaller sample.

communities. For animal communities, Williams (1947, 1964) elucidated these same patterns using species-abundance models and computer simulations. Although their work was ignored by ecologists for several decades (Järvinen 1982), re-analyses of species-to-genus ratios now suggest that island communities harbour slightly *more* species per genus than expected by chance, in spite of the lower absolute number of species per genus expected in smaller samples (Simberloff 1970). This finding is the opposite of what competition theory predicts, perhaps reflecting instead the similar dispersal potential and ecological requirements of congeneric species (the Icarus Effect of Colwell & Winkler 1984). Despite the periodic rediscovery of this classic pitfall, sample-size dependence of taxonomic ratios continues to trap the unwary (e.g. Ashton 1998).

SPECIES RICHNESS VS. SPECIES DENSITY

We have emphasized the importance of using taxon sampling curves (both individual- and sample-based) to standardize datasets to a common number of individuals for the purposes of comparing species richness. In contrast, most community ecology studies standardize on the basis of area or sampling effort. Thus, most ecological comparisons of biodiversity are actually comparisons of *species density*: the number of species per unit area (Simpson 1964). Such studies hinge on the assumption that samples are drawn from populations of individuals that are at comparable densities. However, species density depends on both species richness and on the mean density of individuals (disregarding species), as discussed in relation to the example of old-growth vs. second-growth forest above (Fig. 2). Consequently, the ordering of communities may differ when ranked by species richness vs. species density (James & Wamer 1982; McCabe & Gotelli 2000).

Both species richness and species density can be compared using sample- and individual-based rarefaction curves (Fig. 7). Individual-based rarefaction curves standardize each of two or more samples on the basis of the number of individuals, for the purpose of comparing species richness. Sample-based rarefaction curves can be used to compare richness in the same way, as long as the *x*-axis is re-scaled in units of individuals. In contrast, to compare species density when samples are derived from incommensurate areas, the *x*-axis of individual-based rarefaction curves can be rescaled from individuals to area, based on average density. Likewise, if sample-based rarefaction curves are simply left scaled by number of accumulated samples (instead of re-scaling to individuals), then comparisons among datasets will be in terms of species density, instead of species richness (Fig. 7), assuming samples are space-based.

In comparisons of species density, a familiar pitfall awaits the unwary, but in a new guise. For a constant density, area is

(a)

(b)
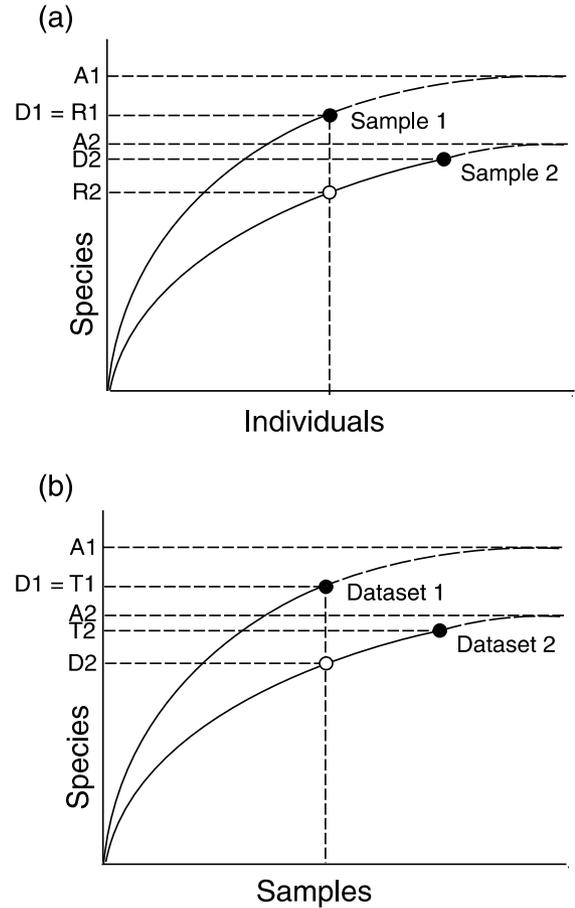
Figure 7 Species richness vs. species density. Part (a) shows individual-based rarefaction curves for two contrasting samples, whereas (b) shows sample-based rarefaction curves for two contrasting datasets. A1 and A2 indicate the asymptotic richness for the two curves in each panel. In (a), raw species totals for the two samples (black dots) measure species *density* (D1 and D2 – assuming each sample covers the same area), whereas Sample 2 must be rarefied to the same number of individuals as Sample 1 (the open dot) to allow a valid comparison of species *richness* (R2 vs. R1). In (b), raw species totals for the two datasets (black dots) measure total species *richness* (T1 and T2) for the datasets. Assuming each sample covers the same amount of space, Dataset 2 must be rarefied to the same number of samples as Dataset 1 (the open dot) to allow a valid comparison of species *density* (D2 vs. D1). (For a valid comparison of richness between the two datasets in the lower panel, the *x*-axis would have to be re-scaled to individuals, as in Figs 2 and 4.)

a proxy for number of individuals. Thus, "normalizing" species density data of two unequal areas by dividing the number of species by the area measured is subject to the very same pitfalls as the species-per-individual ratio, as shown in Fig. 4. Although it sounds paradoxical, the ratio of richness

to area is not a valid measure of species density, because the number of species increases nonlinearly with area. Instead, species density is validly compared only with the appropriate taxon sampling curves (e.g. James & Warner 1982).

Which measure is more appropriate, species richness or species density? In other words, should communities be compared on the basis of a standardized number of individuals (species richness) or a standardized area or sampling unit (species density)? For conservation purposes and applied problems that focus on large areas, species density is probably of more interest because it measures the number of species within a specified area. However, rarefaction should nevertheless be used when comparing species density in different regions, to assess the degree to which differences in species density can be attributed to patterns of individual abundance (which determines the position of the community on the x-axis of the species accumulation curve), and how much can be attributed to the shape and magnitude of the species accumulation curve (which determines the species richness achieved at a particular level of individual abundance).

On the other hand, for testing models and evaluating theoretical predictions in ecology, species richness may be more appropriate. Most theoretical models in community ecology do not contain explicit terms for area or density. Instead, the currency of these models is abundance (N) and population growth rates ($\mathrm{d}N/\mathrm{d}t$), which are modified by *per capita* coefficients that describe interactions with other species (Gotelli 2001). *Per capita* interactions may be expressed in samples that are based on common numbers of individuals, which is how species richness is measured with individual-based rarefaction or with sample-based rarefaction scaled to individuals. We stress that neither species density nor species richness is necessarily the "correct" way to measure diversity, but that patterns of diversity will be very sensitive to which measure is used. Conservation decisions may be complicated when some reserves or candidate areas contain higher species density and others contain higher species richness. Disturbance or management regimes that affect abundance might have to be considered in choosing among such areas.

Recent studies of species richness and species density patterns have led to a re-evaluation of some familiar patterns of diversity in natural communities. For example, many experimental and correlative studies have documented that disturbances reduce the diversity of benthic invertebrate assemblages in streams (Lake 1990; Vinson & Hawkins 1998). However, most of these studies have quantified species diversity as species density, the number of species per unit area. Because ecological disturbances reduce abundance, we would expect disturbance to decrease species density, simply because there will be fewer individuals present to be sampled after a disturbance.

In an experimental study of northern U.S. stream assemblages, McCabe & Gotelli (2000) manipulated the area, intensity, and frequency of disturbance on artificial substrates in an orthogonal 3-way design. Macroinvertebrates were collected from substrate surfaces after 6 weeks of treatment application. Species density (number of taxa per sample) was significantly reduced in all disturbance treatments compared to unmanipulated controls. However, when the treatments were compared by individual-based rarefaction, the patterns were completely reversed: for a fixed number of individuals, taxon richness was higher in all disturbance treatments than in the undisturbed controls (Fig. 8). This example demonstrates the importance of using the species accumulation curve to carefully quantify taxon richness – even in experimental studies in which sampling effort is carefully standardized.

Similarly, plant ecologists have repeatedly made the error of comparing richness per quadrat (species density) among stands differing in overall plant density (Fig. 2). These comparisons have confounded or equated differences in density with the differences in disturbance, successional, or productivity regimes that are being compared. As discussed above, attempting to correct for this error by computing species per stem (Fig. 4) leads to the same pitfall as computing species per genus for samples differing in numbers of species (Fig. 6).

For example, a frequent ecological pattern is the hump-shaped diversity curve, in which species richness peaks at intermediate productivity levels (DiTomasso & Aarssen 1989; Rosenzweig & Abramsky 1993). Many models of plant competition assume that mortality is not equal among species, so that interspecific competition leads to species losses at high levels of fertility (Tilman 1982, 1988; Huston & DeAngelis 1994; but see Abrams 1995).

However, the *assemblage-level thinning hypothesis* (Oksanen 1996) accounts for the hump-shaped diversity curve by variation in total plant density. As fertility increases, individuals get larger, crowding occurs, and density (number of plants/area) decreases. Therefore, rare species are "lost" at high densities because they are represented by few individuals, not because of differential mortality or interspecific differences in competitive ability.

To test the assemblage-level thinning hypothesis, Stevens & Carson (1999) established an experimental productivity gradient in 1-year-old fields in the north-eastern U.S. They found that both species number and density of herbaceous plants declined at high fertility. A simulation model similar to individual-based rarefaction established that random survivorship of individuals could largely account for the decline in diversity at high productivity. Although many plant assemblages are characterized by strong pairwise competitive interactions (Shipley 1993), net competitive effects in multispecies communities may be weak (Miller 1994), and
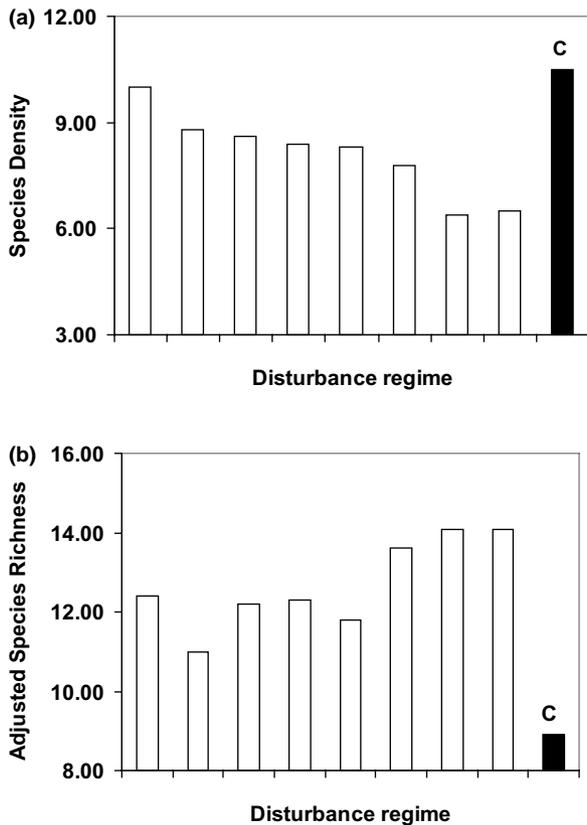
Figure 8 Contrasting results for species density vs. species richness in assessing patterns of response to disturbance among aquatic invertebrate assemblages. Each open bar is the average diversity in one of eight experimental disturbance regimes, and the solid bar is the average diversity in unmanipulated controls (C) ($n = 7$ replicates/treatment). The eight regimes are derived from a fully crossed three-factor experiment with two levels of disturbance frequency (one or two disturbances/week), disturbance area (50% or 100%), and disturbance intensity (light vs. heavy scraping) applied to artificial substrates in a Vermont stream; (a) shows the conventional measure of species density (species number/sample); (b) shows the same data, but the response variable has been calculated from an individual-based rarefaction curve constructed for each replicate then standardized to a common number of randomly subsampled individuals. In both analyses, treatment means differ significantly by ANOVA ($P < 0.01$). However, the patterns of diversity are opposite for species density vs. species richness measures. Figure adapted and simplified from McCabe & Gotelli (2000).

simple changes in density may be the primary determinant of species richness across productivity gradients.

ASYMPTOTIC ESTIMATORS OF SPECIES RICHNESS

Estimates of asymptotic species richness may be especially important in biotic inventories and surveys, where it is impractical to exhaustively sample species rich communities, such as tropical invertebrate, microbial or plant communities (e.g. Cannon et al. 1998; Fisher 1999; Novotny & Basset 2000). Rarefaction (either individual-based or sample-based) is a method for interpolating to smaller samples and estimating species richness in the rising part of the taxon sampling curve. However, rarefaction cannot be used for extrapolation; it does not provide an estimate of asymptotic richness (Tipper 1979).

Statistical studies have produced a large number of estimators of the asymptotic number of "classes" for samples of classified objects (reviewed by Bunge & Fitzpatrick 1993), of which species richness is one example. The most promising of these are nonparametric estimators based on mark and recapture statistics (Colwell & Coddington 1994; Nichols & Conroy 1996; Boulinier et al. 1998; Chazdon et al. 1998; Colwell 2000a). The nonparametric estimators use information on the distribution of rare species in the assemblage – those represented by only one (singletons), two (doubletons) or a few individuals. The greater the number of rare species in a dataset, the more likely it is that other species are present that were not represented in the dataset. In addition, asymptotic (and nonasymptotic) richness may be estimated by curve-fitting extrapolation methods (e.g. Palmer 1990; Lamas et al. 1991; Soberón & Llorente 1993; Mawdsley 1996; Keating & Quinn 1998; Fisher 1999).

Although extrapolation is inherently more risky than interpolation, some of these asymptotic estimators have so far performed well when tested on exhaustively censused, benchmark datasets in which the species sampling curve reaches a stable asymptote [such as the tropical seedbank dataset of Butler & Chazdon 1998 (analysed by Colwell & Coddington 1994) or the parasite data of Walther et al. 1995]. A richness estimator is tested on such a benchmark dataset by computing the sample-based rarefaction curve, then computing the estimator for each cumulative level of sample pooling, following Pielou (1966, 1975; Colwell & Coddington 1994). By repeating the computations for all levels of sample accumulation, a continuous plot of the estimator can be displayed along with the sample-based rarefaction. Resampling and recomputing the estimators repeatedly and taking means produces smooth curves. An ideal estimator would (1) reach its *own* asymptote much sooner than the sample-based rarefaction curve levels off, and (2) approximate the empirical asymptote in an unbiased way, when tested over many benchmark datasets (Anderson & Ashe 2000 provide numerous examples for tropical beetles).

Of course, aside from testing estimators, there is no reason to use an estimator for a dataset that reaches a steady asymptote. The datasets that need richness estimators are those that, as yet, are nowhere near an asymptote, such as

most tropical arthropod datasets (e.g. Stork 1991; Wolda *et al.* 1998; Fisher 1999; Novotny & Basset 2000). The tricky issue is whether the performance of the estimators on benchmark datasets – which usually consist of relatively small numbers of species – accurately predicts the performance of the same estimators on not-yet-asymptotic datasets, which usually consist of very large numbers of species. One indication of the failure of the existing catalogue of estimators for hyperdiverse taxa is that they often fail to reach any asymptote at all, rising more or less in parallel with the still-steep sample-based rarefaction curve (e.g. Fisher 1999). In these cases, the estimators must be viewed as providing only lower-bound estimates of species richness (Anne Chao, personal communication). On the other hand, restricting datasets to ecologically more homogenous subsets of samples sometimes does produce well-behaved, asymptotic richness estimates (J. Longino *et al.*, in press). This is still an ongoing area of research, and there is much need for comparative studies of the performance of asymptotic species estimators on different empirical and theoretically derived data sets.

CONCLUSIONS

The principles of species accumulation, rarefaction, species richness, and species density have been established for many decades. However, ecologists have only recently begun in earnest to incorporate these concepts into their measurements of species diversity patterns and evaluation of theory in community ecology and biogeography. These tasks are especially important as ecologists attempt to inventory species-rich communities and document the loss of species diversity from habitat destruction and global climate change. Ecologists may have avoided individual-based and sample-based rarefaction curves because they are computationally intensive, but public-domain software is now available for these calculations (Colwell 2000a; Gotelli & Entsminger 2001).

ACKNOWLEDGEMENTS

REFERENCES

Abele, L.G. & Walters, K. (1979). The stability-time hypothesis: Reevaluation of the data. *Am. Naturalist*, 114, 559–568.

Abrams, P. (1995). Monotonic or unimodal diversity-productivity gradients: what does competition theory predict? *Ecology*, 76, 2019–2027.

Anderson, R.S. & Ashe, J.S. (2000). Leaf litter inhabiting beetles as surrogates for establishing priorities for conservation of selected tropical montane cloud forests in Honduras, Central America (Coleoptera; Staphilidae, Curculionidae). *Biodiversity Conservation*, 9, 617–653.

Ashton, P.S. (1998). Niche specificity among tropical trees: a question of scales. In: *Dynamics of Tropical Communities*, eds Newbery D.M., Brown N. & Prins H.H.T. BES Symposium, Vol. 37, pp. 491–514. Blackwell Science, Oxford, U.K.

Boulinier, T., Nichols, J.D., Sauer, J.R., Hines, J.E. & Pollock, K.H. (1998). Estimating species richness: the importance of heterogeneity in species detectability. *Ecology*, 79, 1018–1028.

Brewer, A. & Williamson, M. (1994). A new relationship for rarefaction. *Biodiversity Conservation*, 3, 373–379.

Bunge, J. & Fitzpatrick, M. (1993). Estimating the number of species; a review. *J. Am. Statist. Assoc.*, 88, 364–373.

Butler, B.J. & Chazdon, R.L. (1998). Species richness, spatial variation, and abundance of the soil seed bank of a secondary tropical rain forest. *Biotropica*, 30, 214–222.

Cannon, C.H., Peart, D.R. & Leighton, M. (1998). Tree species diversity of commercially logged Bornean rainforest. *Science*, 281, 1366–1368.

Chazdon, R.L., Colwell, R.K. & Denslow, J.S. (1999). Tropical tree richness and resource-based niches. *Science*, 285, 1459a http://www.sciencemag.org/cgi/content/full/285/5433/1459a

Chazdon, R.L., Colwell, R.K., Denslow, J.S. & Guariguata, M.R. (1998). Statistical methods for estimating species richness of woody regeneration in primary and secondary rain forests of NE Costa Rica. In: *Forest Biodiversity Research, Monitoring and Modeling: Conceptual Background and Old World Case Studies*, eds Dallmeier F. & Comiskey J.A.), pp. 285–309. Parthenon Publishing, Paris, France.

Clench, H. (1979). How to make regional lists of butterflies: Some thoughts. *Journal Lepidopterist's Society*, 33, 216–231.

Coddington, J.A., Griswold, C.E., Silva Dávila, D., Peñaranda, E. & Larcher, S.F. (1991). Designing and testing sampling protocols to estimate biodiversity in tropical ecosystems. In: *The Unity of Evolutionary Biology: Proceedings of the Fourth International Congress of Systematic and Evolutionary Biology*, ed. Dudley E.C., pp. 44–60. Dioscorides Press, Portland, Oregon, U.S.A.

Coddington, J.A., Young, L.H. & Coyle, F.A. (1996). Estimating spider species richness in a southern Appalachian cove hardwood forest. *J. Arachnol.*, 24, 111–128.

Coleman, B.D. (1981). On random placement and species–area relations. *Mathemat Biosciences*, 54, 191–215.

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

Colwell, R.K. (2000a). *EstimateS: Statistical Estimation of Species Richness and Shared Species from Samples* (Software and User's Guide), Version 6. http://viceroy.eeb.uconn.edu/estimates

Colwell, R.K. (2000b). Rensch's Rule crosses the line: Convergent allometry of sexual size dimorphism in hummingbirds and flower mites. *Am. Naturalist*, 156, 495–510.

Colwell, R.K. & Coddington, J.A. (1994). Estimating terrestrial biodiversity through extrapolation. *Phil. Trans. Royal Soc. London B*, 345, 101–118.

Colwell, R.K. & Winkler, D.W. (1984). A null model for null models in biogeography. In: *Ecological Communities: Conceptual Issues and the Evidence*, eds Strong D.R. Jr, Simberloff D., Abele L.G. & Thistle A.B.), pp. 344–359. Princeton University Press, Princeton, U.S.A.

Connell, J.H. (1978). Diversity in tropical rain forests and coral reefs. *Science*, 199, 1302–1303.

Cook, R.E. (1969). Variation in species density of North American birds. *Syst. Zool.*, 18, 63–84.

Cornell, H.V. (1999). Unsaturation and regional influences on species richness in ecological communities: a review of the evidence. *Ecoscience*, 6, 303–315.

Darwin, C. (1859). *The Origin of Species by Means of Natural Selection.* Murray, London, U.K.

Denslow, J. (1995). Disturbance and diversity in tropical rain forests: the density effect. *Ecological Applications*, 5, 962–968.

DiTomasso, A. & Aarssen, L.W. (1989). Resource manipulations in natural vegetation: a review. *Vegetatio*, 84, 9–29.

Elton, C. (1946). Competition and the structure of ecological communities. *J. Anim. Ecol.*, 15, 54–68.

Engstrom, R.T. & James, F.C. (1981). Plot size as a factor in winter bird-population studies. *Condor*, 83, 34–41.

Fisher, B.L. (1999). Improving inventory efficiency: a case study of leaf-litter ant diversity in Madagascar. *Ecol. Applications*, 9, 714–731.

Gaston, K.J. (2000). Global patterns in biodiversity. *Nature*, 405, 220–227.

Gotelli, N.J. (2001). *A Primer of Ecology*, 3rd edn. Sinauer Associates, Inc, Sunderland, MA, U.S.A.

Gotelli, N.J. & Arnett, A.E. (2000). Biogeographic effects of red fire ant invasion. *Ecol. Lett*, 3, 257–261.

Gotelli, N.J. & Entsminger, G.L. (2001). *Ecosim: Null Models Software for Ecology*, Version 6.0. Acquired Intelligence Inc, & Kesey-Bear http://homepages.together.net/gentsmin/ecosim.htm

Gotelli, N.J. & Graves, G.R. (1996). *Null Models in Ecology.* Smithsonian Institution Press, Washington, DC, U.S.A.

Grassle, J.F. & Maciolek, N.J. (1992). Deep-sea species richness: regional and local diversity estimates from quantitative bottom samples. *Am. Naturalist*, 139, 313–341.

Heck, K.L. Jr, van Belle, G. & Simberloff, D. (1975). Explicit calculation of the rarefaction diversity measurement and the determination of sufficient sample size. *Ecology*, 56, 1459–1461.

Howard, R.A. & Moore, A. (1984). *A Complete Checklist of Birds of the World.* Macmillan, London, U.K.

Hubbell, S.P. (1999). Tropical tree richness and resource-based niches. *Science*, 285, 1459a http://www.sciencemag.org/cgi/content/full/285/5433/1459a

Hubbell, S.P., Foster, R.P., O'Brien, S.T., Harrms, K.E., Condit, R., Wechsler, B., de Wright, S.J. & Lao, S.L. (1999). Light-gap disturbances, recruitment limitation, and tree diversity in a neotropical forest. *Science*, 283, 554–557.

Hurlbert, S.H. (1971). The nonconcept of species diversity: a critique and alternative parameters. *Ecology*, 52, 577–585.

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

Huston, M.A. & DeAngelis, D.L. (1994). Competition and coexistence: the effects of resource transport and supply rates. *Am. Naturalist*, 144, 954–977.

Irwin, T.L. (1997). Biodiversity at its utmost: tropical forest beetles. In: *Biodiversity II*, eds Reaka-Kudla M.L., Wilson D.E. & Wilson E.O., pp. 27–40. Joseph Henry Press, Washington, DC, U.S.A.

James, F.C. & Wamer, N.O. (1982). Relationships between temperate forest bird communities and vegetation structure. *Ecology*, 63, 159–171.

Järvinen, O. (1982). Species-to-genus ratios in biogeography: a historical note. *J. Biogeogr.*, 9, 363–370.

Keating, K.A. & Quinn, J.F. (1998). Estimating species richness: the Michaelis–Menton model revisited. *Oikos*, 81, 411–416.

Kobayashi, S. (1982). The rarefaction diversity measurement and the spatial distribution of individuals. *Jap. J. Ecol.*, 32, 255–258.

Kobayashi, S. (1983). Another calculation for the rarefaction diversity measurement for different spatial distributions. *Jap. J. Ecol.*, 33, 101–102.

Kobe, R.K. (1999). Tropical tree richness and resource-based niches. *Science*, 285, 1459a http://www.sciencemag.org/cgi/content/full/285/5433/1459a

Lake, P.S. (1990). Disturbing hard and soft bottom communities: a comparison of marine and freshwater environments. *Aust. J. Ecol.*, 15, 477–488.

Lamas, G., Robbins, R.K. & Harvey, D.J. (1991). A preliminary survey of the butterfly fauna of Pakitza, Parque Nacional del Manu, Peru, with an estimate of its species richness. *Publicaciones del Museo Historia Natural (Universidad San Marcos, Peru)*, 40, 1–19.

Lawton, J.H., Bignell, D.E. & Bolton, B. (1998). Biodiversity inventories, indicator taxa, and effects of habitat modification in tropical forest. *Nature*, 391, 72–76.

Longino, J. & Colwell, R.K. (1997). Biodiversity assessment using structured inventory: capturing the ant fauna of a tropical rain forest. *Ecol. Applications*, 7, 1263–1277.

Longino, J., Colwell, R.K. & Coddington, J.A. (in press). The ant fauna of a tropical rainforest: estimating species richness three different ways. *Ecology*.

MacArthur, R.H. & Wilson, E.O. (1967). *The Theory of Island Biogeography.* Princeton University Press, Princeton, U.S.A.

Magurran, A.E. (1988). *Ecological Diversity and its Measurement.* Princeton University Press, Princeton, U.S.A.

Maillefer, A. (1929). Le Coefficient generique de P. Jacard et sa signification. *Mem. Soc. Vaudoise Sc. Nat.*, 3, 113–183.

Mawdsley, N. (1996). The theory and practice of estimating regional species richness from local samples. In: *Tropical Rainforest Research – Current Issues: Proceedings of the Conference Held in Bandar Seri Gegawan, April (1993)*, eds Edwards D.S., Booth W.E. & Choy S.C., Kluwer Academic Publishers, Dordrecht, The Netherlands.

May, R.M. (1988). How many species on earth? *Science*, 241, 1441–1449.

McCabe, D.J. & Gotelli, N.J. (2000). Effects of disturbance frequency, intensity, and area on assemblages of stream invertebrates. *Oecologia*, 124, 270–279.

Melhop, P. & Lynch, J.F. (1986). Bird/habitat relationships along a successional gradient in the Maryland coastal plain. *Am. Midland Naturalist*, 116, 225–239.

Miller, T.E. (1994). Direct and indirect interactions in an early old-field plant community. *Am. Naturalist*, 143, 1007–1025.

Nichols, J.D. & Conroy, M.J. (1996). Estimation of species richness. In: *Measuring and Monitoring Biological Diversity. Standard Methods for Mammals*, eds Wilson D.E., Cole F.R., Nichols J.D.,

Rudran R. & Foster, M., pp. 226–234. Smithsonian Institution Press, Washington, DC, U.S.A.

Novotny, V. & Basset, Y. (2000). Rare species in communities of tropical insect herbivores: pondering the mystery of singletons. *Oikos*, 89, 564–572.

Oksanen, J. (1996). Is the humped relationship between species richness and biomass an artifact due to plot size? *J. Ecol.*, 84, 293–295.

Palmer, M.W. (1990). The estimation of species richness by extrapolation. *Ecology*, 71, 1195–1198.

Pielou, E.C. (1966). The measurement of diversity in different types of biological collection. *J. Theoret. Biol.*, 13, 131–144.

Pielou, E.C. (1975). *Ecological Diversity*. Wiley Interscience, New York, NY, U.S.A.

Poulsen, B.O., Krabbe, N., Frølander, A., Hinojosa, M.B. & Quiroga, C.O. (1997). A rapid assessment of Bolivian and Ecuadorian montane avifaunas using 20-species lists: efficiency, biases and data gathered. *Bird Conservation Int.*, 7, 53–67.

Purvis, A. & Hector, A. (2000). Getting the measure of biodiversity. *Nature*, 405, 212–219.

Raup, D.M. (1979). Size of the Permo-Triassic bottleneck and its evolutionary implications. *Science*, 206, 217–218.

Robbins, C.S., Sauer, J.R., Greenberg, R.S. & Droege, S. (1989). Population declines in North American birds that migrate to the neotropics. *Proc. Natl Acad. Sci. (USA)*, 86, 7658–7662.

Rosenzweig, M.L. & Abramsky, Z. (1993). How are diversity and productivity related?. In: *Species Diversity in Ecological Communities: Historical and Geographical Perspectives*, eds Ricklefs R.E. & Schluter D., pp. 52–65. University of Chicago Press, Chicago, IL, U.S.A.

Sanders, H. (1968). Marine benthic diversity: a comparative study. *Am. Naturalist*, 102, 243–282.

Scheiner, S.M., Cox, S.B., Willig, M., Mittelbach, G.G., Osenberg, C. & Kaspari, M. (2000). Species richness, species-area curves, and Simpson's paradox. *Evol Ecol. Res.*, 2, 791–802.

Shipley, B. (1993). A null model for competitive hierarchies in competition matrices. *Ecology*, 74, 1693–1699.

Silva, D. & Coddington, J.A. (1996). Spiders of Pakitza (Madre de Diós, Perú): species richness and notes on community structure. In: *Manu: the Biodiversity of Southeastern Perú*, eds Wilson D.E. & Sandoval A., pp. 253–311. Smithsonian Institution Press, Washington, DC, U.S.A.

Simberloff, D. (1970). Taxonomic diversity of island biotas. *Evolution*, 24, 23–47.

Simberloff, D. (1972). Properties of the rarefaction diversity measurement. *Am. Naturalist*, 106, 414–418.

Simberloff, D. (1978). Use of rarefaction and related methods in ecology. In: *Biological Data in Water Pollution Assessment: Quantitative and Statistical Analyses*, eds Dickson K.L., Cairns J. Jr & Livingston R.J., pp. 150–165. American Society for Testing and Materials, Philadelphia, PA, U.S.A.

Simberloff, D. (1986). Are we on the verge of a mass extinction in tropical rain forests? In: *Dynamics of Extinction*, ed. Elliott D.K., pp. 165–180. John Wiley & Sons, New York.

Simpson, G.G. (1964). Species density of North American recent mammals. *Syst. Zool.*, 13, 57–73.

Smith, E.P., Stewart, P.M. & Cairns, J. Jr (1985). Similarities between rarefaction methods. *Hydrobiologia*, 120, 167–169.

Soberón, J. & Llorente, J. (1993). The use of species accumulation functions for the prediction of species richness. *Conservation Biol.*, 7, 480–488.

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

Stevens, M.H.H. & Carson, W.P. (1999). Plant density determines species richness along an experimental fertility gradient. *Ecology*, 80, 455–465.

Stork, N.E. (1991). The composition of the arthropod fauna of Bornean lowland rain forest trees. *J. Trop. Ecol.*, 7, 161–180.

Tilman, D. (1982). *Resource Competition and Community Structure*. Princeton University Press, Princeton, NJ, U.S.A.

Tilman, D. (1988). *Plant Strategies and the Dynamics and Structure of Plant Communities*. Princeton University Press, Princeton, NJ, U.S.A.

Tipper, J.C. (1979). Rarefaction and rarefiction – the use and abuse of a method in paleoecology. *Paleobiology*, 5, 423–434.

Vandermeer, J., Cerda, I.G.d.l., Boucher, D., Perfecto, I. & Ruiz J. (2000). Hurricane disturbance and tropical tree species diversity. *Science*, 290, 788–791.

Vinson, M.R. & Hawkins, C.P. (1998). Biodiversity of stream insects: variation at local, basin, and regional scales. *Annu. Rev. Entomol.*, 43, 271–293.

Walther, B.A., Cotgreave, P., Price, R.D., Gregory, R.D. & Clayton, D.H. (1995). Sampling effort and parasite species richness. *Parasitol. Today*, 11, 30–310.

Washington, H.G. (1984). Diversity, biotic and similarity indices. *Water Res.*, 18, 653–694.

Williams, C.B. (1947). The generic relations of species in small ecological communities. *J. Anim. Ecol.*, 16, 11–18.

Williams, C.B. (1964). *Patterns in the Balance of Nature*. Academic Press, New York, NY, U.S.A.

Wolda, H., O'Brien, C.W. & Stockwell, H.P. (1998). Weevil diversity and seasonality in tropical Panama as deduced from light-trap catches (Coleoptera: Curculionoidea). *Smithson. Contr. Zool.*, 590, 1–79.

BIOSKETCH

Nicholas J. Gotelli is a population and community ecologist with interests in null models, biogeography, community assembly, metapopulation dynamics, and demography.