Evolutionary Ecology Research, 2006, **8**: 195–211

Effects of taxon abundance distributions on expected numbers of sampled taxa

Matthew A. Kosnik[1]* and Peter J. Wagner[2]

[1]*Department of Geophysical Sciences, University of Chicago, Chicago, IL 60637 and*
[2]*Department of Geology, Field Museum of Natural History, Chicago, IL 60605, USA*

ABSTRACT

Question: Holding both the true number of taxa and their evenness constant, what is the effect of the shape of the abundance distribution on the number of sampled taxa?

Method: We examine the effects of three types of abundance distribution (geometric, log-normal and Zipf) on the expected number of sampled taxa using Hurlbert's equation (1971, equation 14). First, we examine the differences in the number of sampled taxa for the three distributions given the same true number of taxa and true evenness. Second, we determine the sample sizes needed to find more taxa from a taxon-rich, low-evenness collection than found in a taxon-poor, high-evenness collection with the same model distribution.

Conclusions: Independently of the true number of taxa and evenness, the shape of the abundance distribution affects the number of taxa expected in a sample. Given moderate to large sample sizes, a Zipf distribution will yield the most taxa, whereas a geometric distribution will yield the fewest. When comparing collections with the same model distributions, it takes the smallest sample sizes to recognize that a taxon-rich, low-evenness Zipf distribution has more taxa than does a taxon-poor, high-evenness Zipf distribution. It requires the largest sample sizes to do this when both are geometric distributions. A necessary implication of these results is that no simple evenness metric can predict the same number of sampled taxa given the same true number of taxa, true evenness and sample size but different model distributions.

Keywords: abundance distributions, evenness, geometric, log-normal, Pielou's *J*, sampled diversity, sampled richness, Zipf.

INTRODUCTION

Hypotheses about diversity patterns, extinction and origination dynamics, and ecological complexity frequently require ecologists and palaeontologists to contrast the number of taxa in different locations or time intervals. It has long been known that the number of observed taxa depends on sample size as well as the true number of taxa present (e.g. Fisher *et al.*, 1943; Sanders, 1968; Raup, 1975). In addition, variation in the relative abundance of taxa affects the

---

* Address all correspondence to M.A. Kosnik, Centre for Coral Reef Biodiversity, School of Marine Biology and Aquaculture, James Cook University, Townsville, QLD 4811, Australia. e-mail: mkosnik@alumni.uchicago.edu
Consult the copyright statement on the inside front cover for non-commercial copying policies.

number of sampled taxa (Hurlbert, 1971; Simberloff, 1978; Tipper, 1979; Gotelli and Colwell, 2001). Thus, whether one is sampling in the field or resampling with a computer, the underlying taxon abundance distribution affects the number of taxa sampled in the resulting collection. If we take equal sample sizes from two assemblages with the same true number of taxa but different relative abundance distributions, then we expect the assemblage with more nearly equal taxon abundances to yield collections containing more taxa. Simulations have shown that evenness exerts a much stronger influence on the number of sampled taxa than does the true number of taxa present at sample sizes typical of palaeontological studies (Wagner, 2000; Peters, 2001).

Ecology has a strong tradition of discussing the effects of sampling and evenness on the expected number of sampled taxa in a collection of a given size (e.g. Fisher *et al.*, 1943; Hurlbert, 1971; Hill, 1973). We build on this literature by examining how different model taxon abundance distributions affect the number of sampled taxa independently of evenness. This is potentially important because numerous possible taxon abundance distributions (TADs) exist (see Tipper, 1980; Hayek and Buzas, 1997; Etter, 1999). In this paper, we address two specific questions:

1. How do three general TAD models (geometric, log-normal and Zipf) affect the expected number of sampled taxa when the TADs have the same true evenness and number of taxa?
2. Do particular differences in evenness and the number of taxa have the same implications for our ability to determine which sample contains more taxa when we compare two log-normal distributions or two geometric distributions or two Zipf distributions?

We demonstrate that evenness is only a partial measure of the effect of TADs on the number of sampled taxa, and that the implications of particular differences in evenness and true numbers of taxa differ among model TADs.

METHODS

Biodiversity terms are used inconsistently across disciplines. Palaeontologists (e.g. Raup, 1975; Sepkoski, 1978) and conservation biologists (e.g. Williams and Gaston, 1994) typically define diversity as the number of taxa. However, many ecologists use richness for the number of taxa and define diversity as a compound of richness and evenness of those taxa (e.g. Hurlbert, 1971; Hayek and Buzas, 1997). To avoid potential confusion, we use the unambiguous phrase ‘number of taxa’.

Model taxon abundance distributions

We examine three model taxon abundance distributions (TADs):

- geometric (Motomura, 1932)
- log-normal (Preston, 1948)
- Zipf (i.e. log–log) (Zipf, 1949)

A semi-logarithmic plot (Fig. 1) illustrates the general differences between these models given distributions with the same number of taxa ($n = 300$) and evenness (Pielou’s $J = 0.775$; see ‘Evenness metrics’ below).

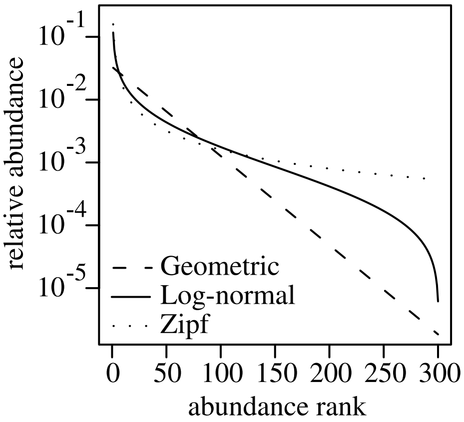

Fig. 1. Log-linear plots illustrating rank order abundances under geometric ($d = 1.033$), log-normal ($m = 5.358$) and Zipf distributions ($d = 0.999$). All three distributions have the same evenness ($J = 0.775$) and the same number of taxa ($S = 300$).

Geometric

The geometric distribution represents a simple decay where each taxon rank $i + 1$ has a constant fraction of the number of individuals that taxon rank i has. This yields a simple exponential decay of abundances that is linear on a log-linear plot (Fig. 1). The expected frequency of taxon rank i (f_i) under a geometric distribution is given by:

$$f_i = \frac{d^i}{\sum_{j=1}^{S} d^j} \tag{1}$$

where d is simply an exponential decay slope (i.e. f_{i+1}/f_i) and S is the number of taxa (see May, 1975).

Theoretically, a geometric series can contain an infinite number of members, with the most abundant member having $f_1 = (1 - d)$. However, there are a finite number of organisms and biological organisms are integer units, so an infinite series is not biologically possible. The number of taxa has much less influence on geometric distributions than on either log-normal or Zipf distributions as f_i rapidly approaches d even at low true S.

The geometric distribution is nearly identical to the log-series distribution (Fisher *et al.*, 1943), although the log-series distribution can curve slightly on log-linear plots. Buzas and Hayek (2005) suggest that log-series distributions are common among foraminifera assemblages. We use the geometric distribution instead simply for ease of computation.

Log-normal

Log-normal distributions are not described by simple decays. Typically, workers illustrate a log-normal distribution as an octave plot – that is, a histogram shaped like a normal curve, with the height of each bar giving the number of taxa with abundances in a single order of magnitude (Fig. 2). No simple formula yields a log-normal that can be plotted on a log-linear plot. However, we partition normal curves into $S + 1$ units having equal area under that curve. The position of each partition i on the x-axis then gives the logarithm of

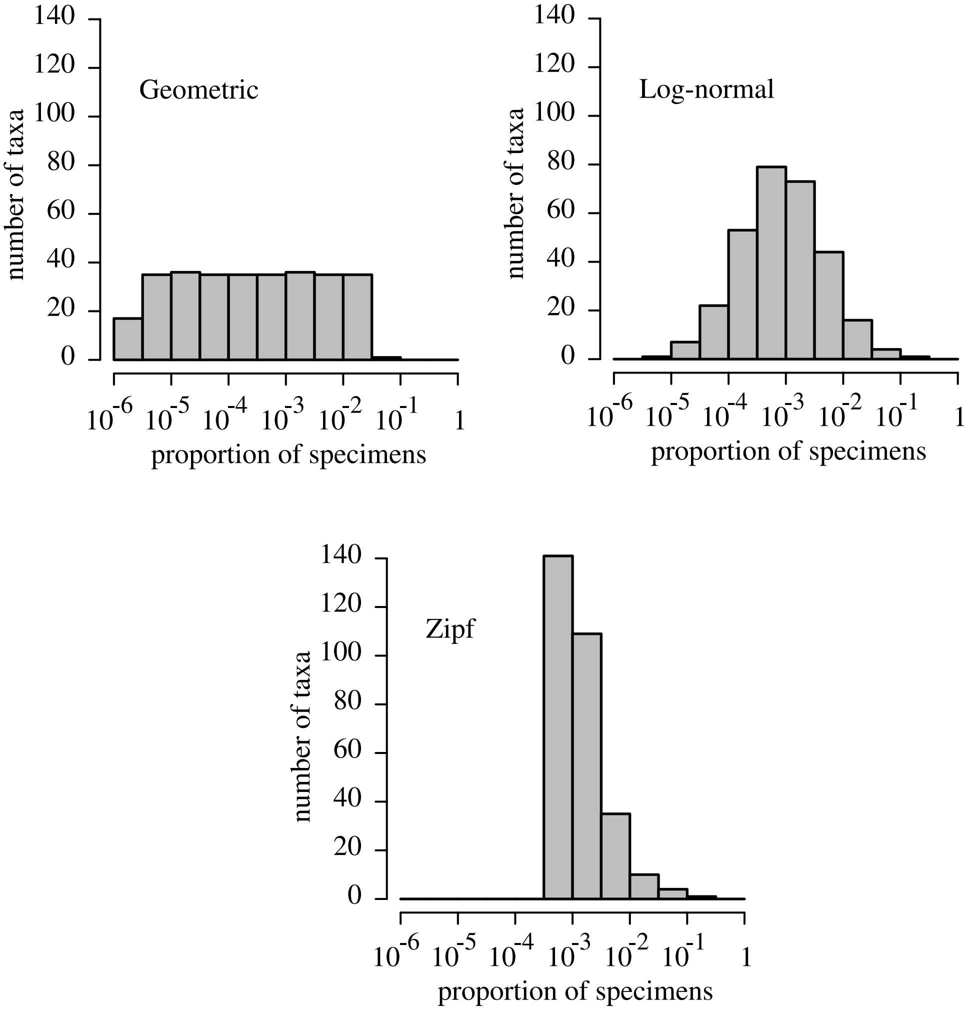

Fig. 2. Octave plots of the geometric, log-normal and Zipf distributions from Fig. 1. The geometric distribution has a nearly uniform number of taxa in each octave. The log-normal distribution has a large number of taxa in the modal octave, with fewer exceptionally rare and exceptionally abundant taxa than the geometric distribution. The Zipf distribution is strongly skewed, with no exceptionally rare and fewer common taxa present.

the relative abundance of each taxon *i*. The resulting TAD yields a curve with an inflection point when plotted on a log-linear plot (Fig. 1). Workers frequently have suggested that empirical data fit log-normal distributions (e.g. Preston, 1948).

Zipf

The Zipf distribution (Zipf, 1949) represents a decay where abundance decreases exponentially as taxon rank increases exponentially. This yields a distribution with a very abundant first taxon followed by a rapid exponential decline in the abundance of subsequent taxa (Fig.1). The expected frequency of taxon rank *i* (f_i) under a Zipf distribution is:

$$f_i = \frac{i^{-\gamma}}{\sum_{j=1}^{S} j^{-\gamma}} \tag{2}$$

(Mouillot and Leprêtre, 1999; see May, 1975). S is again the number of taxa. The slope of the TAD is determined by parameter γ, and reflects the average probability of sampling a taxon (Frontier, 1985). Like the geometric distribution, a theoretical Zipf distribution can encompass an infinite series, but the number of taxa has a more drastic effect on a Zipf f_i than it does on a geometric f_i.

Other abundance distributions

We examine these three distributions for two reasons. First, empirical studies suggest that each of these distributions describes empirical assemblages in some examples. Second, these three distributions illustrate the range of distribution shapes that workers have proposed, and the different shapes of the three TADs provide different proportions of 'common' and 'rare' taxa at a given evenness. Re-plotting the distributions illustrated in Fig. 1 as octave plots (i.e. numbers of taxa with abundances within a certain order of magnitude) helps illustrate these differences (Fig. 2).

Many other model TADs exist. For example, log-normal distributions can be truncated (Slocomb *et al.*, 1977), where the modal (peak) octave is closer to the rarest or most common octave than it is to the most common average abundance octave. If the mode is closest to the rarest taxa, then the truncated log-normal distribution will resemble a Zipf distribution. Similarly, the Zipf-Mandelbrot distribution modifies the Zipf with a second parameter (b) that creates a sigmoidal distribution on a log-linear plot, one that is similar to the log-normal. The negative binomial distribution (Gleason, 1922) tends to behave similarly to a log-normal distribution. The broken-stick distribution (MacArthur, 1957) also resembles the geometric distribution. Finally, the zero-sum multinomial (Hubbell, 2001) typically yields a distribution between the log-normal and the geometric distributions. Because these other distributions tend to have shapes like one or another of the three distributions we examine, we do not illustrate the effects on the expected number of sampled taxa for all of these additional model distributions.

Finally, workers have advanced many theoretical reasons for expecting many of these distributions. However, our sole concern here is the effect of different model distributions on the number of sampled taxa. Many studies have shown that each of these three models describes different empirical collections well, which we view as ample reason to be concerned with the potential effects of different TADs on the number of sampled taxa.

Evenness metrics

There is no universally accepted evenness metric, although many evenness metrics exist (e.g. Pielou, 1966; Buzas and Gibson, 1969; Hurlbert, 1971; Smith and Wilson, 1996; Powell and Kowalewski, 2002; Peters, 2004). However, different evenness metrics are strongly correlated when applied to the same data sets (Powell and Kowalewski, 2002; M.A. Kosnik and P.J. Wagner, unpublished analyses). Some evenness metrics are independent of richness (e.g. Smith and Wilson, 1996) or are little affected by sampling intensity among collections (e.g. Peters 2004). However, the conclusions presented here are the same regardless of which evenness metric we used.

We will present results using Pielou's (1966) *J*, but in the Appendix we show the results using E′ (Camargo, 1993) and E_{var} (Smith and Wilson, 1996). We use Pielou's *J* because it is one of the most commonly used evenness metrics in palaeoecological studies (e.g. Levinton and Bambach, 1975; Pryor and Gastaldo, 2000; Edinger *et al.*, 2001; Wing and Harrington, 2001; Jaramillo, 2002; Powell and Kowalewski, 2002). Pielou's *J* is calculated as:

$$J = \frac{\sum_{i=1}^{S} f_i \ln f_i}{\ln S} \tag{3}$$

where *S* is the number of taxa and f_i is the frequency of taxon *i*. The numerator here is the Shannon-Weiner index (e.g. Simpson, 1960), and is sometimes labelled 'Shannon's Evenness' (e.g. Gardiner, 2001). This index often is used to measure heterogeneity without describing a model distribution, which means that evenness here is simply the logarithmic mean of heterogeneity. *J* ranges from zero to one, with one representing perfectly even abundance distributions (i.e. 1/*S* for all taxa).

The results presented here are based on the true evenness of the TAD. Empirical measures of evenness are biased estimates of true evenness because empirical measures of evenness are calculated using the number of sampled taxa, and the number of sampled taxa is affected by sample size (Peters, 2004; but see Smith and Wilson, 1996). Moreover, rare taxa are more likely to be missed, so empirical measures are biased towards overestimating evenness. However, metric 'distortions' due to finite sampling are much less than the ranges of evenness values that are reported in the literature (e.g. Powell and Kowalewski, 2002). Thus, our exercise should offer adequate comparisons to empirical data sets.

Calculating the expected number of sampled taxa

Hurlbert (1971, equation 14) gives the expected number of sampled taxa for a collection of *N* individuals as:

$$E(S'_N) = \sum_{i=1}^{S} \left[1 - (1 - f_i)^N \right] \tag{4}$$

where f_i is the probability of sampling taxon *i*. In an unbiased sample, the sampling probability (f_i) is the same as the occurrence frequency. In collections where the sampling probability is not uniform for all individuals, the sampling frequency (f_i) is the occurrence frequency weighted by the sampling probability (i.e. preservation potential in palaeontological studies or collection probability in ecological studies). This equation assumes that the sample is well mixed (spatially homogeneous over the scale of sampling) and that all individuals have the same sampling probability. Although these assumptions are likely to be violated in real data, they will serve as the simplest case. Moreover, violations of these assumptions do not appear to favour any particular model used here. Because $(1 - f_i)^N$ is the probability of not sampling taxon *i* given *N* individuals, equation (4) simply sums the probabilities of sampling each taxon at least once.

We use equation (4) to answer two distinct questions that arise when comparing assemblages:

1. Given two equal-sized samples from different TADs with the same true evenness and true number of taxa, should we expect the same number of taxa in the samples?
2. When one compares two log-normal distributions or two geometric distributions or two Zipf distributions, should we expect similar differences in evenness and the number of taxa to have the same effects on the number of taxa in the sample?

To address the first question, we calculate the expected number of taxa sampled using equation (4) for the three models of taxon abundance (TADs) at various sample sizes ($N = 1$ to 10,000), true numbers of taxa ($S_T = 50$, 100 and 200) and evenness values ($J = 0.90$, 0.75 and 0.60). The values for the number of taxa and the evenness are realistic for certain palaeontological studies (Powell and Kowalewski, 2002). We present our results as standard collection curves for each TAD and each of the nine combinations of true number of taxa and true evenness.

To address the second question, we determine the sample size at which the expected number of sampled taxa for a taxon-rich, low-evenness TAD exceeds that of a reference assemblage with the same model TAD but with fewer taxa ($S_T = 100$) and high evenness ($J = 0.95$). As a general summary, we present our results as contour plots for each TAD type. The plots show the sample size required to determine that a sample of known true number of taxa (y-axis) and known true evenness (x-axis) has more taxa than the reference sample ($S_T = 100$ and $J = 0.95$). This is equivalent to plotting the sample size at which Hurlbert's sampling curves intersect. This extends Simberloff's (1978, figure 1) finding that the sample size at which the expected number of taxa from an uneven, taxon-rich assemblage surpasses that of a taxon-poor, even assemblage increases as the difference in evenness increases.

RESULTS

Expected number of taxa in a sample: comparing across model TADs

Model TAD has a noticeable effect on the expected number of sampled taxa given identical true numbers of taxa (S_T) and identical true evenness (J). Figure 3 presents collection curves for three different values of true evenness (columns; $J = 0.90$, 0.75, 0.60) and three different values of the true number of taxa (rows; $S_T = 50$, 100, 200). The bottom row illustrates the TADs from the bottom row of collection curves ($S_T = 200$). (See the Appendix for discussion and analogous plots of E' and E_{var}.) Not surprisingly, the most nearly even TADs (Fig. 3, left column) show the smallest difference in the expected number of sampled taxa among TADs, whereas the least nearly even TADs (Fig. 3, right column) show the greatest difference in the expected number of sampled taxa among TADs. Differences in the expected number of sampled taxa among different model TADs with the same evenness increase as the true number of taxa increases (Fig. 3, bottom row versus top row). At low sample sizes ($N < 100$), sample size has the strongest effect on the number of sampled taxa (Fig. 3, all panels). The essentially complete sampling of an assemblage seen at sample sizes of about 10,000 individuals is going to be exceptionally rare in practice (Fig. 3, left column). Taxa from a Zipf distribution are more thoroughly sampled (Fig. 3, all panels) due to the large number of moderately rare taxa that all have a reasonable chance of being sampled (Figs. 1, 2 and Fig. 3, bottom row). Geometric TADs have slightly higher numbers of taxa sampled at very low sample sizes ($N < 100$), but they are slowest to accumulate taxa at larger

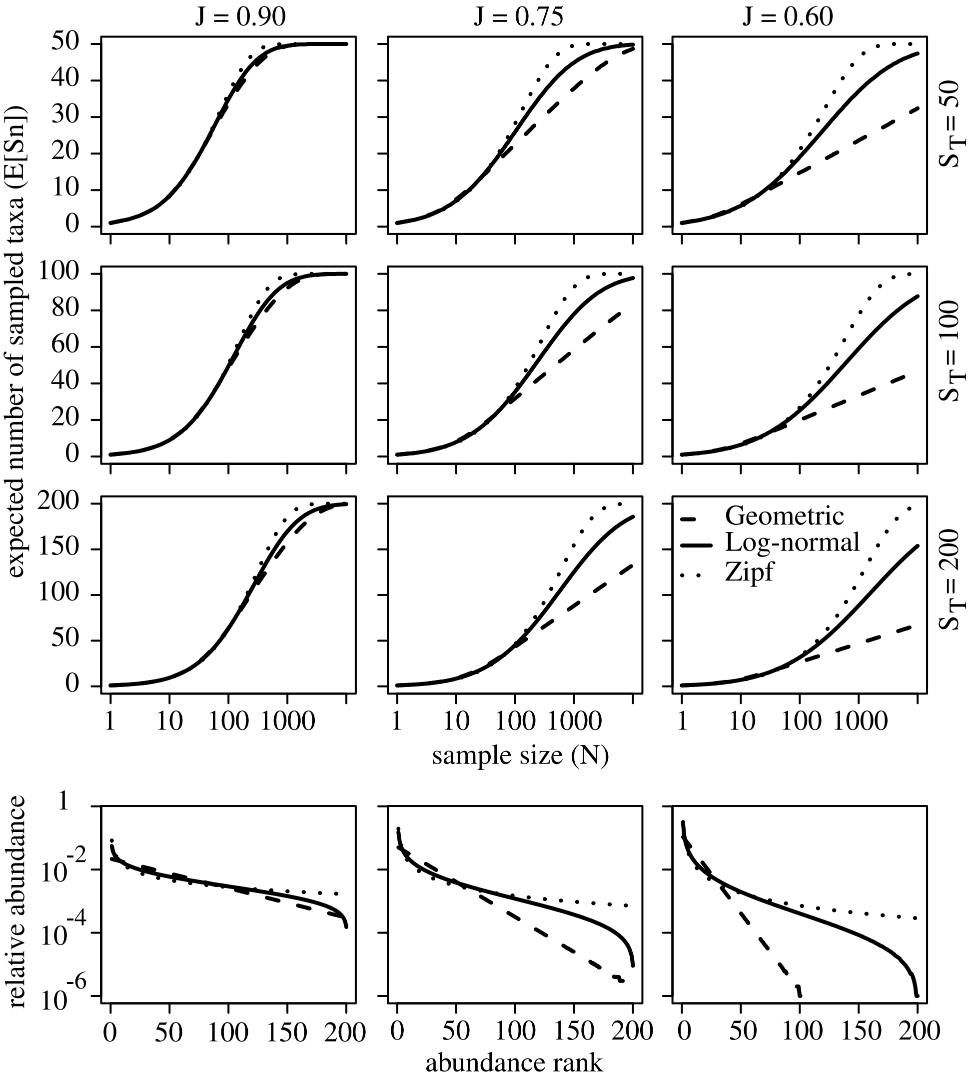

Fig. 3. Collection curves for geometric, log-normal and Zipf distributions and representative TADs. Columns with a given evenness (Pielou's *J*: 0.90, 0.75, 0.60) and rows of given true number of taxa (S_T: 50, 100, 200). The bottom row plots the TADs corresponding to the collection curve row immediately above it (Pielou's *J* as per the column and $S_T = 200$). Both the evenness and number of taxa have strong effects on the shape of the collection curves, but the effects are distribution specific. Zipf distributions are always highest and geometric distributions are always lowest at larger sample sizes.

sample sizes (Fig. 3, all panels); this is due to the large number of very rare taxa (Figs. 1, 2 and Fig. 3, bottom row). Log-normal TADs lie in the middle (Fig. 3, all panels), due to the large number of intermediate taxa but small number of very rare taxa (Figs. 1, 2 and Fig. 3, bottom row). Note that although the sequence of lines is the same for all parameters tested (Fig. 3, all panels), the magnitude of the differences in the expected number of sampled taxa among TADs is sensitive to both the number of taxa and to evenness.

Recognizing differences in the true number of taxa: comparing within-model TADs

When comparing collections with the same model TAD, geometric distributions require the highest sample sizes to determine that the richer sample is indeed richer. Many of the comparisons require thousands of individuals to correctly determine which assemblage contains more taxa (Fig. 4). The log-normal and Zipf distributions require similar sampling effort. However, most of the comparisons require sample sizes of more than 500 individuals to correctly determine that the richer/less even assemblage is indeed richer than the reference sample (Fig. 4). The Zipf distribution requires the lowest sample sizes to

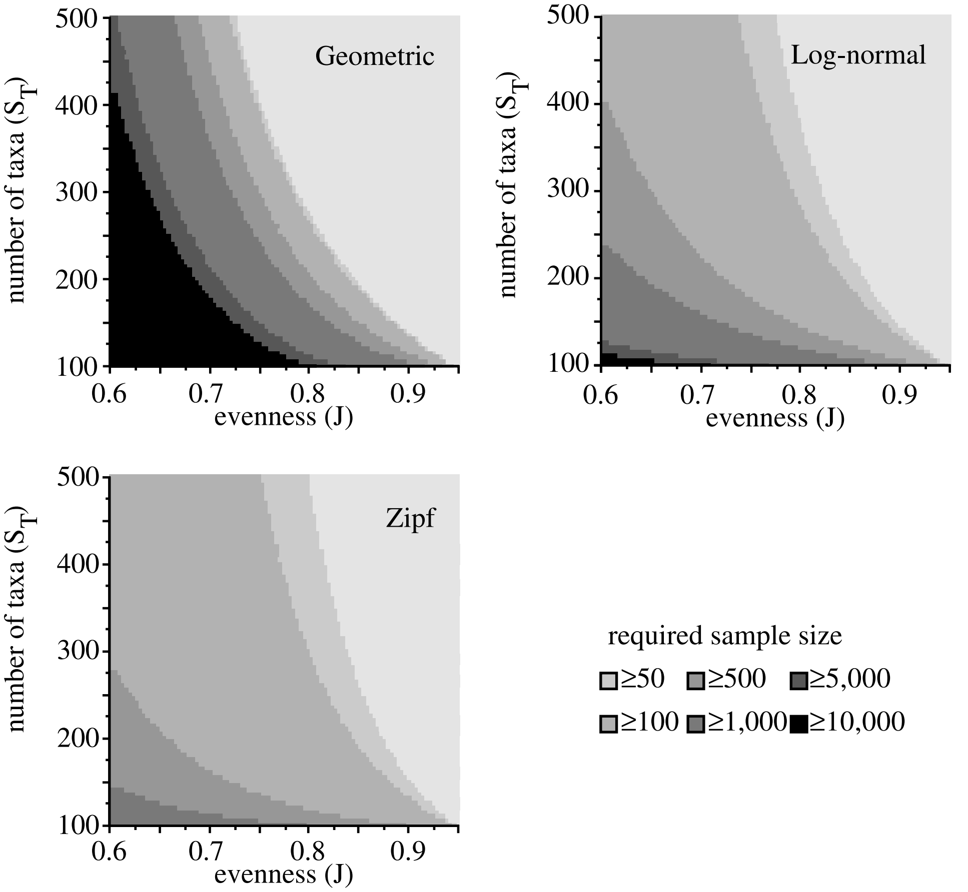

Fig. 4. Contour plots showing the sample sizes (N) necessary for the expected number of sampled taxa, $E(S'_N)$, of a distribution with a given number of taxa and evenness to match the $E(S'_N)$ from the same model distribution with 100 taxa and evenness of $J = 0.95$. In other words, this shows the sample sizes at which sampling curves for a taxon-rich, low-evenness distribution match those of a taxon-poor, high-evenness distribution. Geometric distributions of low evenness require a much greater sample size to yield the same $E(S'_N)$ as the 100-taxon more even, geometric distribution than log-normal or Zipf require to match the $E(S'_N)$ of 100-taxon more even, log-normal or Zipf distributions. Similarly, log-normal distributions of low evenness require greater sample sizes to yield the same $E(S'_N)$ as the 100-taxon more even, log-normal distributions.

determine that the richer sample is indeed richer (Fig. 4). When sampling Zipf or log-normal distributions of similar true numbers of taxa (i.e. $S_T = 150$) but different evenness (i.e. $J = 0.80$), sample sizes of 500 individuals will correctly rank the number of taxa in two collections, whereas correctly ordering geometric distributions with the same parameter values requires more than 1000 individuals.

Although the results illustrate comparisons of the same model distribution, these results will be very similar across TAD types given the similarity between highly even TADs of different TAD types (Fig. 3, left column, middle row). Thus comparable differences in evenness have very different effects on our ability to recognize the richer sample given different TADs.

DISCUSSION

Can we build a better evenness metric?

We show effects of model distributions on the expected number of sampled taxa that are largely independent of evenness. We used J, E', E_{var} and other evenness metrics. One might wonder whether an evenness metric might be devised such that given the same evenness, two collections of the same true number of taxa will have the same expected number of sampled taxa for all possible sample sizes regardless of TAD. However, we think that this is not possible.

Consider the three model TADs with $S_T = 300$ and $J = 0.775$ (Fig. 5). The geometric has 35 taxa with $f_i \geq 0.01$ (i.e. a relative abundance of more than 1 in 100 individuals), the log-normal has 21 taxa with $f_i \geq 0.01$ and the Zipf has 15 taxa with $f_i \geq 0.01$ (Fig. 5). Given that we expect to sample taxa with $f_i \geq 0.01$ at least once at a sample size (N) of 100, we expect to find over twice as many taxa at $N = 100$ from the geometric than from the Zipf and nearly half again as many taxa at $N = 100$ from the log-normal despite their having the same true number of taxa and true evenness (J).

The scenario changes markedly when we look at the number of taxa we expect to find at higher sampling levels from the same distributions (Fig. 5). The geometric has 109 taxa with $f_i \geq 0.001$, the log-normal has 140 taxa with $f_i \geq 0.001$, and the Zipf has 162 taxa with $f_i \geq 0.001$. Finally, all 300 taxa from the Zipf have $f_i \geq 0.0001$, whereas only 248 taxa from the log-normal and 181 taxa from the geometric have $f_i \geq 0.0001$.

The example above illustrates why there cannot be a single evenness metric that can predict the same number of sampled taxa at all sample sizes. Because of the fundamentally different shapes of the TAD, there will always be different numbers of taxa that we expect to have sampled from different model TADs. At most, there will be only two values of N at which Zipf and geometric distributions have the same number of sampled taxa and only three values of N at which log-normal distributions and either geometric or Zipf distributions will have the same number of sampled taxa. These results are consistent with other investigations of evenness metrics.

Implications for diversity studies

Workers often are interested in testing hypotheses about the relative numbers of taxa among collections (e.g. Miller and Foote, 1996) or in estimating the number of taxa, either for its own sake (e.g. Longino *et al.*, 2002) or to use such estimates as models for testing macroevolutionary

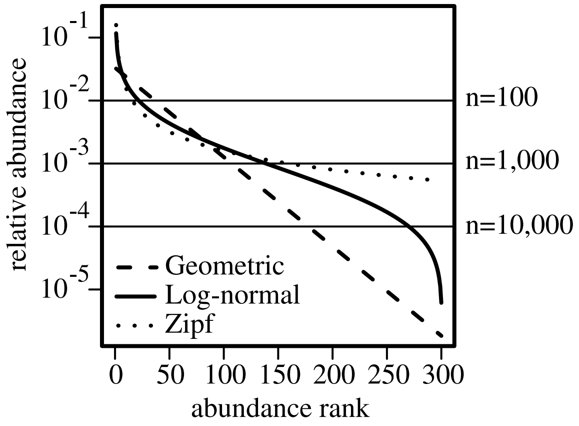

Fig. 5. Why the expected number of sampled taxa, $E(S'_N)$, from different model distributions will differ even when the true number of taxa and evenness are the same. The horizontal lines denote taxon relative abundance (frequencies) of 0.01, 0.001 and 0.0001. At any sample size N, we expect to sample at least one individual from any taxon representing $1/N^{th}$ of the population or more. Thus, at 100 individuals we expect to sample taxa above the highest horizontal line, whereas at 10,000 individuals we expect to sample taxa above the lowest horizontal line. The geometric distribution has both the greatest number of taxa above the highest horizontal line and the greatest number of taxa below the middle and lowest horizontal lines. Thus, it has the most taxa that we expect to sample at 100 individuals but the fewest that we expect to sample at 1000 or 10,000 individuals. Conversely, the Zipf has the fewest taxa above the highest line but the most above the lowest line. Because the model distributions have fundamentally different shapes, they will have the same number of taxa above $1/N$ for only two or three values of N. Thus, it is impossible for two different model distributions to have the same expected number of sampled taxa for all sample sizes. A corollary of this is that there can be no measure of evenness for which the expected number of sampled taxa will be the same for all sample sizes across different model distributions.

hypotheses about diversification (e.g. Purvis *et al.*, 1995). Because of the effects of TADs on the expected number of sampled taxa, simply standardizing sample sizes (e.g. Sanders, 1968) does not allow one to accurately estimate the relative number of taxa (Hurlbert, 1971). Moreover, because distribution types affect the number of sampled taxa independently of evenness, we cannot devise a simple correction based on measured evenness.

The effects of distributions that we illustrate here affect the number of sampled taxa in both primary field studies and analytical resampling studies. Also, Brose *et al.* (2003) showed that broken-stick and random-assortment distributions affected the number of taxa inferred by non-parametric estimators such as the Chao-2 metric (Chao, 1984). How other distributions affect such metrics certainly merits investigation. Ultimately, it is possible that we must explicitly fit model distributions when estimating the number of taxa present (e.g. Efron and Thisted, 1976).

CONCLUSIONS

Model distributions have a strong effect on the expected number of sampled taxa. At a given number of taxa and evenness and at a small sample size, we expect to sample more taxa from a geometric distribution than from a log-normal distribution or a Zipf distribution, and more taxa from a log-normal distribution than from a Zipf distribution. At a large sample size, however, we expect to sample more taxa from a Zipf distribution than from a log-normal distribution or geometric distribution, and more taxa from a log-normal distribution than from a geometric distribution. Thus, evenness metrics are inadequate and workers should account for differences in taxonomic abundance distributions when comparing the number of taxa between faunas. Future work should concentrate on incorporating information about distributions when testing hypotheses concerning differences in the number of taxa.

ACKNOWLEDGEMENTS

M.A.K. was funded by a NSF predoctoral fellowship. P.J.W.'s contributions were funded in part by NSF EAR-990328. Thanks to T. Felix, M. Foote, E.G. Hunt, D. Jablonski, A.P. Kerswell, S.M. Kidwell, R. Lockwood, R.A. Lupia, S.K. Lyons, T.E. Roberts and K.J. Wilson for comments on drafts of this manuscript. Critical comments by M.A. Buzas, S.R. Connolly, L.-A. Hayek and M.L. Rosenzweig on drafts of this paper helped us to clarify its focus. This is Paleobiology Database publication #37.

REFERENCES

Brose, U., Martinez, N.D. and Williams, R.J. 2003. Estimating species richness: sensitivity to sample coverage and insensitivity to spatial patterns. *Ecology*, **84**: 2364–2377.

Buzas, M.A. and Gibson, T.G. 1969. Species diversity: benthonic foraminifera in western North Atlantic. *Science*, **163**: 72–75.

Buzas, M.A. and Hayek, L.-A.C. 2005. On richness and evenness within and between communities. *Paleobiology*, **31**: 199–220.

Camargo, J.A. 1993. Must dominance increase with the number of subordinate species in competitive interactions? *J. Theor. Biol.*, **161**: 537–542.

Chao, A. 1984. Nonparametric-estimation of the number of classes in a population. *Scand. J. Stat.*, **11**: 265–270.

Edinger, E.N., Pandolfi, J.M. and Kelley, R.A. 2001. Community structure of Quaternary coral reefs compared with Recent life and death assemblages. *Paleobiology*, **27**: 669–694.

Efron, B. and Thisted, R. 1976. Estimating the number of unseen species: how many words did Shakespeare know? *Biometrika*, **63**: 435–447.

Etter, W. 1999. Community analysis. In *Numerical Palaeobiology* (D.A.T. Harper, ed.), pp. 285–360. Chichester: Wiley.

Fisher, R.A., Corbet, A.S. and Williams, C.B. 1943. The relation between the number of species and the number of individuals in a random sample of an animal population. *J. Anim. Ecol.*, **12**: 42–48.

Frontier, S. 1985. Diversity and structure in aquatic ecosystems. In *Oceanography and Marine Biology: An Annual Review* (M. Barnes, ed.), pp. 253–312. Aberdeen: Aberdeen University Press.

Gardiner, L. 2001. Stability of Late Pleistocene reef mollusks from San Salvador Island, Bahamas. *Palaios*, **16**: 372–386.

Gleason, H. 1922. On the relation between species and area. *Ecology*, **3**: 156–162.

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

Hayek, L.-A.C. and Buzas, M.A. 1997. *Surveying Natural Populations*. New York: Columbia University Press.

Hill, M.O. 1973. Diversity and evenness: a unifying notation and its consequences. *Ecology*, **54**: 427–431.

Hubbell, S.P. 2001. *The Unified Neutral Theory of Biodiversity and Biogeography*. Monographs in Population Biology #32. Princeton, NJ: Princeton University Press.

Hurlbert, S.H. 1971. The nonconcept of species diversity: a critique and alternative parameters. *Ecology*, **52**: 577–586.

Jaramillo, C.A. 2002. Response of tropical vegetation to Paleogene warming. *Paleobiology*, **28**: 222–243.

Levinton, J.S. and Bambach, R.K. 1975. A comparative study of Silurian and Recent deposit-feeding bivalve communities. *Paleobiology*, **1**: 97–124.

Longino, J.T., Coddington, J.A. and Colwell, R.K. 2002. The ant fauna of a tropical rain forest: estimating species richness three different ways. *Ecology*, **83**: 689–702.

MacArthur, R.A. 1957. On the relative abundance of bird species. *Proc. Natl. Acad. Sci., USA*, **43**: 293–295.

May, R.M. 1975. Patterns of species abundance and diversity. In *Ecology and Evolution of Communities* (M.L. Cody and J.M. Diamond, eds.), pp. 87–120. Cambridge, MA: The Belknap Press of Harvard University Press.

Miller, A.I. and Foote, M. 1996. Calibrating the Ordovician radiation of marine life: implications for Phanerozoic diversity trends. *Paleobiology*, **22**: 304–309.

Motomura, I. 1932. A statistical treatment of associations. *Zool. Mag., Tokyo*, **44**: 379–383.

Mouillot, D. and Leprêtre, A. 1999. A comparison of species diversity estimators. *Res. Pop. Ecol.*, **41**: 203–215.

Peters, S.E. 2001. The evenness and richness components of taxonomic diversity. *Geological Society of America, Abstracts with Program*, **33**: A-100.

Peters, S.E. 2004. Evenness of Cambrian-Ordovician benthic marine communities in North America. *Paleobiology*, **30**: 325–346.

Pielou, E.C. 1966. Species diversity and pattern diversity in the study of ecological succession. *J. Theor. Biol.*, **10**: 370–383.

Powell, M.G. and Kowalewski, M. 2002. Increase in evenness and sampled alpha diversity through the Phanerozoic: comparison of early Paleozoic and Cenozoic marine fossil assemblages. *Geology*, **30**: 331–334.

Preston, W.H. 1948. The commonness and rarity of species. *Ecology*, **29**: 254–283.

Pryor, J.S. and Gastaldo, R.A. 2000. Paleoecological analysis of two Early Pennsylvanian mineral-substrate wetlands. *Palaios*, **15**: 3–13.

Purvis, A., Nee, S. and Harvey, P.H. 1995. Macroevolutionary inferences from primate phylogeny. *Proc. R. Soc. Lond. B*, **260**: 329–333.

Raup, D.M. 1975. Taxonomic diversity estimation using rarefaction. *Paleobiology*, **1**: 333–342.

Sanders, H.L. 1968. Marine benthic diversity: a comparative study. *Am. Nat.*, **102**: 243–282.

Sepkoski, J.J., Jr. 1978. A kinetic model of Phanerozoic taxonomic diversity. I. Analysis of marine orders. *Paleobiology*, **4**: 223–251.

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

Simpson, G.G. 1960. Notes on the measurement of faunal resemblance. *Am. J. Sci.*, **258a**: 300–311.

Slocomb, J., Stauffer, B. and Dickson, K.L. 1977. On fitting the truncated lognormal distribution to species-abundance data using maximum likelihood estimation. *Ecology*, **58**: 693–696.

Smith, B. and Wilson, J.B. 1996. A consumer's guide to evenness indices. *Oikos*, **76**: 70–82.

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

Tipper, J.C. 1980. Some distributional models for fossil animals. *Paleobiology*, **6**: 77–95.

Wagner, P.J. 2000. The effects of evenness on subsampling analyses of taxic richness. *Geological Society of America, Abstracts with Program*, **32**: A-140.

Williams, P.H. and Gaston, K.J. 1994. Measuring more of biodiversity: can higher-taxon richness predict wholesale species richness? *Biol. Conserv.*, **67**: 211–217.

Wing, S.L. and Harrington, G.J. 2001. Floral response to rapid warming in the earliest Eocene and implications for concurrent faunal change. *Paleobiology*, **27**: 539–563.

Zipf, G.K. 1949. *Human Behaviour and the Principle of Least-Effort*. Cambridge, MA: Addison-Wesley.

APPENDIX

Effects of different evenness metrics

We completed analyses using multiple evenness measures. Smith and Wilson (1996) suggest using E_{var}. Based on 14 test criteria, they suggest that E_{var} is generally superior to the 13 other metrics they evaluated, especially when equal sensitivity to minor and abundant taxa is required and the shape of the TAD is important. They also recommend E' as a measure with 'good mid-range behaviour'. They found that both E' and E_{var} were independent of the number of taxa in the sample.

Collection curves, such as those plotted in Fig. 3 using J, can be recalculated using E' (Fig. A1). Comparing the bottom row of Fig. 3 (TADs) with Fig. A1 reveals that the two metrics are very similar. For a fixed E', the geometric distribution has a slightly shallower slope than for a fixed J, but the log-normal and Zipf distributions are basically unchanged. These similarities are also seen in the nine collection curves. The overall pattern in Fig. A1 is identical to that seen in Fig. 3. The slightly shallower slopes of the geometric distributions result in slightly steeper geometric collection curves.

Collection curves recalculated using E_{var} have a large number of superficial differences, but the conclusion is still the same as with the other evenness metrics (Figs. 3 and A2). Comparing the bottom row of Fig. A1 and Fig. A2 (TADs) reveals key differences in these evenness metrics. Zipf distributions have many more very rare taxa than log-normal or geometric distributions with the same E_{var}. 'Similarly even' TADs as measured using E_{var} are, in fact, quite different and have a large number of very rare taxa. The impact of this difference is seen in the collection curve panels (Fig. A2). Using E_{var} the geometric and log-normal distributions have very similar collection curves when evenness is very high (distributions with $E_{var} > 0.5$ typically have $J > 0.9$), but the Zipf distribution is extremely depressed as expected given the shape of the TAD seen in the bottom row. Moreover, the difference in expected sampled richness becomes more pronounced for geometric and log-normal TADs as E_{var} decreases. Although the ordering of the distributions is opposite to that seen when plotted using J or E', the resulting collection curves are strongly TAD dependent.

The conclusions from analyses using E' and E_{var} are substantively the same as the results presented using J (Fig. 3). Figure A3 plots the relative abundance, standardized by the most abundant taxon, versus taxon rank, standardized by the number of taxa, for TADs of the same distribution type and evenness but a different number of taxa. If an evenness measure is independent of the number of taxa, then all four lines should fall on top of each other and only one line should be visible in each panel. However, this is true only for the geometric distributions with the same E' or E_{var}. Log-normal distributions with the same E' or E_{var} overlap at the first and last points, but differ in between. Zipf distributions with the same E' or E_{var} have no overlap. Moreover, this richness independence for geometric distributions is

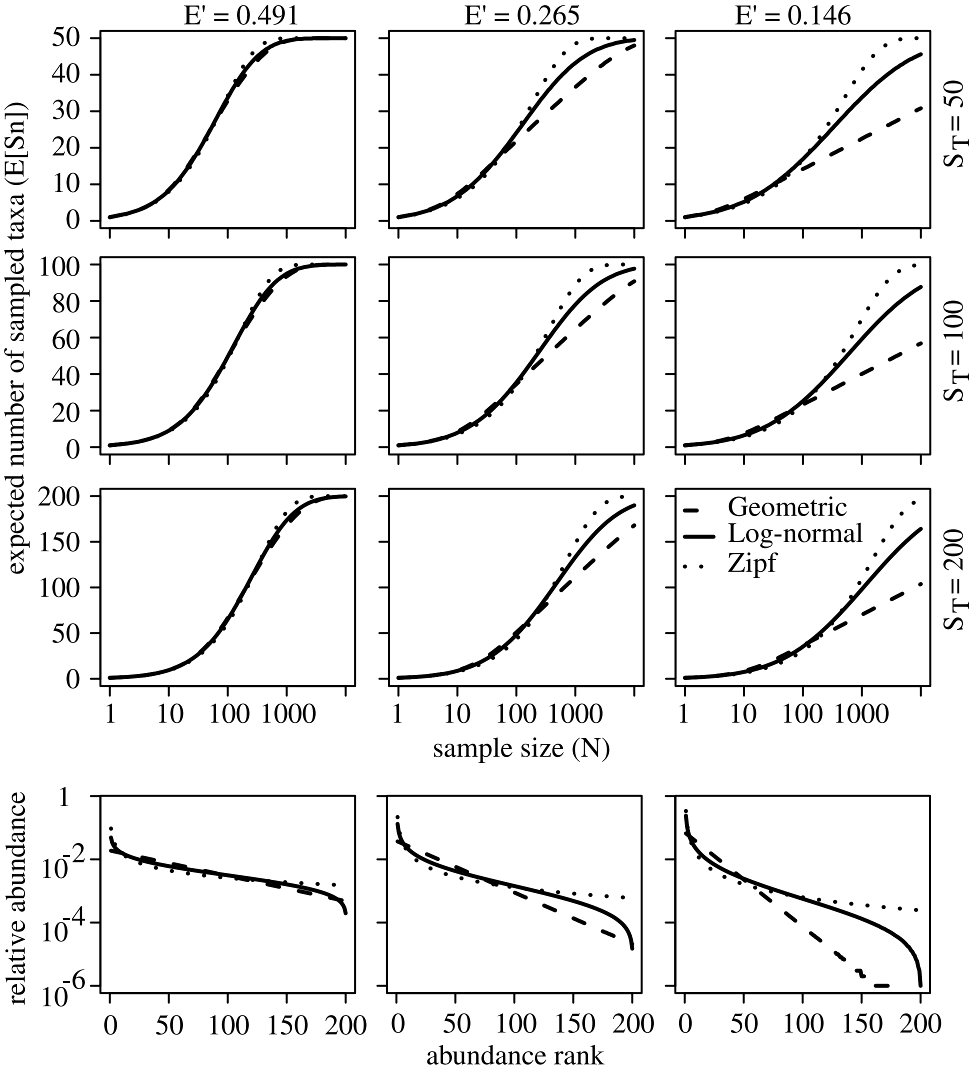

Fig. A1. Collection curves for geometric, log-normal and Zipf distributions and representative TADs. Columns with a given evenness (Camargo's E′: 0.491, 0.265, 0.146) and rows of given true number of taxa (S_T: 50, 100, 200). The bottom row plots the TADs corresponding to the collection curve row immediately above it (Camargo's E′ as per the column and $S_T = 200$). Both evenness and the number of taxa have strong effects on the shape of the collection curves, but the effects are distribution specific. Zipf distributions are always highest and geometric distributions are always lowest at larger sample sizes.

somewhat misleading. For example, two geometric distributions with the same semi-logarithmic slope *d* but different numbers of taxa will not have the same E_{var} or E′. Similarly, suppose that a geometric series with 300 taxa shares the same E_{var} or E′ with a geometric distribution with 400 taxa. Because these two distributions have different slopes, if we add one taxon to both distributions, then they will not only cease to retain the same

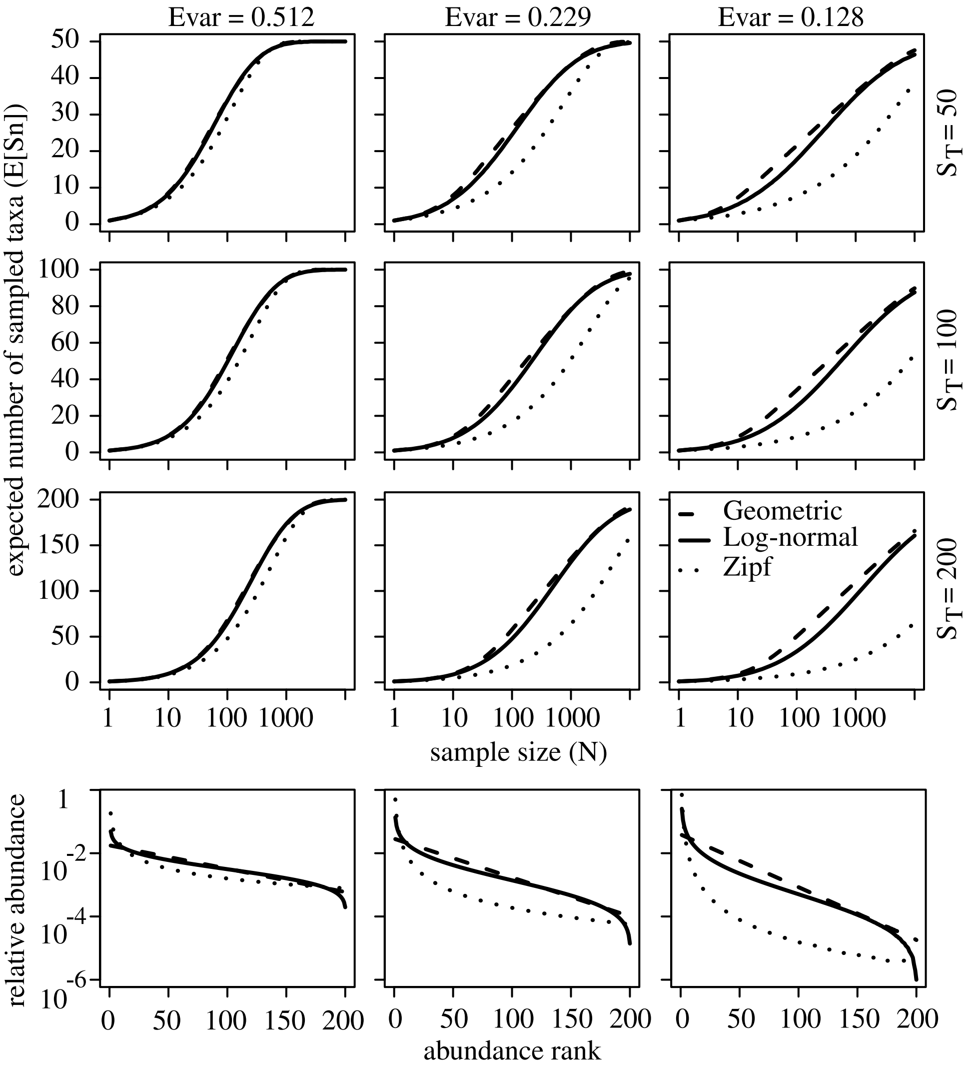

Fig. A2. Collection curves for geometric, log-normal and Zipf distributions and representative TADs. Columns with a given evenness (Smith and Wilson's E_{var}: 0.512, 0.229, 0.128) and rows of given true number of taxa (S_T: 50, 100, 200). The bottom row plots the TADs corresponding to the collection curve row immediately above it (Smith and Wilson's E_{var}: as per the column and $S_T = 200$). Both evenness and the number of taxa have strong effects on the shape of the collection curves, but the effects are distribution specific. This evenness measure is very different than the other two presented (comparing the TADs plotted in the bottom row between figures illustrates this quite clearly). The resulting collection curves have the opposite ordering as in the other evenness measures. Zipf distributions are always lowest and geometric distributions are always highest at larger sample sizes.

E_{var} and E′, but they also will cease to share the E_{var} or E′ with each other. This extends our argument that a perfect evenness measure is beyond reach, and strongly suggests that if two collections fit different model distributions, then they simply should not be compared using evenness metrics.

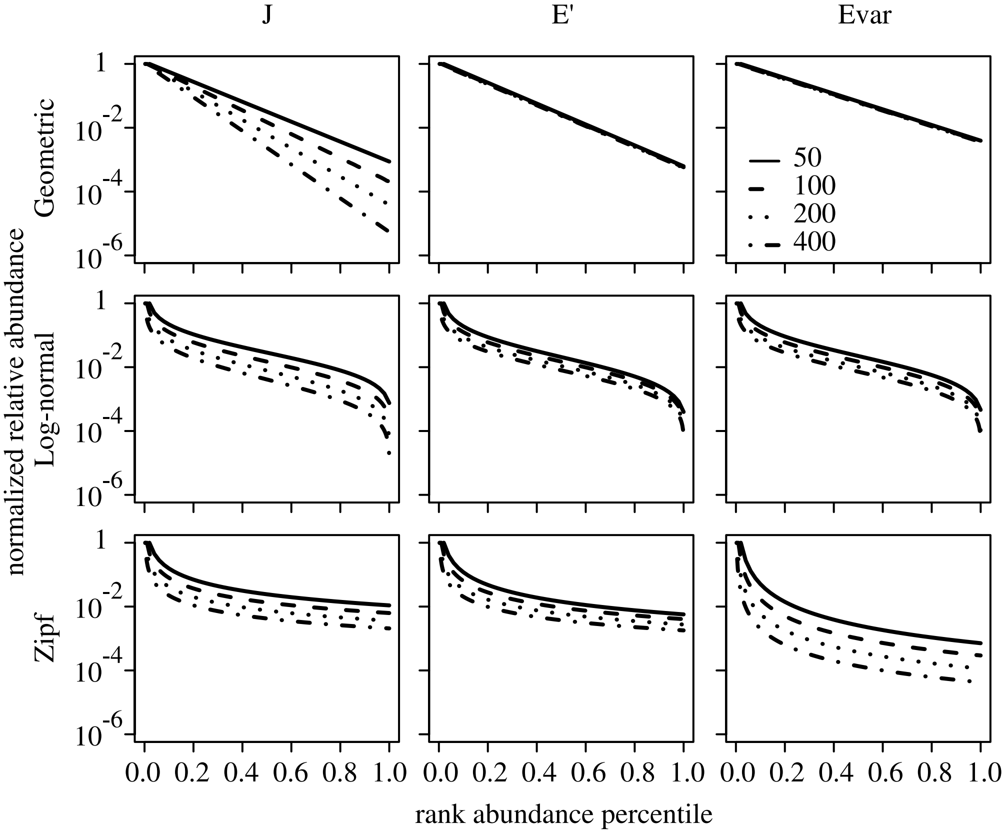

Fig. A3. Standardized taxon abundance distributions. Columns are a single evenness metric and value ($J = 0.750$, $E' = 0.265$, $E_{var} = 0.229$) and rows of single TAD. Each panel has four TADs ($S_T = 50$, 100, 200, 400) where the relative abundance (y-axis) is standardized by the abundance of the most abundant taxon and the taxon rank (x-axis) is standardized by the number of taxa. Plotting the TADs on these axes clearly illustrates the independence of E' and E_{var} from the number of taxa for these geometric distributions, but also shows the lack of independence of these metrics on the number of taxa for both log-normal and Zipf TADs.