Unveiling rare diversity by integrating museum, literature, and field data

1 author:

Paul G Harnik
Franklin and Marshall College
31 PUBLICATIONS **424** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   citizen science View project

Project   authentic geology and paleontology experiences in urban schools View project

Unveiling rare diversity by integrating museum, literature, and field data

Paul G. Harnik

Abstract.—Estimates of taxonomic richness and abundance are complicated by sampling biases. The failure to sample rare taxa is most often attributed to inadequate sampling and to removal during the process of sample-size standardization. Here I present two methods for unveiling rare diversity by integrating species presence/absence data from museum collections and the literature with quantitative estimates of species richness and abundance gathered from field-based bulk samples. Combining museum, literature, and field data can provide a more comprehensive estimate of taxonomic richness and abundance without substantial increase in current sampling effort. First, in a given bulk sample, the lowest proportional abundance value observed can be used to estimate the maximum abundance of rare species known to have occurred at the locality at least once but not recorded in the current sample. Second, a model-selection approach can be used, in which a set of relative abundance distribution models are fit to the bulk-sample abundance data and the parameter estimates for the best model used to calculate the abundance distribution for all species known from the locality. The Paleogene marine fossil record of the U.S. Gulf Coastal Plain is suitable for applying these methods, because (1) the molluscan fauna is well represented in museum collections and the literature, (2) the molluscan fauna has been taxonomically standardized, and (3) many classic localities remain accessible for standardized bulk sampling. I introduce these methods by applying them to a single locality and then, using the faunas of the Gosport, Moodys Branch, and Red Bluff Formations, I demonstrate how the model-fitting approach can be used to compare taxonomic richness among multiple localities. A substantial fraction of the molluscan richness known from each locality is not captured in bulk samples and much of this unobserved richness may be attributed to the rarity of species. The multiple-locality comparison suggests that the greatest Paleogene decline in standing richness occurred in the middle Eocene and that the recovery of richness following the Eocene-Oligocene extinction was quite rapid despite substantial loss of taxa. These analyses underscore the magnitude of veiled diversity in marine fossil assemblages and the potential of existing sources of data to unveil rare taxa, allowing them to be incorporated into quantitative diversity studies.

Paul G. Harnik. Committee on Evolutionary Biology, University of Chicago, Chicago, Illinois 60637. E-mail: pharnik@uchicago.edu

Incorporating Rare Taxa into Biodiversity Estimates

The major approaches to measuring biodiversity, such as bulk sampling, sampling standardization, and accumulation of historical collection records, weight richness and abundance differently and capture rare species with varying effectiveness. How can these be combined to develop a robust, synthetic biodiversity estimate? Both richness and abundance are informative, providing insights, albeit at times indirectly, into species interactions, niche partitioning, and community assembly (e.g., Pandolfi 1996; Olszewski and Erwin 2004; Jackson and Erwin 2006; Wagner et al. 2006; Bush et al. 2007; Novack-Gottshall 2007), as well as diversification rates and the selectivity of extinctions and originations (e.g., Budd and Johnson 2001; Lockwood 2003;

Kiessling and Baron-Szabo 2004). However, attempts to quantify richness and abundance in both Recent and fossil communities are complicated by sampling biases. Rare species are, by definition, encountered infrequently and their observed occurrences are strongly influenced by sampling effort (Preston 1948; Signor and Lipps 1982; Meldahl 1990; Gaston 1994; Hayek and Buzas 1997; Holland and Patzkowsky 2002; Thompson 2004). In a classic paper, Preston (1948) described the sample-size dependency of richness estimates by using the analogy of a veil behind which rare species are concealed from observation. The position of the veil line is controlled by the number of individuals sampled and the shape of the underlying abundance distribution (Preston 1948; Chisholm 2007, although see Dewdney 1998).

Given the sensitivity of biodiversity estimates to sampling effort, many have advocated the use of rarefaction or similar sampling standardization protocols in comparing patterns of relative richness (Raup 1975; Bush et al. 2004; Crampton et al. 2006; Kowalewski et al. 2006; Alroy et al. 2008). In recent years, analyses of evenness have also been increasingly emphasized (Powell and Kowalewski 2002; Olszewski 2004; Peters 2004, 2006; Bulinski 2007), in part because of the insensitivity of some evenness metrics to sample size (Hurlburt 1971; Peters 2004; Bulinski 2007), and because sample-level differences in evenness affect taxonomic richness estimates (Sanders 1968; Powell and Kowalewski 2002; Peters 2006). Although these methodological advances have been critical in evaluating changes in biodiversity while minimizing some sampling biases, they have also tended to focus attention on the dynamics of abundant and common taxa. Sampling standardization frequently results in comparison among only those taxa sufficiently common to be found in sample sizes of 100 or fewer specimens (e.g., Kowalewski et al. 2006), whereas sample-size ''independent'' evenness metrics provide information primarily about the numerical dominance of top taxa, largely neglecting the abundances of most taxa in any given assemblage (Kosnik and Wagner 2006).

Surveys of Recent shallow marine habitats (Sanders 1968; Schlacher et al. 1998; Ellingsen 2001; Bouchet et al. 2002; Shin and Ellingsen 2004; Zuschin and Oliver 2005) underscore how much benthic richness in the global oceans is rare. In New Caledonia, for example, approximately 50% of macrobenthic molluscan species were represented by five or fewer specimens in a survey of over 127,000 individuals (Bouchet et al. 2002). Similarly, in Hong Kong, 38% of macrobenthic invertebrate species were known from only one or two individuals out of a sample size of 16,334 specimens (Shin and Ellingsen 2004). Although elevated tropical richness may accentuate the patterns cited here, similar results have also been reported for temperate shelf faunas (Ellingsen 2001). In most Recent marine and terrestrial communities that have been surveyed, rare species compose the bulk of taxonomic diversity (Gaston 1994; Brown 1995; Kunin and Gaston 1997; Gaston and Blackburn 2000, and references therein).

The magnitude of rare richness prompts questions about the ecological and evolutionary properties of rare species. For example, to what extent does rarity increase extinction risk? How rapidly does richness recover following extinction, and how is this partitioned with respect to abundance? Do rare and abundant taxa differ in per capita rates of origination? Given that ubiquitous sampling limitations prevent quantitative field-sampling approaches from fully addressing such questions, is there a way to study diversity that takes into account the often considerable number of rare species? In this paper, I present two methods for unveiling rare diversity by integrating species presence/absence data available in museum collections and the literature (hereafter referred to as ''faunal lists'') with species abundance data gathered from field-based random samples (hereafter referred to as ''bulk samples'').

To explore diversity patterns of species beyond those recovered in bulk samples of limited size, we can capitalize on historical sampling effort by augmenting bulk-sample abundance data with faunal-list occurrence data in two ways. First, for a given bulk sample (or set of pooled samples), we can estimate the maximum relative abundance of rare species found at least once before at that spatial and temporal scale but not observed in a new bulk sample of individuals by dividing one by the current sampling effort (i.e., the sample size of individuals). Second, using a model-selection approach, we can fit a set of relative abundance distributions (RADs) to the bulk-sample data and use the parameter estimates for the best RAD to model the abundances of all species reported previously from the locality including those not found in the new bulk sample. These approaches enable rare species to be incorporated into quantitative studies of macroecology (e.g., body size-abundance relationships), macroevolution (e.g., abundance-extinction relationships), and taxonomic richness without substantially increasing new sampling effort. Although gathering new diversity data is essential in regions where little

TABLE 1. The sampling domains of bulk collections at the Gosport, Moodys Branch, and Red Bluff localities. For each study, the number of sampled outcrops is one with the exception of Red Bluff in which bulk samples were pooled by the original authors from two outcrops (see Hansen et al. 2004 for details). Sample spacing is the mean vertical spacing of bulk samples. Sampling extent is a measure of the overlap in coverage of bulk-sample and faunal-list data sets, measured for a given locality as the maximum vertical separation between bulk samples divided by the thickness of the focal stratigraphic unit; faunal lists are resolved to the locality scale and are assumed to sample the complete spatial and temporal extent of the focal stratigraphic unit at the locality scale. Stratigraphic thicknesses were compiled from the following references: Gosport (Harrison 1994); Moodys Branch (Dockery 1986a); the stratigraphic thickness for the Red Bluff locality was estimated using the maximum recorded thickness of that unit in Wayne Co., MS as a proxy due to the lack of a detailed measured section. None of the authors collected lateral replicate bulk samples, but the Gosport data at CB pooled across studies probably includes a lateral component.

	No. of outcrops	No. of bulk samples	No. of individuals	Sample spacing	Sampling extent (maximum sample spacing/ unit thickness)	Reference
Red Bluff	2	7	3746	0.33 m	82% (3.5 m/4.25 m)	Hansen et al. (2004)
Moodys Branch	1	5	6171	0.45 m	100% (4.3 m/4.3 m)	Elder (1981)
Gosport	1	7	6733	<0.50 m	7% (0.50 m/7.05 m); samples collected ~2 m above base of Formation	CoBabe & Allmon (1994)
	1	18	16,898	Contiguous	100% (7.05 m/7.05 m)	Harrison (1994)
	1	3	63,499	Not reported	Not reported	Swindel (1986)

prior research has been conducted, I will show here that for well-studied systems (e.g., Paleogene and Neogene of the Gulf and Atlantic Coastal Plains, Ordovician of the Cincinnati Arch) we can benefit from the emphasis on richness of historical data by using these to extend bulk-sample diversity estimates.

Data and Methods

Two methods are presented for estimating the abundances of ''list-only'' species (i.e., those species reported previously from the locality but not found in bulk samples). The first involves the assignment of maximum proportional abundance values to specific ''list-only'' species, and the second uses the RAD observed in the bulk-sample data to model the abundances of all species reported from that stratigraphic unit at that locality. These methods require quantitative abundance data, ideally derived from replicate bulk samples, and faunal lists compiled at comparable spatial and temporal scale from museum records and published occurrences. The Paleogene fossil record of the Gulf Coastal Plain is suitable for exploring these methods because the molluscan fauna is well represented in museum collections and the literature and has been the subject of extensive taxonomic standardization (Palmer and Brann 1965, 1966; Dockery 1977, 1982; MacNeil and Dockery 1984), and because many historical localities remain ac-

cessible for standardized bulk sampling. Furthermore, extensive work on the Paleogene stratigraphy of the Gulf Coastal Plain has resulted in a sequence stratigraphic framework (Dockery 1986b; Mancini and Tew 1992; Davidoff and Yancey 1993; Tew and Mancini 1995; Ivany 1998; Jaramillo and Oboh-Ikuenobe 1999) resolved to approximately one-million-year (Myr) intervals using nannoplankton and foraminiferal zones. Mollusks are a particularly well-suited group in which to apply these methods because they are diverse and abundant in Recent and ancient shallow marine environments and exhibit significant agreement in species rank-order abundances between life and death assemblages (Kidwell 2001, 2002; Kowalewski et al. 2003; Lockwood and Chastant 2006).

Data Compilation.—To illustrate these methods, I compiled bulk-sample data and faunal lists from the literature (see supplemental data online at http://dx.doi.org/10.1666/07062.s1). Bulk-sample data for gastropods and bivalves from the Gosport Formation at Claiborne Bluff (CB), Monroe County, Alabama, were compiled from studies by Swindel (1986), CoBabe and Allmon (1994), and Harrison (1994). The authors of each study bulk-sampled the Gosport at CB in somewhat different ways (Table 1): Harrison (1994) divided the stratigraphic section into contiguous 30 cm units and then trenched and sieved each of

these; Swindel (1986) collected stratigraphic samples from the lower, middle, and upper portions of the stratigraphic section; CoBabe and Allmon (1994) collected multiple replicates from the same stratigraphic position in the lower third of the section. Gastropod individuals were counted as numbers of apices, and bivalves were corrected for the number of individuals either by dividing the total count of valves for each species by two (CoBabe and Allmon 1994; Harrison 1994) or by using the greater number of either left or right valves (Swindel 1986); this correction was made by the original authors with the exception of the CoBabe and Allmon data which I corrected prior to analysis. Minimum sieve size varied among studies, ranging from 1 mm (Swindel 1986) to 3 mm (CoBabe and Allmon 1994; Harrison 1994).

A faunal list for the Gosport at CB was compiled from publications by Palmer and Brann (1965, 1966), with the addition of ten species found in bulk samples by CoBabe and Allmon, Harrison, and Swindel, but not previously listed as occurring in that unit at that locality. The faunal list included only those species described as occurring in the Gosport at CB; Gosport taxa with uncertain locality information and CB taxa with uncertain stratigraphic assignments were excluded. Taxonomic inconsistencies among the three studies were minimized by removing all indeterminate occurrences and standardizing the remaining bulk-sample data following the taxonomy of Palmer and Brann (1965, 1966), which synthesized a century of Paleocene and Eocene molluscan systematics in the Gulf and Atlantic Coastal Plains. Few taxonomic changes were required because Palmer and Brann's publications were the primary references used for the three studies examined here. Following taxonomic standardization, the bulk-sample data set contained 150 species from a total of 28 samples, comprising 87,130 individuals, with a median sample size of 920 individuals. When pooled, these bulk-sample data are assumed to sample the underlying species pool over the temporal and spatial extent of the Gosport exposure at CB. This assumption is reasonable given that bulk samples encompass the complete vertical thickness of the

Gosport at CB (Table 1). Possible spatial variation is less well constrained because none of the authors collected laterally spaced replicate samples, but it is likely that the data pooled across studies encompass both lateral and vertical variation if such exists. The spatial and temporal distributions of species occurrences in the associated faunal list are resolved to the locality scale, and specimens may have come from anywhere in the exposure.

In addition to the data for the middle Eocene Gosport Formation at CB, bulk-sample data were compiled from the literature for two other localities (Table 1): the late Eocene Moodys Branch Formation at Fossil Gulch, Hinds County, Mississippi (Elder 1981), and the early Oligocene Red Bluff Formation along the Chickasawhay River near Hiwanee, Wayne County, Mississippi (Hansen et al. 2004). The Red Bluff bulk-sample data were pooled by the original authors from samples collected from two localities, Mississippi Geological Survey localities MGS-35 and MGS-38 (MacNeil and Dockery 1984), located less than 2 km apart with equivalent stratigraphy (see discussion in Hansen et al. 2004). Counting protocols for the Moodys Branch and Red Bluff localities were comparable to those described above for the Gosport at CB and the minimum sieve size used was 1 mm. Faunal lists were generated from published species occurrences for each of these localities and both bulk-sample data and faunal lists were taxonomically standardized prior to analysis (Palmer and Brann 1965; Dockery 1977, 1982; MacNeil and Dockery 1984; Dockery and Lozouet 2003). Bulk samples were distributed over more than 80% of the vertical extent of the stratigraphic units exposed at each of these localities (Table 1) and provide a representative sample of the diversity and abundance of marine mollusks at the locality scale.

Samples at the middle and late Eocene localities come from well-mixed fossiliferous sands deposited on the shelf during maximum transgression (Mancini and Tew 1992). The early Oligocene samples come from sandy fossiliferous lenses preserved in silty clays deposited in shelf and delta-margin environments (Dockery 1982), likely during sea level highstand (Mancini and Tew 1992; Jar-

amillo and Oboh-Ikuenobe 1999, though see Pasley and Hazel 1995 and Hansen et al. 2004 for alternative sequence stratigraphic interpretations). Molluscan preservation at these three localities is exceptional, with primary aragonite preserved at all localities, making these data well suited for comparisons of richness and abundance within largely iso-taphonomic conditions.

Relative Abundance Distributions and Model Selection.—To incorporate ''list-only'' species into comparative analyses of richness it is necessary to estimate the sampling effort that would be necessary to unveil them. To determine the bulk-sample size that would be necessary to capture the cumulative known richness, I estimated the relative abundances of all species known from a locality by using a model of the RAD derived from the observed bulk-sample data. For each locality, I used the maximum-likelihood method to fit a candidate set of RAD models (broken stick, geometric, lognormal, Zipf, Zipf-Mandelbrot) to the observed bulk-sample data, providing estimates of the parameter values and the log-likelihood of each model given the data. The candidate set of models included representatives from most of the major families of RAD models, and the candidates have shapes similar to many RADs not considered (Marquet et al. 2003; McGill et al. 2007); the zero-sum multinomial (Hubbell 2001) was not included in the candidate set of models because many algorithms for evaluating the zero-sum fail to converge in a reasonable time given sample sizes as large as those analyzed here. The equations and routines used to calculate the expected abundances of taxa for each model are outlined in several works on RADs (Frontier 1985; Mouillot and Lepretre 1999; Kosnik and Wagner 2006). Akaike's Information Criterion (AIC) was then used to select the best model (Akaike 1974; Burnham and Anderson 2002) through the calculation of the Akaike weights. AIC is a parsimony-based approach and its calculation can be thought of as a measure of the fit of a given model to the data penalized by the complexity of the model. AIC is calculated as follows:

$$\text{AIC} = -2\ell + 2k \qquad (1)$$

where ℓ is the log-likelihood of the model and k is the number of estimated parameters. The Akaike weights are calculated as

$$w_i = \exp\left(-\frac{1}{2}\Delta_i\right) \bigg/ \left[\sum_{r=1}^{R} \exp\left(-\frac{1}{2}\Delta_i\right)\right] \qquad (2)$$

where w_i is the weight of the ith model, Δ_i is the difference in AIC values between the ith model and the best model (identified as the model with the smallest AIC value), and this is then scaled to the sum of Δ_i values for all models in the set. Akaike weights sum to 1 and are a measure of the relative support for each model in the candidate set given the observed data. An Akaike weight >0.95 was used as the criterion for model selection. The percentage of explained deviance for each model, calculated as $100 \cdot ([\text{null deviance} - \text{residual deviance}]/\text{null deviance})$, was used as a measure of goodness-of-fit (McCullagh and Nelder 1989).

Using the estimated RAD model parameters, I then calculated the expected abundance of each species known for each locality, including those not found in bulk samples. Assuming the rarest species at each locality was known from only a single specimen, I estimated the sample size that would be necessary to recover the known bivalve and gastropod species.

Assessing the Sensitivity of Model Selection and Parameter Estimation to Variations in Sample Size.—I illustrate these methods with data from extensive bulk sampling and more than a century of qualitative collecting in the Gulf Coastal Plain. But are tens of thousands of specimens required to apply this approach? The question of sample-size sensitivity has two components: model selection and parameter estimation. To address model-selection uncertainty, I used the CB Gosport data and compared the mean Akaike weights for the five different RAD models as a function of sample size, using a bootstrap resampling procedure with 1000 replicates per sample size. I evaluated the effects of sample size on parameter estimation in a similar fashion, fitting the set of RADs to a range of sample sizes and comparing the resulting parameter estimates with those obtained from models fit to the total sample.

Assessing Whether Individual Localities Shared the Same RAD.—Unless observations are identical, the hypothesis that two or more localities had unique RADs will always be the most likely. However, given uncertainty in parameter estimation, a null hypothesis that localities shared the same RAD requires rejection in this case. This step is necessary because the modeled RADs are then used to compare the relative differences in taxonomic richness among localities. To address this, I used a bootstrap resampling routine in which the RADs were fit repeatedly to the bulk-sample data at each locality in question, and then contrasted the mean model parameters and 95% confidence intervals.

All analyses were conducted using the statistical programming environment R, version 2.3.1 (R Development Core Team 2006). RAD models were fit to observed abundances of species using the community ecology package ''vegan'' (Oksanen et al. 2007), with the Zipf-Mandelbrot distribution (Frontier 1985) fit using code modified for application to large data sets provided by J. Oksanen (personal communication 2006).

Results

Comparing Faunal-List and Bulk-Sample Richness.—The expectation is that the total known richness compiled from more extensive and, in many cases, nonrandom sampling should exceed that observed in bulk samples. An observed deficit in the richness of bulk samples may be due to the rarity of ''list-only'' species. These expectations were assessed by comparing the richness of faunal-list and bulk-sample data sets and then examining the distribution of rarity among species observed in bulk samples. Total known richness for the Gosport at CB (i.e., the sum of faunal-list and bulk-sample richness) greatly exceeds bulk-sampled richness, with only 41% of known species recovered in bulk samples (Fig. 1). A small number of species recovered in bulk samples were not previously known from the Gosport at CB. These ''bulk-sample-only'' species were all rare (ten species with a median numerical abundance of 22 specimens and proportional abundance = 0.0002).

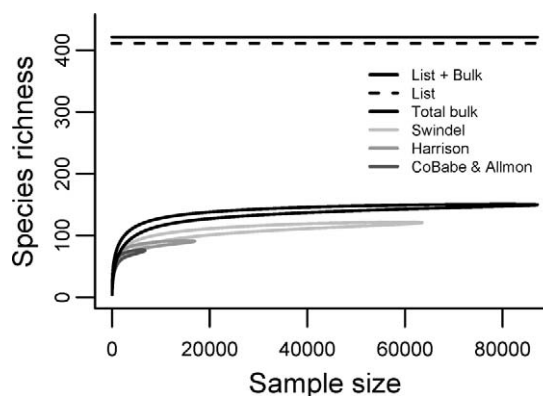The large discrepancy in richness between

FIGURE 1. Rarefaction curves for the middle Eocene Gosport Formation at Claiborne Bluff (CB), Alabama. The horizontal dashed line indicates the species richness reported from the Gosport at CB in museum collections and the literature ($S = 411$). The slightly elevated solid black line indicates the cumulative species richness from the Gosport at CB, including species found in bulk samples as well as those found in museum collections and the literature ($S = 421$). Rarefaction curves from top to bottom follow the order presented in the legend. Over 50% of species known from the Gosport at CB are not captured in bulk samples, even those of substantial size.

the faunal-list and bulk-sample data cannot be attributed simply to taxonomic inconsistency because all species names were standardized, or to species lumping due to poor monographic coverage because over 97% of bivalve and gastropod species known from the Gosport at CB have been figured (Palmer and Brann 1965, 1966). Nonetheless, some of the variation between the faunal-list and bulk-sample data undoubtedly still results from differences in taxonomic practice. Paleoecologists may tend to lump species more than systematists, or conversely systematists may split species more often, which would contribute to the discrepancy in richness observed between the faunal-list and bulk-sample data. However, differing taxonomic practices are unlikely to account for the nearly three-fold difference in richness observed between the faunal-list and bulk-sample data. To mitigate any remaining taxonomic effects, the most conservative application of the methods presented below would involve taxonomically standardizing all identifications at the specimen level, both in bulk samples and faunal lists, by using museum collections and monographic plates.

Individual bulk-sample data sets for the Gosport at CB recovered between 19% and

TABLE 2. Agreement in rank abundance and taxonomic composition between bulk-sample data sets for the Gosport at CB.

	Spearman Rho	Bray-Curtis dissimilarity
CoBabe & Allmon vs. Swindel	0.24**	0.83
CoBabe & Allmon vs. Harrison	0.67***	0.43
Swindel vs. Harrison	0.19*	0.68

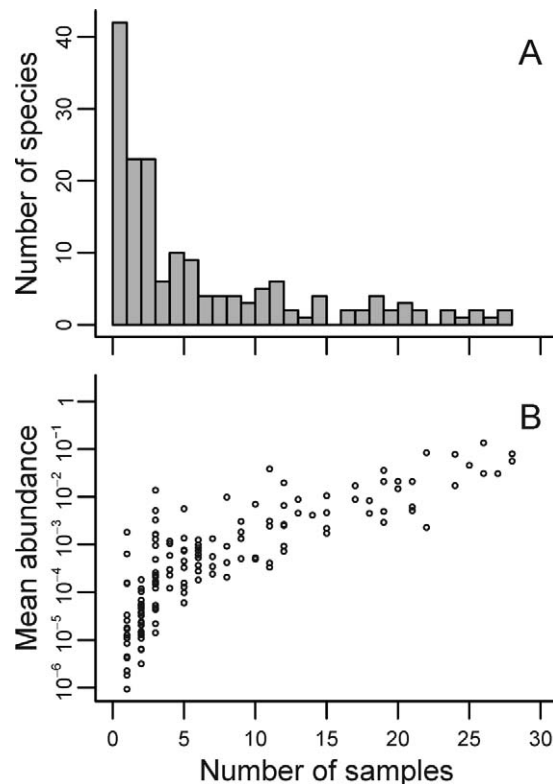*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

FIGURE 2. A, Frequency distribution of bulk-sampled occurrences for mollusk species from the Gosport at CB. Most species are known from relatively few samples although a few species are common throughout. B, The relationship between sampled occurrences and the mean proportional abundances of species in the samples in which they occur. Common species occur in more samples, and with declining abundance the frequency of occurrence also declines.

33% of the known richness, with higher recovery found in studies of greater spatial and temporal extent and larger sample size (Swindel 1986; Harrison 1994). Species rank abundances were positively correlated across all bulk-sample data sets using Spearman rank correlation tests, although the strengths of these correlations varied (Spearman rho values of 0.19 to 0.67; Table 2). Bray-Curtis dissimilarity values also varied among the three data sets, ranging from 0.43 to 0.83, reflecting differences in the taxonomic composition of bulk samples (Table 2). These differences may reflect variation in sampling methods, spatial or temporal variation in the underlying species pool being sampled, taxonomic practice, or some combination of these different factors.

Low correlations among bulk samples are expected if species are patchily distributed and/or occur at low densities. Examination of the frequency of species occurrences in bulk samples corroborates this expectation, with 46% of Gosport mollusk species found in three or fewer samples and 15% restricted to a single sample (Fig. 2A). To determine whether species found in few samples were rare or common, I plotted the relationship between the number of samples in which a species occurs and its mean proportional abundance in those samples (Fig. 2B). Locally abundant species occurred in more samples, and with decreasing abundance, the frequency of occurrence also declined (Spearman rho = 0.90, $p < 0.00001$). Positive abundance-occupancy relationships are one of the most ubiquitous macroecological patterns and may be generated through true covariance between local density and spatial occurrence and/or through poor sampling of rare taxa (Brown 1984; Gaston et al. 1997; Schlacher et al. 1998). Either mecha-

nism implies that taxa sampled infrequently were rare.

Assuming that the abundance-occupancy relationship observed in the bulk-sample data is general for the Gosport fauna at CB as a whole, we can assign an abundance value to species found previously but absent in the new bulk sample of individuals. These "list-only" species were probably no more abundant than the rarest bulk-sampled species and potentially much rarer. The maximum proportional abundances of rare "list-only" species can be estimated as $1/N$ where N is the bulk-sample size of individuals. The precision of these abundance estimates depends on the accuracy of bulk-sample estimates of the underlying species RAD. For the Gosport at CB, "list-only" species would be assigned an

TABLE 3. Model selection criteria used to identify the best model among the set of five RADs fit to the CB Gosport bulk-sample data. k, number of model parameters; AIC, AIC values; Δ_i, AIC differences between the model and the best model with low values indicating greater support; w_i, Akaike weights for each model; ED, percent explained deviance. A model with an Akaike weight >0.95 is unambiguously identified as the best model in the set given the data.

RAD Model	k	AIC	Δ_i	w_i	ED
Zipf-Mandelbrot	3	2,792.5	0	>0.999	99.1%
Lognormal	2	15,299.1	12,506.6	≪0.0001	95.0%
Geometric	1	15,453.1	12,660.6	≪0.0001	94.9%
Zipf	2	30,371.4	27,578.9	≪0.0001	90.1%
Broken stick	0	107,113.4	104,320.9	≪0.0001	64.9%

abundance of 1/87,130. Although this will likely overestimate the proportional abundance of some rare species, the estimated abundance value for these ''list-only'' species will still be ''rare,'' given sufficiently large samples. This approach is particularly useful for clade or guild-based analyses in which the total sample size is easily calculated (e.g., the total number of bivalve individuals) but species-level abundance and faunal-list occurrence data are gathered only for a subset of the preserved fauna (e.g., species of scallops). If one were to estimate the abundances of all ''list-only'' species as 1/N, the sum of proportional abundances would exceed 100%, which may necessitate recalculating species proportional abundances depending on the specific research questions being addressed. Estimating the maximum proportional abun-
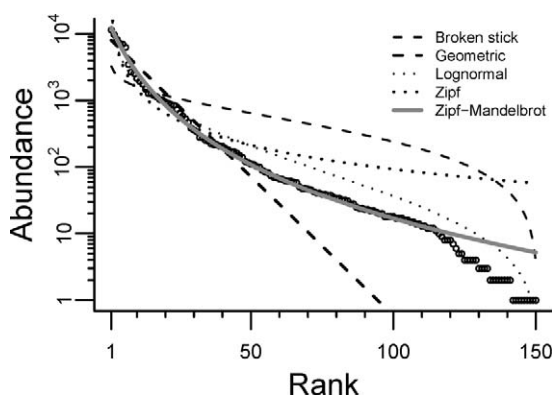

FIGURE 3. Set of five relative abundance distribution (RAD) models fit to bulk-sample data from the Gosport at CB. For these data the best-supported model is the Zipf-Mandelbrot distribution (see Table 3 and discussion in text). Absolute rather than relative abundance is plotted here to illustrate the number of species known from only one or two specimens even at large sample sizes.

dances of ''list-only'' species as 1/N allows poorly sampled rare taxa to be incorporated into macroecological and macroevolutionary analyses such as abundance–body size and abundance-extinction relationships.

To incorporate ''list-only'' species into comparative analyses of richness it is necessary to determine the sampling effort that would be necessary to unveil them with greater precision than that provided by the maximum proportional abundance approach described above. To determine the bulk-sample size that would be necessary to capture the cumulative known richness, I estimated the relative abundances of all species known from a locality using a model of the RAD derived from the observed bulk-sample data. Using the model-fitting approach, the Zipf-Mandelbrot RAD was identified as the best model in the set given the observed data for the Gosport at CB (w_{ZM} ≫ 0.99; Table 3, Fig. 3). The Zipf-Mandelbrot RAD model explains >99% of the observed deviance in the Gosport CB data and is largely an unbiased estimator of observed species abundances (Fig. 4). Using the estimated Zipf-Mandelbrot model parameters, I calculated the expected abundance of each species known from the Gosport at CB, including those not found in bulk samples (Frontier 1985; Mouillot and Lepretre 1999).

Assuming the rarest species at CB was known from only a single specimen and that the best fit RAD is an accurate description of the underlying RAD at that spatial and temporal scale, an estimated sample size of 416,548 specimens would be necessary to recover the 421 known bivalve and gastropod species. The estimated sampling effort required to unveil the known richness is ap-
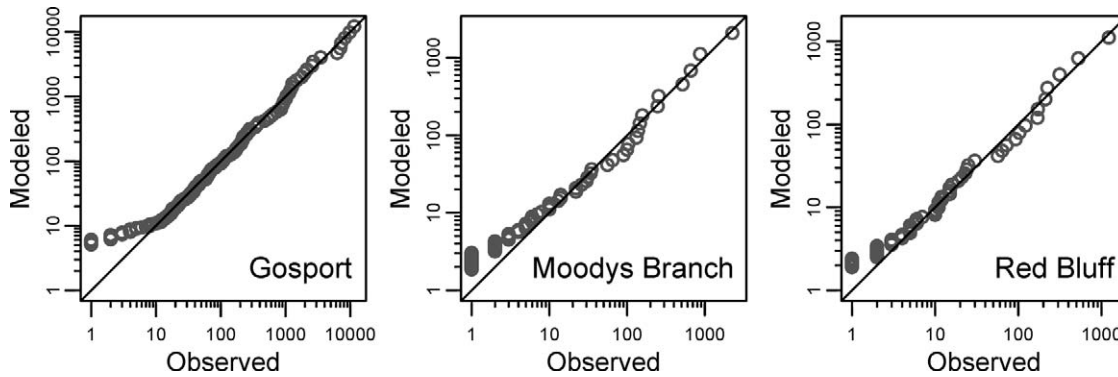
FIGURE 4. Observed versus modeled abundance data for the Gosport, Moodys Branch, and Red Bluff localities. Species abundances modeled using the Zipf-Mandelbrot RAD are largely unbiased estimators of the observed abundances of species, with most values falling along the line of unity. Where modeled and observed values systematically differ is in the rare tail, in which modeled values overestimate the observed abundances of the rarest species (though the magnitude of this offset is accentuated by plotting the data on logarithmic axes).

proximately five times the current sampling effort and unlikely to be met using only quantitative field-based approaches. By estimating the RAD of the cumulative faunal list and the sampling effort necessary to unveil it, this second approach enables rare species to be incorporated into sample-standardized comparisons of richness.

Comparing Richness among Multiple Localities.—The three Paleogene localities considered here vary in bulk-sampling effort (N), bulk-sampled richness (S_{Bulk}), and total known richness ($S_{Cumulative}$) (Table 4). As with the middle Eocene example, bulk-sampled richness for both the late Eocene and early Oligocene data falls short of the cumulative richness known from each locality ($S_{Cumulative} > S_{Bulk}$). Approximately 100 species are known from the late Eocene Moodys Branch and early Oligocene Red Bluff Formations at these localities, of which approximately 60% are recovered in the bulk-sample

TABLE 4. Summary of the taxonomic richness of the three Paleogene Gulf Coastal Plain localities exposing the Red Bluff, Moodys Branch, and Gosport Formations respectively. Localities vary in bulk-sampling effort (N), bulk-sampled richness (S_{Bulk}), sample-standardized richness at a sample size of 3700 individuals ($S_{Bulk(N=3700)}$), and the cumulative known richness compiled from museum, literature, and bulk-sample data ($S_{Cumulative}$).

Locality	N	S_{Bulk}	$S_{Bulk(N=3700)}$	$S_{Cumulative}$
Red Bluff	3746	65	64	106
Moodys Branch	6171	62	55	101
Gosport	87,130	150	101	421

data. How much additional sampling would be needed to unveil the known local richness? Furthermore, do the patterns of relative richness among localities change when these poorly sampled rare species are incorporated into comparisons?

Using the model-fitting approach, the lognormal RAD is better supported than the Zipf-Mandelbrot for both the Moodys Branch and Red Bluff data sets (Table 5). However, when the lognormal is used to estimate the number of veiled species (i.e., the number of species represented by zero specimens), these estimates fall short of the cumulative richness known for each of these localities (e.g., total richness for the Red Bluff is estimated to be 90 using the lognormal parameters although 106 species are already known). If the lognormal model is rejected post hoc as being inconsistent with the excess of rare species observed at each of these localities, then the next best model in both cases is the Zipf-Mandelbrot. Examining the evidence ratio for the Zipf-Mandelbrot relative to the third ranked model, calculated as the ratio of Akaike weights for these two models (Burnham and Anderson 2002), confirms that the Zipf-Mandelbrot is much better supported than the next best model for both the Moodys Branch and Red Bluff (evidence ratios of 2.07e+81 and 5.88e+40 respectively). For these two localities the Zipf-Mandelbrot RAD explains >96% of the deviance in the observed data (Table 5, Fig. 4).

TABLE 5. Model selection criteria used to identify the best model among the set of five RADs fit to the middle Eocene (Gosport, G), late Eocene (Moodys Branch, MB), and early Oligocene (Red Bluff, RB) data sets. k, number of model parameters; AIC, AIC values; Δ_i, AIC differences between the model and the best model, with low values indicating greater support; w_i, Akaike weights for each model; ED, percent explained deviance. A model with an Akaike weight >0.95 is unambiguously identified as the best model in the set given the data.

RAD model	k	AIC			Δ_i			w_i			ED		
		G	MB	RB	G	MB	RB	G	MB	RB	G	MB	RB
Zipf-Mandelbrot	3	2792.5	472.97	403.81	0	38.21	32.13	>0.999*	≪0.0001	≪0.0001	99.1	97.8	96.6
Lognormal	2	15,299.1	434.76	371.68	12,506.6	0	0	≪0.0001	>0.999*	>0.999*	95.0	98.0	96.9
Zipf	2	30,371.4	847.44	591.56	27,578.9	412.68	219.88	≪0.0001	≪0.0001	≪0.0001	90.1	96.1	95.0
Geometric series	1	15,453.1	1818.76	1338.93	12,660.6	1384	967.25	≪0.0001	≪0.0001	≪0.0001	94.9	91.7	88.7
Broken stick	0	107,113.4	8163.62	4135.45	104,320.9	7728.86	3763.77	≪0.0001	≪0.0001	≪0.0001	64.9	62.8	65.1

Using the parameter estimates for the Zipf-Mandelbrot RAD, the sample sizes necessary to capture the known richness can be calculated for each locality. For the Moodys Branch, a sample size of 9341 individuals is expected to reveal the 101 species known from the locality, and for the Red Bluff, 5345 individuals are expected to reveal the known richness of 106 species. Sampling at the late Eocene Moodys Branch and early Oligocene Red Bluff localities would need to increase by a factor of 1.5 and 1.4 respectively over current sampling efforts. For both of these localities the known richness and sample sizes necessary to recover it pale in comparison to the middle Eocene locality. This is likely an artifact of sampling effort (i.e., the middle Eocene Gosport at CB has been sampled intensively for over a century) as well as a real difference in the underlying diversities of these stratigraphic units.

Treating the modeled species abundances as the sample for each locality, I calculate the species richness expected for a given sample size using analytical rarefaction (Raup 1975; Tipper 1979). Sample-standardized richness estimates that incorporate both faunal-list and bulk-sample data exhibit many similarities and key differences from patterns derived exclusively from bulk samples (Fig. 5). Both data sets exhibit the same rank ordering of localities with respect to species richness, but the relative differences in richness change as sample size increases and rare, ''list-only'' species are included in the comparisons (Figs. 5, 6). For example, the middle Eocene locality still exhibits the highest standing richness, but the decline in richness to the late Eocene is damped somewhat by the inclusion of rare late Eocene species. Similarly, although both the modeled and bulk-sample data sets indicate that the early Oligocene locality was marginally more diverse than the late Eocene locality, this difference becomes more apparent when rare early Oligocene species are incorporated into the comparison (Fig. 5). All else being equal, comparable degrees of time-averaging among localities may further strengthen the pattern of elevated early Oligocene richness because the early Oligocene samples come from highstand deposits whereas the middle and late Eocene samples are derived from transgressive-systems-tract shell beds (Dockery 1982; MacNeil
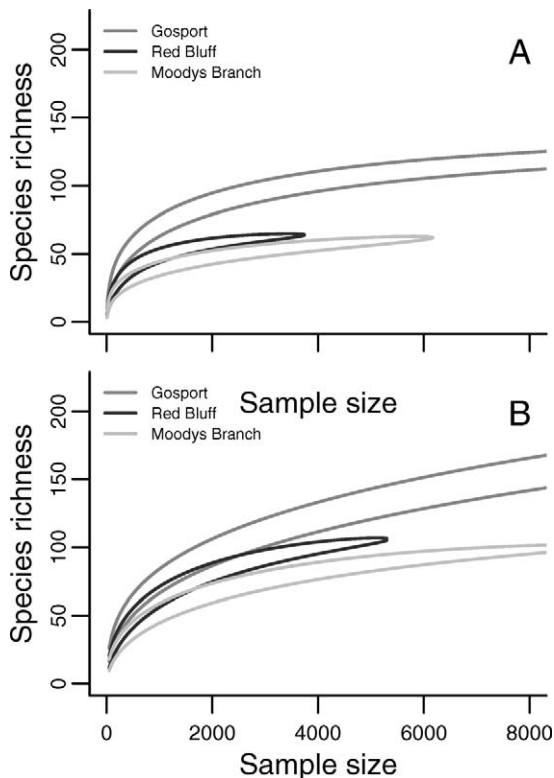
FIGURE 5. Rarefaction curves for three Paleogene Gulf Coastal Plain localities. A, Rarefaction curves derived solely from bulk-sample data. B, Rarefaction curves derived from modeled abundance data that incorporates both bulk and faunal-list data. Rarefaction curves from top to bottom follow the order presented in the legend. Relative differences in richness among localities are observed between the bulk and modeled data sets, although the rank ordering of localities by richness does not change.
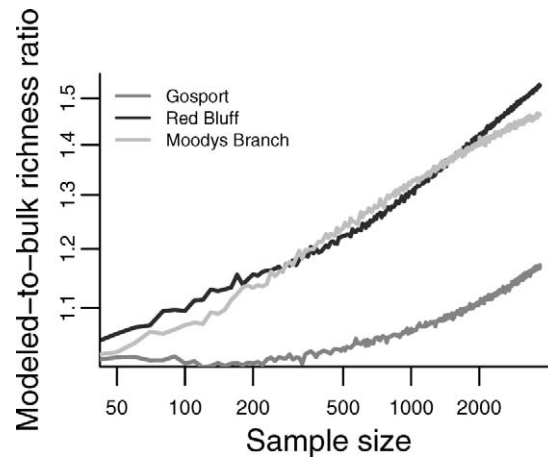

FIGURE 6. The mean ratio of modeled to bulk-sampled richness calculated by using a bootstrap resampling approach with 1000 replicates per sample size over a range of sample sizes for the three Paleogene localities. Little difference is observed in the expected richness ratio at low sample sizes, but as sampling effort increases, rare species are progressively unveiled in the modeled diversity data.

and Dockery 1984; Mancini and Tew 1992) and are expected to contain more species simply as a result of increased time-averaging (but see Scarponi and Kowalewski 2007). Although estimates of richness are sensitive to temporal binning (i.e., longer time bins should sample more species [Foote 2000]), these three stratigraphic

units were deposited under approximately equivalent lengths of time (1 Myr or less).

A regional history of molluscan taxonomic richness from synoptic Formation-level data (Dockery 1986b; Dockery and Lozouet 2003) implied a secular decline from the middle Eocene through the early Oligocene in the Gulf Coastal Plain (Table 6). In contrast, sample-standardized local richness estimates based on either bulk-sample or modeled diversity data suggest that the primary drop in richness occurred between the middle and late Eocene, with early Oligocene richness slightly elevated over the preceding interval. This is consistent with recent findings that the early Oligocene recovery interval following the Eocene-Oligocene extinction was relatively brief, with Oligocene regional richness reaching values close to that observed for the late Eocene by

TABLE 6. Comparison of species richness estimates under several different data treatments. Regional richness values are from a synoptic compilation of species stratigraphic occurrences in the central and western Gulf Coastal Plain (Dockery 1986b; Dockery and Luzouet 2003); $Local_{list}$ are the known faunal lists for the three Paleogene localities examined; $Local_{Bulk}$ are the expected richness values for bulk-sample sizes of 100 and 3700 specimens; $Local_{Model}$ are the expected richness values for the modeled diversity data at a sample size of 5300 specimens.

	Regional	$Local_{list}$	$Local_{Bulk(N=100)}$	$Local_{Bulk(N=3700)}$	$Local_{Model(N=5300)}$
Red Bluff	196	106	21 ± 5	64 ± 1	106 ± 1
Moodys Branch	346	101	18 ± 4	55 ± 4	89 ± 6
Gosport	495	421	30 ± 6	101 ± 7	134 ± 11

TABLE 7. Evenness values calculated using Hurlburt's PIE (Hurlburt 1971) for the bulk-sample versus modeled abundance data for each Paleogene locality.

Locality	PIE_{Bulk}	PIE_{Model}
Red Bluff	0.85	0.86
Moodys Branch	0.82	0.83
Gosport	0.94	0.94

the top of the early Oligocene Vicksburg Group (Hansen et al. 2004). This study corroborates that result, and finds support for rapid recovery of local richness to pre-extinction late Eocene levels within the oldest Oligocene macrobenthic assemblages in the Red Bluff Formation. This result is robust even when rare species are incorporated into the comparison. The regional decline in standing richness from the late Eocene to early Oligocene may be an artifact of more extensive preservation of shallow marine facies during deposition of the late Eocene Moodys Branch Formation (Dockery 1977, 1982).

In contrast to species richness, evenness does not differ dramatically between the bulk-sampled and modeled abundance data (Table 7). This lack of difference is expected because evenness metrics such as PIE and Ess are primarily measures of the numerical dominance of top taxa, which is not influenced greatly by the inclusion of additional rare species.

Sample-Size Sensitivity of Model Selection and Parameter Estimation.—Although differences in the shapes of RAD models can be difficult
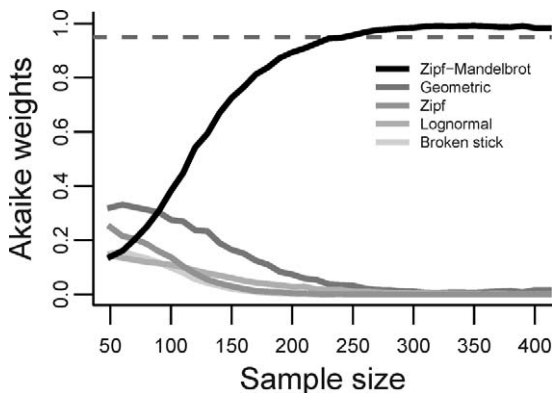

FIGURE 7. Akaike weights as a function of sample size for each of the relative abundance distribution models fit to the Gosport bulk-sample data from CB. Even at relatively small sample sizes ($N > 300$) the Zipf-Mandelbrot RAD is unequivocally identified as the best model.
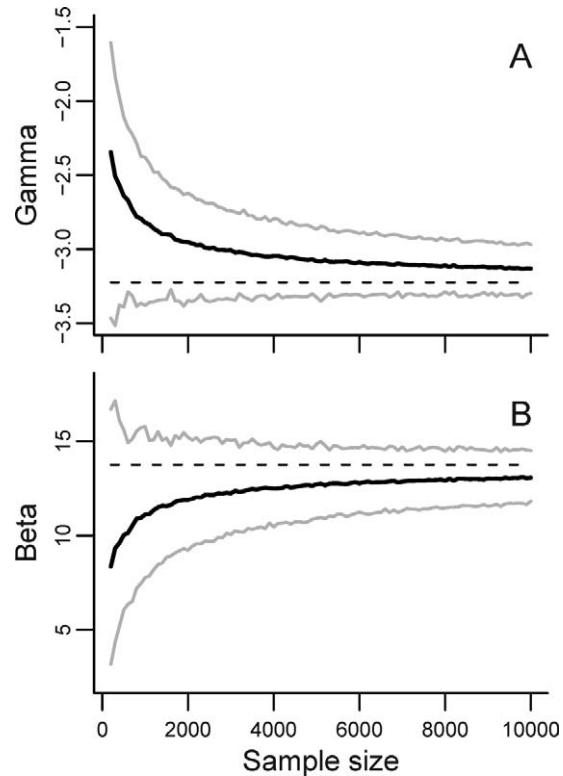

FIGURE 8. Sensitivity of parameter value estimation to sample size for the gamma and beta parameters of the Zipf-Mandelbrot model fit to the CB Gosport bulk-sample data. Solid black lines are the mean model parameter values calculated from 1000 bootstrap replicates for each sample size. Dashed lines are the maximum-likelihood parameter values estimated from fitting the model to the complete data set. Light-gray lines are the 95% bootstrap confidence intervals. A, Bootstrapped values of the gamma parameter of the Zipf-Mandelbrot distribution. B, Bootstrapped values of the beta parameter of the Zipf-Mandelbrot distribution.

to distinguish with small samples (Kosnik and Wagner 2006; Wagner et al. 2006; McGill et al. 2007), it is clear for these data from the CB Gosport that even at relatively small sample sizes (<300 specimens) differences in model support are sufficient to identify the best model unequivocally (Fig. 7). This result is consistent with a study by Wagner and collaborators (2006), who found that RAD models could be distinguished in fossil samples containing as few as 100 specimens when the underlying RAD is relatively uneven.

The relationship between parameter estimates and sample size is illustrated in Figure 8 by two of the parameters of the Zipf-Mandelbrot RAD for the CB Gosport data. For a
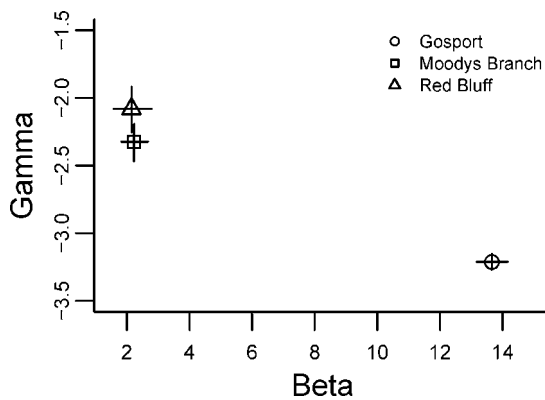
FIGURE 9. Mean bootstrapped parameter values and 95% confidence intervals for the Zipf-Mandelbrot RAD fit separately to the three Paleogene localities. The gamma and beta parameter values for the Moodys Branch and Red Bluff are significantly different and both are entirely nonoverlapping with those observed for the Gosport, providing support for unique model parameters for each locality.

given RAD, considerable variation in parameter values is seen at small sample sizes. However, with increased sample size the variance in parameter values decreases dramatically and the mean bootstrapped model parameters begin to converge on the maximum likelihood parameter estimates fit to the entire data set.

Support for Unique RAD Models for Each Locality.—Although the Red Bluff and Moodys Branch parameters are more similar to each other than either is to the Gosport, each parameter combination is significantly different ($p < 0.0001$ for t-tests of the bootstrapped gamma and beta parameters for the Moodys and Red Bluff localities), supporting the use of unique models for each locality (Fig. 9).

Discussion

The methods presented here are demanding with respect to data, requiring not only species presence/absence but also abundance data. However, they provide an opportunity to incorporate rare, poorly sampled species into paleobiological studies. These rare species can be integrated into quantitative analyses by modeling their abundances either by assigning maximum proportional abundance values or by estimating their abundances using parametric models fit to the abundance data gathered from bulk samples. This latter approach should not be miscon-

strued as an attempt to extrapolate species richness from rarefaction curves (Tipper 1979). This important point is evident by comparing Figure 5A and 5B, in which the rarefaction curves calculated using bulk-sample versus modeled abundance data differ in their overall shapes as well as endpoints. The different trajectories of modeled versus bulk-sample rarefaction curves are not due to changing the abundance structure of the observed RAD, because modeled abundances are largely unbiased estimates of the observed abundances of species (Fig. 4) and modeled and bulk sampled data sets do not differ substantially in evenness (Table 7). Rather the different trajectories between modeled and bulk-sample rarefaction curves result from increasing the frequency of species known from few specimens that constitute the long rare tails of the modeled RADs.

Rather than attempting to estimate the total number of species present at a given sampling scale, I have addressed here the much more modest problem of how to incorporate species already known to occur at a given sampling scale but which lack the quantitative abundance data necessary to incorporate them into comparative studies of macroecology and taxonomic richness. However, species occurrence data in museum collections and the literature can be used to test the accuracy of existing methods of species richness estimation (Gotelli and Colwell 2001; Petersen et al. 2003). As has been shown in other studies (e.g., Petersen et al. 2003), species richness estimators such as the Chao and lognormal underestimate the known richness for the samples considered here, and by extension fall short of the total richness, which includes additional species that have yet to be sampled.

Rarity provides a plausible explanation for the observed differences in taxonomic richness between bulk samples and faunal lists. However, spatial and temporal heterogeneity in species abundance and occurrence might contribute to the observed discordance. How confident can we be that species not recovered in bulk samples were rare and not simply abundant elsewhere at a given locality? To address this question the focal scale of interest must be explicitly defined—are we interested in estimating the diversity of a bulk sample, a

TABLE 8. Rank-order correlation between species abundances in stratigraphic subsets at each locality. Rho values are the average Spearman correlation coefficient for 1000 bootstrap iterations in which samples containing N individuals were drawn from each stratigraphic subset and compared.

	Rho	p	N
Red Bluff	0.49 ± 0.09	0.001	1570
Moodys Branch	0.17 ± 0.07	0.21	2400
Gosport (lower versus middle)	0.67 ± 0.02	$\ll 0.00001$	20,160
Gosport (lower versus upper)	0.47 ± 0.03	$\ll 0.00001$	20,160
Gosport (middle versus upper)	0.59 ± 0.03	$\ll 0.00001$	20,160

biofacies exposed at a single outcrop, or a stratigraphic unit exposed regionally? The sampling domains of the bulk-sample and faunal-list data should then be compared. At broad scales, such as those considered here (e.g., an outcrop or greater), pooling bulk samples provides the best estimate of the RAD, encompassing finer-scale variation (Bennington 2003; Zuschin et al. 2006; Scarponi and Kowalewski 2007) while minimizing differences in the sampling domains of bulk samples and faunal lists.

For the three Paleogene data sets analyzed here, I examined heterogeneity at the locality scale in three ways. First, I compared the sampling domains of bulk samples and faunal lists. Second, I compared the rank-order correlation between the average abundance and sampled occurrence of species. A strong positive relationship would provide support for ''list-only'' species being rare whereas a weak relationship would suggest greater uncertainty in abundance estimates for species with few or no bulk sampled occurrences. Species found at high frequencies within a single sample may still have been rare at the outcrop scale, but a weak relationship between abundance and occurrence would suggest that more extensive sampling is needed to accurately describe the RAD at the outcrop scale. Third, I subdivided the bulk sample data into lower and upper stratigraphic bins and compared the rank-order abundances of species in these stratigraphic subsets; for the Gosport at CB a middle stratigraphic unit was also included. To minimize differences due simply to sample-size variation, I subsampled with replacement the abundance data from each stratigraphic bin and calculated the Spearman correlation coefficient between subsamples and repeated this procedure 1000 times.

The sampling domains of bulk samples and faunal lists largely overlap (Table 1), although our ability to measure this is constrained by the resolution of historical locality information. At each locality, average abundance and sampled occurrence were significantly positively correlated (Rho values from 0.85 to 0.92, all significant at the $p < 0.0001$); species restricted to few samples did not occur at high frequencies. For the middle Eocene and early Oligocene localities, species abundances were also significantly correlated among stratigraphic subdivisions (Table 8). These results are not surprising because the stratigraphic context of these deposits also suggests that outcrop-scale spatial and temporal heterogeneity should generally be low. Shellbeds formed during the Transgressive Systems Tract may subsume greater periods of time-averaging, which likely reduces local-scale spatial and temporal heterogeneity (Zuschin and Stanton 2002; Scarponi and Kowalewski 2004, 2007). At the late Eocene locality, positive correlations are observed among some individual samples but abundance data between the lower and upper portions of the outcrop are not correlated, suggesting the presence of vertical heterogeneity in species abundance and occurrence; additional bulk samples at the late Eocene locality may change abundance estimates derived from the observed RAD.

Does the Zipf-Mandelbrot RAD fit the observed abundance data sufficiently to warrant its use in modeling the abundances of ''list-only'' species? For these three localities the Zipf-Mandelbrot explains greater than 96% of the deviance in the observed abundance data and is largely an unbiased estimator of species abundances (Table 5, Fig. 4). The modeled versus observed abundance values for species fall along the line of unity for each locality (Fig.

4), though closer inspection reveals a consistent offset in the abundances of the rarest taxa (those known from ten or fewer specimens), with modeled values always slightly elevated over observed values. Although the Zipf-Mandelbrot model fails to replicate the rarest portion of the observed RAD, it is arguable that the observed abundances of these rare species are those most subject to uncertainty, and previous studies that have noted this offset have suggested not weighting these data heavily in the model-fitting procedure (Frontier 1985; Mouillot and Lepretre 2000). Furthermore, although the Zipf-Mandelbrot RAD fails to fit the observed rare tail of species, the predictive failure of the model is consistent in direction with the excess of unobserved species known from each locality. Another RAD model might fit the observed abundance data even better, but the improved fit to the rare tail of species would generate distributions inconsistent with the plethora of ''list-only'' species known from these localities. If the modeled RAD replicated the observed abundances of the rarest taxa, the abundance estimates generated for unobserved ''list-only'' species (and the sample sizes necessary to unveil them) would be implausible.

Further work is necessary to understand the causes of rarity in fossil macrobenthic assemblages. Do fossil concentrations accurately capture the long rare tail of the abundance distributions of the living communities they are derived from, or might rare ''list-only'' species have different taphonomic half-lives than bulk-sampled species and disproportionately accrue during time-averaging (Kidwell 2002)? For the three localities considered here, the frequency of species known from few specimens is within the range of that observed in Recent marine benthic surveys. For example, 30–50% of species observed in fossil bulk samples from these localities are known from five or fewer specimens. This frequency is slightly elevated in the modeled abundance data (40–60% of species) but not considerably greater than the 50% of species observed in some Recent marine benthic surveys (Bouchet et al. 2002; Zuschin and Oliver 2005).

Alternatively, were rare ''list-only'' species simply aberrant forms of existing species that would be synonimized given taxonomic revision? Although it has been argued that species-level taxonomic revisions tend to eliminate taxa rather than add them (e.g., Alroy 2002; Wagner et al. 2007), generalities are wanting. Wagner and collaborators (2007), for example, documented a net increase in the richness of Cenozoic bivalve genera following taxonomic revision and net declines in the genus-level diversity of Jurassic bivalves and Paleozoic gastropods. The growing molecular systematics literature also contains many examples in which diversity has increased following taxonomic revision (e.g., Nunes et al. 2008). For the data analyzed here, it is unlikely that further taxonomic revision would greatly reduce the discordance between bulk-sampled and faunal-list richness for these Paleogene localities, because each fauna has been the subject of previous taxonomic standardization (Palmer and Brann 1965, 1966; Dockery 1977, 1982; MacNeil and Dockery 1984) and subsequent taxonomic revisions of individual clades have both added and eliminated species (e.g., Heaslip 1968; Dockery 1977; Allmon 1996; Garvie 1996).

Alternatively, ''list-only'' species may have been ecological relicts resulting from habitat-averaging during the accrual of the death assemblage, or ephemeral members of either the living community or the resulting death assemblage via postmortem transport. Several studies of Recent communities have suggested that excess rare diversity may be explained in part by increased migration rates among species and/or autocorrelation resulting from the pooling of small spatial or temporal samples (McGill 2003). In some instances, when ''transient'' species are removed (using criteria other than rarity) the observed abundance data are best fit by the lognormal RAD, which lacks the fat tail of rare species characteristic of both Zipf-Mandelbrot and zero-sum RAD models (Gregory 2000; Magurran and Henderson 2003; Ulrich and Ollik 2004; McGill et al. 2007). Taphonomic studies comparing the age-distributions, taphonomic condition, and habitat breadth of rare versus abundant species within skeletal concentrations would help test these hypotheses and are necessary to determine if the magnitude of rare richness in

fossil assemblages is a signature of the diversity of the living community, spatial and temporal averaging, and/or differential decay.

Integrating data from museum collections, the literature, and quantitative bulk samples can unveil diversity data, allowing for more comprehensive comparative biodiversity studies. Such an approach enables us to capitalize on the historical sampling effort represented by museums and the literature. Although some workers have emphasized the importance of gathering new data for analyses of fossil diversity (Jackson and Johnson 2001; Adrain and Westrop 2003; Johnson 2003), citing in part the historical emphasis of museum collections on taxonomic richness rather than ecological structure, it is evident that relying solely on a field-based approach cannot eliminate the sampling biases affecting rare taxa and constrains studies to consideration of the dynamics of abundant and common taxa. Understanding the temporal and spatial distribution of rarity as well as its ecological and phylogenetic context is critical for evaluating patterns of extinction and recovery, and for addressing questions related to the evolutionary ecology of species and clades. If, for example, taxa are rare early or late in their histories—with respect to either spatial occurrences (Foote 2007; Foote et al. 2007; Liow and Stenseth 2007) or local densities of individuals—then bulk-sample diversity data will tend to overemphasize the stability of taxon ecology through time. At a different scale, identifying covariances between species abundance and other variables such as body size, geographic range, or trophic level may depend upon adequately sampling the full range of abundance values within a clade or community, with distributions truncated by inadequate sampling perhaps being more likely to exhibit an apparent lack of structure. Comparisons between museum, literature, and field data have typically concentrated on the misfit between these sources of data, focusing primarily on the negative effects of biases on the reconstruction of ecological and evolutionary patterns (e.g., Hunter and Donovan 2005; Davis and Pyenson 2007). However, historical archives are a rich resource for comparative analyses (Webster et al. 2003; Bieler and Mik-kelsen 2004; Allmon 2005; Guralnick and Van Cleve 2005; Krause et al. 2006), given an understanding of the directions and magnitudes of biases contained within them.

Conclusions

This paper provides two approaches for estimating the abundances of rare species by capitalizing on the extensive, and often nonrandom, sampling effort housed within museum collections and the literature. Species occurrences in museums and monographs can be integrated into quantitative diversity analyses by using the relative abundance distributions observed in quantitative samples compiled at comparable spatial and temporal scales to estimate species abundances. Rather than dismissing museum and literature data as biased records unfit for study, the direction of this bias (i.e., an emphasis on taxonomic diversity over ecological structure) facilitates the sampling of rare species that would otherwise be veiled from observation under most reasonable quantitative sampling schemes. Data for three Paleogene Gulf Coastal Plain localities demonstrate both the magnitude of local diversity that can be missed when only bulk samples are used and the substantial sampling effort that would be necessary to unveil this diversity. From the observed RAD and the cumulative faunal list a modeling approach can be used to include rare species into comparative analyses of richness. Although there are risks involved in using simple models to extend the observed RAD to include ''list-only'' species, such risks are not unique to this approach. In practice, sample-standardized richness estimates are also used as models of the relative differences in diversity among samples, localities, and biotas, with the exact form of the extrapolation not made explicit. When combined, these two approaches to estimating taxonomic richness provide useful end-members for examining patterns of biodiversity, differing in their comprehensiveness of sampling and underlying assumptions. When these methods are applied to three Paleogene data sets, the effects of the Eocene-Oligocene extinction on local richness in the Gulf Coastal Plain are less dramatic than expected from examination of synoptic regional-scale patterns (Dockery 1986b). Furthermore, these results suggest that the primary

drop in local richness in the Gulf Coastal Plain occurred earlier during the middle Eocene. Most broadly, these results stress the value of exploring novel approaches for extracting new information from old data. Even data gathered under quite different sampling regimes can be integrated to better understand patterns of diversity and distribution through time.

Acknowledgments

Literature Cited

Adrain, J. M., and S. R. Westrop. 2003. Paleobiodiversity: we need new data. Paleobiology 29:22–25.

Akaike, H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control 19:716–723.

Allmon, W. D. 1996. Systematics and evolution of Cenozoic American Turritellidae (Mollusca: Gastropoda) I: Paleocene and Eocene Coastal Plain species related to ''*Turritella mortoni* Conrad'' and ''*Turritella humerosa* Conrad.'' Palaeontographica Americana 59:1–134.

———. 2005. The importance of museum collections in paleobiology. Paleobiology 31:1–5.

Alroy, J. 2002. How many named species are valid? Proceedings of the National Academy of Sciences USA 99:3706–3711.

Alroy, J., M. Aberhan, D. J. Bottjer, M. Foote, F. T. Fursich, P. J. Harries, A. J. W. Hendy, S. M. Holland, L. C. Ivany, W. Kiessling, M. A. Kosnik, C. R. Marshall, A. J. McGowan, A. I. Miller, T. D. Olszewski, M. E. Patzkowsky, S. E. Peters, L. Villier, P. J. Wagner, N. Bonuso, P. S. Borkow, B. Brenneis, M. E. Clapham, L. M. Fall, C. A. Ferguson, V. L. Hanson, A. Z. Krug, K. M. Layou, E. H. Leckey, S. Nurnberg, C. M. Powers, J. A. Sessa, C. Simpson, A. Tomasovych, and C. C. Visaggi. 2008. Phanerozoic trends in the global diversity of marine invertebrates. Science 321:97–100.

Bennington, J. B. 2003. Transcending patchiness in the comparative analysis of paleocommunities: a test case from the Upper Cretaceous of New Jersey. Palaios 18:22–33.

Bieler, R., and P. M. Mikkelsen. 2004. Marine bivalves of the Florida Keys: a qualitative faunal analysis based on original collections, museum holdings and literature data. Malacologia 46:503–544.

Bouchet, P., P. Lozouet, P. Maestrati, and V. Heros. 2002. Assessing the magnitude of species richness in tropical marine environments: exceptionally high numbers of molluscs at a New Caledonia site. Biological Journal of the Linnean Society 75:421–436.

Brown, J. H. 1984. On the relationship between abundance and distribution of species. American Naturalist 124:255–279.

———. Macroecology. University of Chicago Press, Chicago.

Budd, A. F., and K. G. Johnson. 2001. Contrasting patterns in rare and abundant species during evolutionary turnover. Pp. 295–325 in J. B. C. Jackson, S. Lidgard, and F. K. McKinney, eds. Evolutionary patterns: growth, form, and tempo in the fossil record. University of Chicago Press, Chicago.

Bulinski, K. V. 2007. Analysis of sample-level properties along a paleoenvironmental gradient: the behavior of evenness as a function of sample size. Palaeogeography, Palaeoclimatology, Palaeoecology 253:490–508.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer, New York.

Bush, A. M., M. J. Markey, and C. R. Marshall. 2004. Removing bias from diversity curves: the effects of spatially organized biodiversity on sampling-standardization. Paleobiology 30:666–686.

Bush, A. M., R. K. Bambach, and G. M. Daley. 2007. Changes in theoretical ecospace utilization in marine fossil assemblages between the mid-Paleozoic and late Cenozoic. Paleobiology 33:76–97.

Chisholm, R. A. 2007. Sampling species abundance distributions: resolving the veil-line debate. Journal of Theoretical Biology 247:600–607.

CoBabe, E. A., and W. D. Allmon. 1994. Effects of sampling on paleoecological and taphonomic analyses in high-diversity fossil accumulations: an example from the Eocene Gosport Sand, Alabama. Lethaia 27:167–178.

Crampton, J. S., M. Foote, A. G. Beu, P. A. Maxwell, R. A. Cooper, L. Matcham, B. A. Marshall, and C. M. Jones. 2006. The ark was full! Constant to declining Cenozoic shallow marine biodiversity on an isolated midlatitude continent. Paleobiology 32:509–532.

Davidoff, A. J., and T. E. Yancey. 1993. Eustatic cyclicity in the Paleocene and Eocene: data from the Brazos River Valley, Texas. Tectonophysics 222:371–395.

Davis, E. B., and N. D. Pyenson. 2007. Diversity biases in terrestrial mammalian assemblages and quantifying the differences between museum collections and published accounts: a case study from the Miocene of Nevada. Palaeogeography, Palaeoclimatology, Palaeoecology 250:139–149.

Dewdney, A. K. 1998. A general theory of the sampling process with applications to the ''veil line.'' Theoretical Population Biology 54:294–302.

Dockery, D. T., III. 1977. Mollusca of the Moodys Branch Formation, Mississippi. Mississippi Geological Survey Bulletin 120:1–212.

———. 1982. Lower Oligocene Bivalvia of the Vicksburg Group in Mississippi. Mississippi Bureau of Geology Bulletin 123:1–261.

———. 1986a. The Cockfield (Claiborne Group), Moodys Branch and Yazoo (Jackson Group) Formations at the Riverside Park locality in Jackson, Mississippi. Geological Society of America centennial field guide, pp. 401–403. Geological Society of America, Boulder, Colo.

———. 1986b. Punctuated succession of Paleogene mollusks in the Northern Gulf Coastal Plain. Palaios 1:582–589.

Dockery, D. T., III., and P. Lozouet. 2003. Molluscan faunas across the Eocene/Oligocene boundary in the North American Gulf Coastal Plain, with comparisons to those of the Eocene and Oligocene of France. Pp. 303–340 in D. R. Prothero, L. C. Ivany, and E. A. Nesbitt, eds. From greenhouse to icehouse: the marine Eocene-Oligocene transition. Columbia University Press, New York.

Elder, S. R. 1981. Fossil assemblages of a marine transgressive sand, Moodys Branch Formation (Upper Eocene), Louisiana and Mississippi. Master's thesis. University of Texas, Austin.

Ellingsen, K. E. 2001. Biodiversity of a continental shelf soft-

sediment macrobenthos community. Marine Ecology Progress Series 218:1–15.

Foote, M. 2000. Origination and extinction components of taxonomic diversity: general problems. *In* D. H. Erwin, and S. L. Wing, eds. Deep time: paleobiology's perspective. Paleobiology 26 (Suppl. to No. 4):74–102.

———. 2007. Symmetric waxing and waning of marine invertebrate genera. Paleobiology 33:517–529.

Foote, M., J. S. Crampton, A. G. Beu, B. A. Marshall, R. A. Cooper, P. A. Maxwell, and I. Matcham. 2007. Rise and fall of species occupancy in Cenozoic fossil mollusks. Science 318:1131–1134.

Frontier, S. 1985. Diversity and structure in aquatic ecosystems. Oceanography and Marine Biology: An Annual Review 23:253–312.

Garvie, C. L. 1996. The molluscan macrofauna of the Reklaw Formation, Marquez Member (Eocene: Lower Claibornian), in Texas. Bulletins of American Paleontology 111:1–177.

Gaston, K. J. 1994. Rarity. Chapman and Hall, London.

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

Gaston, K. J., T. M. Blackburn, and J. H. Lawton. 1997. Interspecific abundance-range size relationships: an appraisal of mechanisms. Journal of Animal Ecology 66:579–601.

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

Gregory, R. D. 2000. Abundance patterns of European breeding birds. Ecography 23:201–208.

Guralnick, R., and J. Van Cleve. 2005. Strengths and weaknesses of museum and national survey data sets for predicting regional species richness: comparative and combined approaches. Diversity and Distributions 11:349–359.

Hansen, T. A., P. H. Kelley, and D. M. Haasl. 2004. Paleoecological patterns in molluscan extinctions and recoveries: comparison of the Cretaceous-Paleogene and Eocene-Oligocene extinctions in North America. Palaeogeography, Palaeoclimatology, Palaeoecology 214:233–242.

Harrison, H. C. 1994. Effects of environmental changes on molluscan evolutionary patterns, Gosport Sand (Middle Eocene), Southwest Alabama. Master's thesis. University of South Florida, Tampa.

Hayek, L.-A. C., and M. A. Buzas. 1997. Surveying natural populations. Columbia University Press, New York.

Heaslip, W. G. 1968. Cenozoic evolution of the alticostate venericards in Gulf and East Coastal North America. Palaeontographica Americana 6:51–135.

Holland, S. M., and M. E. Patzkowsky. 2002. Stratigraphic variation in the timing of first and last occurrences. Palaios 17:134–146.

Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton, N.J.

Hunter, A. W., and S. K. Donovan. 2005. Field sampling bias, museum collections and completeness of the fossil record. Lethaia 38:305–314.

Hurlburt, S. H. 1971. The nonconcept of species diversity: a critique and alternative parameters. Ecology 52:577–86.

Ivany, L. C. 1998. Sequence stratigraphy of the Middle Eocene Claiborne Stage, US Gulf Coastal Plain. Southeastern Geology 38:1–20.

Jackson, J. B. C., and D. H. Erwin. 2006. What can we learn about ecology and evolution from the fossil record? Trends in Ecology and Evolution 21:322–328.

Jackson, J. B. C., and K. G. Johnson. 2001. Measuring past biodiversity. Science 293:2401, 2403–2404.

Jaramillo, C. A., and F. E. Oboh-Ikuenobe. 1999. Sequence stratigraphic interpretations from palynofacies, dinocyst and lithological data of Upper Eocene-Lower Oligocene strata in southern Mississippi and Alabama, US Gulf Coast. Palaeogeography, Palaeoclimatology, Palaeoecology 145:259–302.

Johnson, K. G. 2003. New data for old questions. Paleobiology 29:19–21.

Kidwell, S. M. 2001. Preservation of species abundance in marine death assemblages. Science 294:1091–1094.

———. 2002. Time-averaged molluscan death assemblages: palimpsests of richness, snapshots of abundance. Geology 30:803–806.

Kiessling, W., and R. C. Baron-Szabo. 2004. Extinction and recovery patterns of scleractinian corals at the Cretaceous-Tertiary boundary. Palaeogeography, Palaeoclimatology, Palaeoecology 214:195–223.

Kosnik, M. A., and P. J. Wagner. 2006. Effects of taxon abundance distributions on expected numbers of sampled taxa. Evolutionary Ecology Research 8:195–211.

Kowalewski, M., M. Carroll, L. Casazza, N. Gupta, B. Hannisdal, A. Hendy, R. A. Krause Jr., M. Labarbera, D. G. Lazo, C. Messina, S. Puchalski, T. A. Rothfus, J. Sälgeback, J. Stempien, R. C. Terry, and A. Tomašových. 2003. Quantitative fidelity of brachiopod-mollusk assemblages from modern subtidal environments of San Juan Islands, USA. Journal of Taphonomy 1:43–65.

Kowalewski, M., W. Kiessling, M. Aberhan, F. T. Fürsich, D. Scarponi, S. L. B. Wood, and A. P. Hoffmeister. 2006. Ecological, taxonomic, and taphonomic components of the post-Paleozoic increase in sample-level species diversity of marine benthos. Paleobiology 32:533–561.

Krause, R. A., Jr., J. A. Stempien, M. Kowalewski, and A. I. Miller. 2006. Body size estimates from the literature: utility and potential for macroevolutionary studies. Palaios 22:60–73.

Kunin, W. E., and K. J. Gaston, eds. 1997. The biology of rarity. Chapman and Hall, London.

Liow, L. H., and N. C. Stenseth. 2007. The rise and fall of species: implications for macroevolutionary and macroecological studies. Proceedings of the Royal Society of London B 274:2745–2752.

Lockwood, R. 2003. Abundance not linked to survival across the end-Cretaceous mass extinction: patterns in North American bivalves. Proceedings of the National Academy of Sciences USA 100:2478–2482.

Lockwood, R., and L. R. Chastant. 2006. Quantifying taphonomic bias of compositional fidelity, species richness, and rank abundance in molluscan death assemblages from the upper Chesapeake Bay. Palaios 21:376–383.

MacNeil, F. S., and D. T. Dockery. 1984. Lower Oligocene Gastropoda, Scaphopoda, and Cephalopoda of the Vicksburg Group in Mississippi. Mississippi Bureau of Geology Bulletin 124:1–415.

Magurran, A. E., and P. A. Henderson. 2003. Explaining the excess of rare species in natural species abundance distributions. Nature 422:714–716.

Mancini, E. A., and B. H. Tew. 1992. Paleogene unconformity-bounded depositional sequences of southwest Alabama: lithofacies, systems tracts, and sequence boundaries. Alabama Geological Society Guidebook 29:1–72.

Marquet, P. A., J. A. Keymer, and H. Cofre. 2003. Breaking the stick in space: of niche models, metacommunities and patterns in the relative abundance of species. Pp. 64–86 *in* T. M. Blackburn and K. J. Gaston, eds. Macroecology: concepts and consequences. Blackwell Science, Oxford.

McCullagh, P., and J. A. Nelder. 1989. Generalized linear models. Chapman and Hall, London.

McGill, B. J. 2003. Does Mother Nature really prefer rare species or are log-left-skewed SADs a sampling artefact? Ecology Letters 6:766–773.

McGill, B. J., R. S. Etienne, J. S. Gray, D. Alonso, M. J. Anderson, H. K. Benecha, M. Dornelas, B. J. Enquist, J. L. Green, F. He,

A. H. Hurlbert, A. E. Magurran, P. A. Marquet, B. A. Maurer, A. Ostling, C. U. Soykan, K. I. Ugland, and E. P. White. 2007. Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. Ecology Letters 10:995–1015.

Meldahl, K. H. 1990. Sampling, species abundance, and the stratigraphic signature of mass extinction: a test using Holocene tidal flat mollusks. Geology 18:890–893.

Mouillot, D., and A. Lepretre. 1999. A comparison of species diversity estimators. Researches on Population Ecology 41:203–215.

———. 2000. Introduction of relative abundance distribution (RAD) indices, estimated from the rank-frequency diagrams (RD), to assess changes in community diversity. Environmental Monitoring and Assessment 63:279–295.

Novack-Gottshall, P. M. 2007. Using a theoretical ecospace to quantify the ecological diversity of Paleozoic and modern marine biotas. Paleobiology 33:273–294.

Nunes, F., H. Fukami, S. V. Vollmer, R. D. Norris, and N. Knowlton. 2008. Re-evaluation of the systematics of the endemic corals of Brazil by molecular data. Coral Reefs 27:423–432.

Oksanen, J., R. Kindt, P. Legendre, and R. B. O'Hara. 2007. Vegan: R functions for vegetation ecologists, Version 1. 8.5. URL: http://cc.oulu.fi/~jarioksa/.

Olszewski, T. D. 2004. A unified mathematical framework for the measurement of richness and evenness within and among multiple communities. Oikos 104:377–387.

Olszewski, T. D., and D. H. Erwin. 2004. Dynamic response of Permian brachiopod communities to long-term environmental change. Nature 428:738–741.

Palmer, K. V. W., and D. C. Brann. 1965. Catalogue of the Paleocene and Eocene Mollusca of the southern and eastern United States, Part I. Pelecypoda, Amphineura, Pteropoda, Scaphopoda, and Cepholopoda. Bulletins of American Paleontology 48(218):1–466.

———. 1966. Catalogue of the Paleocene and Eocene Mollusca of the southern and eastern United States, Part II. Gastropoda. Bulletins of American Paleontology 48:467–1057.

Pandolfi, J. M. 1996. Limited membership in Pleistocene reef coral assemblages from the Huon Peninsula, Papua New Guinea: constancy during global change. Paleobiology 22:152–176.

Pasley, M. A., and J. E. Hazel. 1995. Revised sequence stratigraphic interpretation of the Eocene-Oligocene boundary interval, Mississippi and Alabama, Gulf-Coast Basin, USA. Journal of Sedimentary Research 65:160–169.

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

———. 2006. Genus richness in Cambrian-Ordovician benthic marine communities in North America. Palaios 21:580–587.

Petersen, F. T., R. Meier, and M. Nykjaer. 2003. Testing species richness estimation methods using museum label data on the Danish Asilidae. Biodiversity and Conservation 12:687–701.

Powell, M. G., and M. Kowalewski. 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, F. W. 1948. The commonness, and rarity, of species. Ecology 29:254–283.

R Development Core Team. 2006. R: a language and environment for statistical computing, Version 2. 3.1. R Foundation for Statistical Computing, Vienna.

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

Sanders, H. L. 1968. Marine benthic diversity—a comparative study. American Naturalist 102:243–282.

Scarponi, D., and M. Kowalewski. 2004. Stratigraphic paleoecology: bathymetric signatures and sequence overprint of mollusk associations from upper Quaternary sequences of the Po Plain, Italy. Geology 32:989–992.

———. 2007. Sequence stratigraphic anatomy of diversity patterns: Late Quaternary benthic mollusks of the Po Plain, Italy. Palaios 22:296–305.

Schlacher, T. A., P. Newell, J. Clavier, M. A. Schlacher-Hoenlinger, C. Chevillon, and J. Britton. 1998. Soft-sediment benthic community structure in a coral reef lagoon—the prominence of spatial heterogeneity and 'spot endemism.' Marine Ecology Progress Series 174:159–174.

Shin, P. K. S., and K. E. Ellingsen. 2004. Spatial patterns of soft-sediment benthic diversity in subtropical Hong Kong waters. Marine Ecology Progress Series 276:25–35.

Signor, P. W., III., and J. H. Lipps. 1982. Sampling bias, gradual extinction patterns, and catastrophes in the fossil record. In L. T. Silver and P. H. Schultz, eds. Geological implications of impacts of large asteroid and comets on the earth. Geological Society of America Special Paper 190:291–296.

Swindel, D. B. 1986. A paleoecological study of the Gosport Sand (Claibornian: Middle Eocene) in southwest Alabama. Master's thesis. University of Alabama, Tuscaloosa.

Tew, B. H., and E. A. Mancini. 1995. An integrated stratigraphic method for paleogeographic reconstruction: examples from the Jackson and Vicksburg Groups of the eastern Gulf Coastal Plain. Palaios 10:133–153.

Thompson, W. L., ed. 2004. Sampling rare or elusive species. Island Press, Washington, D.C.

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

Ulrich, W., and M. Ollik. 2004. Frequent and occasional species and the shape of relative-abundance distributions. Diversity and Distributions 10:263–269.

Wagner, P. J., M. A. Kosnik, and S. Lidgard. 2006. Abundance distributions imply elevated complexity of post-Paleozoic marine ecosystems. Science 314:1289–1292.

Wagner, P. J., M. Aberhan, A. Hendy, and W. Kiessling. 2007. The effects of taxonomic standardization on sampling-standardized estimates of historical diversity. Proceedings of the Royal Society of London B 274:439–444.

Webster, M., P. M. Sadler, M. A. Kooser, and E. Fowler. 2003. Combining stratigraphic sections and museum collections to increase biostratigraphic resolution. Pp. 95–128 in P. J. Harries, ed. Approaches in high-resolution stratigraphic paleontology. Kluwer Academic, Amsterdam.

Zuschin, M., and P. G. Oliver. 2005. Diversity patterns of bivalves in a coral dominated shallow-water bay in the northern Red Sea—high species richness on a local scale. Marine Biology Research 1:396–410.

Zuschin, M., and R. J. Stanton. 2002. Paleocommunity reconstruction from shell beds: a case study from the Main Glauconite Bed, Eocene, Texas. Palaios 17:602–614.

Zuschin, M., M. Harzhauser, and K. Sauermoser. 2006. Patchiness of local species richness and its implication for large-scale diversity patterns: an example from the middle Miocene of the Paratethys. Lethaia 39:65–88.