Taxonomic Level as a Determinant of the Shape of the Phanerozoic Marine Biodiversity Curve

Abigail Lane and Michael J. Benton[*]

Department of Earth Sciences, University of Bristol, Wills Memorial Building, Queens Road, Bristol BS8 1RJ, United Kingdom

ABSTRACT: Key aims of recent paleobiological research have been the construction of Phanerozoic global biodiversity patterns and the formulation of models and mechanisms of diversification describing such patterns. Two conflicting theories of global diversification have been equilibrium versus expansionist growth of taxonomic diversity. These models, however, rely on accurate empirical data curves, and it is not clear to what extent the taxonomic level at which the data are analyzed controls the resulting pattern. Global Phanerozoic marine diversity curves are constructed at ordinal, familial, and generic levels using several fossil-range data sets. The fit of a single logistic model reduces from ordinal through familial to generic level, while conversely, that of an exponential growth model increases. Three sequential logistic equations, fitted to three time periods during which diversity appears to approach or reach an equilibrium state, provide the best description of the data at familial and generic levels. However, an exponential growth curve describes the diversification of marine life since the end-Permian extinction equally as well as a logistic. A species-level model of global Phanerozoic marine diversification, constructed by extrapolation of the trends from familial to generic level, suggests growth in numbers of marine species was broadly exponential. When smaller subsets of the data are analyzed, the effect of taxonomic level on the shape of the diversity curve becomes more pronounced. In the absence of species data, a consistent signal at more than one higher taxonomic level is required to predict a species-level pattern.

Keywords: diversity, marine, Phanerozoic, model.

There have been many attempts to model long-term global diversity patterns in the hope of understanding their biological underpinning; debate has focused in particular on whether life has followed a growth pattern that is expan-

* E-mail: mike.benton@bristol.ac.uk.

sionist (e.g., Valentine 1969; Walker and Valentine 1984; Signor 1985; Benton 1995, 1997) or equilibrial (e.g., Raup 1972; Carr and Kitchell 1980; Alroy 1998). Most prominent have been the logistic models of Sepkoski (1978, 1979, 1984) culminating in the three-phase, coupled logistic model used to describe Phanerozoic diversification at the familial level. This global hypothesis was derived from earlier attempts to model colonialization dynamics, competition, and evolution on newly emergent islands (MacArthur and Wilson 1967). Equilibrium models of diversification assume that there are limits to the numbers of species and higher taxa sustainable by the earth, limits that have been reached in the past and will be reached in the future. An empirical diversity curve for marine orders fits a logistic equilibrium model well (Sepkoski 1978) and agrees with Raup's (1972) earlier predictions that species diversity through the Phanerozoic increased dramatically in the Cambro-Ordovician, with little subsequent growth for the remainder of the Phanerozoic. Such a single, long diversity equilibrium is also hinted at in recent attempts to uncover the pattern of generic marine Phanerozoic biodiversity using sampling standardization techniques (Alroy et al. 2001). When using familial data, Sepkoski (1981) identified three marine "evolutionary faunas" that, despite sharing origination times in the early Cambrian and diversifying in parallel, still sequentially dominated different periods of the Phanerozoic. Each of these faunas was modeled logistically, with first growth and then stabilization of diversity, followed by decline as another increased in dominance. Hence, turnover of each fauna was affected by the diversity levels of the other two, and while mass extinctions were a feature of the model, they did not fundamentally alter the diversification dynamics of the system. The three interacting faunas were modeled in such a way that, periodically, total diversity appeared to reach or approach an equilibrium level. During these time periods, diversity fluctuated but did not significantly increase or decrease. Taxonomic diversity during the Meso-Cenozoic, dominated by the "'modern" evolutionary fauna, has yet to show signs of reaching such an equilibrium.

A simpler model was more recently proposed by Cour-

tillot and Gaudemer (1996). Instead of modeling individual phases, they fit logistic curves to total marine diversity over discrete periods of the Phanerozoic. Such a fit to total diversity was considered more robust than modeling separate faunal groups. Courtillot and Gaudemer, in contrast to Sepkoski, modeled mass extinction events as "resetting" diversification parameters by the creation of a reduced pool of parent species. They used these events to define boundaries between sequential sections of the curve.

Two fresh attempts at modeling Phanerozoic marine diversity are presented here. First, single equations, both logistic and exponential, are fit to the entirety of the Phanerozoic marine diversity curve. Second, fits of single equations are made to three sequential sections of the global curve. These sections correspond to the three time periods during which diversity levels produced by faunal interactions neared or attained a seeming equilibrium. Hence, this modeling is closer in essence to that of Courtillot and Gaudemer (1996), with mass extinctions resetting the diversification parameters of the Phanerozoic marine diversity system. The current attempt aims to achieve two objectives: first, to describe the empirical data in the most parsimonious way, that is, using the least number of parameters, and to test whether such a description is best achieved using exponential or logistic growth equations or a combination of both; second, to investigate whether such a description is consistent at differing taxonomic levels.

The effect of taxonomic level on apparent diversity patterns has been highlighted in the past (e.g., Raup 1972; Erwin et al. 1987). There are several problems associated with choosing the level at which to conduct a biodiversity study. Higher taxa are less affected by the vagaries of sampling in the fossil record and are thought to provide more robust diversity patterns, but by their long-term nature, they may mask much of the shorter-term or more subtle patterns. Bambach (1989) predicted that genera and families, constructed as assemblages of species with similar morphologies, are useful for diversity studies, while orders, classes, and phyla are defined by key characters and therefore will exhibit diversity histories decoupled from those of species. The majority of recent global paleobiodiversity analysis has been conducted at the level of the family or genus. It can be argued that such higher taxa are biological entities in their own right, in the same way as species, and therefore valid as units of diversity. Each consists of a set of species specialized for a particular mode of life or niche. If this set of niches as a group represents some ecological subdivision, then the higher taxon can be seen as a biological entity occupying that subdivision. Plotting curves at differing taxonomic levels hence illustrates different aspects of biological diversification. Some workers embrace this interpretation (e.g., Van Valen 1984; Bottjer and Jablonski 1988; Valentine 1990), arguing that both mono-

phyletic and paraphyletic higher taxa are biologically meaningful entities. It is often the case, however, that when higher taxa are used in biodiversity studies, it is as proxies for species, and it is the pattern of species diversification through time that we are attempting to discover (Sepkoski 1978). Unfortunately, species data are only available for the most localized of biodiversity analyses. The large number of fossil species that have been described means there is little immediate hope of any global species taxonomic compilation covering the entire Phanerozoic being attempted. Paradoxically, a more fundamental problem in the use of fossil species is the incompleteness of their record. Raup (1995) estimated that only one percent of all the species that have ever existed has been preserved, discovered, and described.

Plotting global paleobiodiversity curves at the level of genera, families, or even orders is therefore necessary. If the distribution of species among the higher taxa is skewed, it is unlikely that a plot at the familial level will accurately match the species-level pattern. Computer simulations (Sepkoski 1978; Sepkoski and Kendrick 1993) have predicted an adequate association between diversity counts at differing taxonomic levels, while a study comparing diversity patterns of extant eastern Pacific molluscs at the level of genus and family with those at species level found a good correlation (Roy et al. 1996). The fossil evidence, however, is equivocal. Sepkoski (1997) claimed a strong overall similarity between the global diversity histories of fossil marine families and genera, although he noted differences in the detail of the curves. McCormick and Owen (2001) noted good correlations between the diversity curves of Ordovician trilobite genera and species from the Welsh Basin. However, a recent study of species richness and higher taxon diversity in both living and fossil plant communities (Enquist et al. 2002) found that although there was a tight relationship between the number of species and the number of genera and families within the samples, this relationship was not one-to-one but a power function. As species richness increased, so generic and familial diversity increased but at a slower rate. There is also the question of timing in the origins of species and higher taxa. Raup (1983) observed that higher taxa should be expected to appear early in the fossil record, while lesser-ranked taxa, genera, and species can appear at any point throughout a group's history. As a consequence, standing diversity at higher taxonomic levels tends to be stable or to decline through time following an initial radiation. This pattern may be totally decoupled from that at species level, which may require a different model of diversification from patterns derived from data at higher taxonomic ranks (Kitchell and Carr 1985).

As a result of these problems, the appropriateness of families as the units of taxon counting has been ques-

tioned. Flessa and Jablonski (1985) noted the regular increase in the number of species per family since the Silurian and concluded that it is this taxonomic distribution, rather than any ecological controls, that causes the decline in familial extinction rates though the Phanerozoic. Similarly, Benton (1995, 1997, 2001) proposed that the multiple equilibria pattern of marine diversification seen in the fossil record is an artifact of the taxonomic level at which the data have been plotted, and he predicted that the equilibria seen at the familial level, in particular the long Paleozoic diversity equilibrium or "plateau," will break down at the lower taxonomic levels of genera and species. Benton noted that the familial diversity plot for nonmarine life has a strongly exponential form and suggested that the real pattern of marine diversification at species level may be similar. The fundamental exponential pattern may be revealed for continental organisms at familial level because such families, especially those in the Paleozoic and Mesozoic, contain fewer species than marine families of the same time. Large marine families could mask the species-level pattern. This idea requires further testing.

While it is currently not possible to do more than model diversity at species level on a global scale, there are now several large data sets containing fossil taxonomic range information, the manipulation of which allows diversity curves to be plotted and compared for many different taxonomic and ecologic groups at various taxonomic levels and over differing time scales. Here, simple logistic and exponential growth equations are used in an attempt to describe the shape of the Phanerozoic global marine diversity curve when plot using orders, families, and genera, in order to examine the relative robustness of equilibrium and expansionist models. In this way, we can aim to recognize the relationship between taxonomic level and the perceived diversity pattern of any particular group. A clear understanding of past diversification patterns is essential to provide the context required for the assessment of current biodiversity and the effect of human activity on the biosphere, in particular on the rate of taxonomic disappearance as compared with past events, and the likely response of global diversity to ongoing environmental change.

Methods

Use of Global Taxonomic Databases

The use of large, computer-held collections of fossil taxonomic occurrence data has become standard in the study of paleobiodiversity dynamics (see Benton 1999 for review), whether the focus is taxonomic radiation periods (e.g., Miller and Foote 1996; McCormick and Owen 2001),

extinction events (e.g., Raup and Sepkoski 1982, 1984; Courtillot and Gaudemer 1996; Sepkoski 1996), or long-term trends within diversity patterns (e.g., Sepkoski 1978, 1979, 1981, 1984; Benton 1995, 1997). Three published compendia of marine fossil animal families (Sepkoski 1992), fossil animal and plant families from all habitats (Benton 1993), and marine genera (Sepkoski 2002) are widely used (e.g., Courtillot and Gaudemer 1996; Sepkoski 1997; Adrain and Westrop 2000; Foote 2000a, 2000b, 2001). Such data collections consist of first and last appearance dates of fossil taxa. From this information, taxonomic time ranges can be calculated, leading to estimates of standing diversity and rates of turnover through time. In recent years, the ease of manipulation of such large data sets has been facilitated by the increasing power of relational database software. The data, however, continue to suffer from the problems of poor taxonomy, uneven stratigraphic resolution, and variations in sampling intensity (Johnson and McCormick 1999), requiring caution to be exercised when interpreting results. A new database currently under construction aims to go some way toward correcting these errors, in particular those of sampling intensity (Alroy 2000; Alroy et al. 2001). The time required for this resource to reach the level of coverage of the Sepkoski and Benton data sets is still undetermined. Until then, the use of the currently available global first and last appearance taxonomic data compilations remains the best way of uncovering the patterns of global biodiversity dynamics.

To investigate the nature of Phanerozoic diversity curves, a relational database system has been created to hold several large taxonomic data sets alongside stratigraphic information. Taxonomic tables hold ordinal, familial, and generic data from Sepkoski's compendium of fossil marine animal families (Sepkoski 1992) and of fossil marine genera (Sepkoski 2002) and familial and a limited amount of ordinal data from the Fossil Record 2 data set (Benton 1993). Stratigraphic tables hold geological interval names and codes as used by the taxonomic data sets, along with chronological information. The dates of Harland et al. (1990) have been updated using the Geological Society of America 1999 Geologic Time Scale (http://rock.geosociety.org/science/timescale/timescl.pdf) and recent published dates for Cambrian stratigraphy (Bowring et al. 1993; Landing 1994; Shergold 1995; Brasier and Sukhov 1998; Davidek et al. 1998; Landing et al. 1998; Palmer 1998).

Marine diversity counts were extracted for stratigraphic intervals through the Phanerozoic at the lowest resolution allowable by the data sets. This resulted in data resolved to either series or stage level (mean duration of intervals: Sepkoski data sets, 6.57 million years; Benton data set, 7.05 million years). Diversity plots have been constructed

using ordinal, familial, and generic data. The diversity of a time interval is the total number of taxa that are present in that interval as either an actual fossil occurrence or an implied presence using range-through data. Each diversity count is plotted at the midpoint of the interval. Total diversity counts and counts excluding uncertain taxonomic assignments and single-interval taxa were produced. Taxa that are known from only a single geological interval (singletons) have a number of undesirable properties that distort diversity patterns (Sepkoski 1993, 1996; Foote and Raup 1996; Foote 2000a, 2000b).

Curve Fitting

To test the applicability of the exponential and logistic models, least squares fits of each of the equations were applied to the plots of marine life. The exponential model is simply a growth curve with a constant rate of multiplication of species or higher taxa through time. Diversity as a function of time (Sepkoski 1978) is

$$D = D_0 \times \exp{(r_d t)},$$

where D = standing diversity, D_0 = initial diversity level at time $t = 0$, t = time, and r_d = rate of diversification (per lineage million years), that is, origination-extinction rate.

The logistic model is an exponential equation with an extra term, the equilibrium parameter, added to curb growth of the system. Diversity as a function of time (Sepkoski 1978) is

$$D = \frac{D_0 D_{eq}}{D_0 + (D_{eq} - D_0) \times \exp{(-r_0 t)}},$$

where r_0 = initial rate of diversification and D_{eq} = equilibrium diversity. These equations were fitted to diversity counts, excluding uncertain taxa and singletons, from the final Precambrian interval up until the present.

In addition, three logistic equations were sequentially fit to three time intervals: the Cambrian, the post-Cambrian Paleozoic, and the Meso-Cenozoic. These correspond to the three periods of the Phanerozoic during which one of the evolutionary faunas dominates and during which total diversity reaches or is predicted to reach equilibrium. The time ranges differed slightly depending on the time scheme of the data set used: Cambrian interval (Fossil Record 2 stratigraphic scheme: Vendian–Tremadoc; Sepkoski stratigraphic scheme: Vendian–Dresbachian), post-Cambrian Paleozoic interval (Fossil Record 2 stratigraphic scheme: Tremadoc–Gzelian; Sepkoski stratigraphic scheme: Dresbachian–Stephanian), and Meso-Cenozoic interval (Fossil Record 2 stratigraphic scheme: Scythian–

Holocene; Sepkoski stratigraphic scheme: Induan–Holocene). Finally, an exponential equation was also fitted to the Meso-Cenozoic interval of the curves to investigate the possibility that diversification of marine life since the Permian has been exponential rather than logistic.

For each fit, the value of the D_0 parameter (initial diversity) was constrained to equal the diversity at the starting point of the curve or phase or set to 1 if a group's Precambrian diversity was 0, assuming that at least one ancestor was present. The other parameters were calculated according to the best fit of the model to the data (cf. Sepkoski 1984; Courtillot and Gaudemer 1996). To demonstrate the goodness-of-fit of the model to the data, R^2 values (coefficient of determination), ANOVA F ratios (ratio of the regression variance to the residual variance), and P values (probability that the inferred association between dependent and independent variables based on F is incorrect) were calculated for each regression.

Modeling Marine Diversity at Species Level

It has been proposed that the apparent logistic pattern of familial marine diversity, in particular the post-Cambrian Paleozoic equilibrium, gradually changes to an exponential curve when data is plotted at progressively lower taxonomic levels (Benton 1995, 1997, 2001). This suggestion was based on empirical family and genus data and a model of species diversity produced by Signor (1985) using various paleontological sampling measures. Signor's model was independent of taxonomy. An alternative model is constructed here using the differences in the apparent diversity pattern between the data at familial and generic levels to infer the pattern at species level. The proportional differences between each of the three parameters (D_0, r_0, D_{eq}) of the model fit to the three diversity phases at the familial level and each of those at the generic level were calculated. The generic parameters were then altered by these same proportions to derive parameters describing a hypothetical species pattern. Sequential plots of three logistic-equation solutions using these derived parameters give a crude approximation of a species-level curve. The familial- and generic-level parameters used to generate the species model are from the curve fitting as applied to the two Sepkoski data sets. The Fossil Record 2 data set has only five data points in the Cambrian phase; therefore, the logistic fit to this period is unreliable. Also, it can be assumed that the two Sepkoski data sets were compiled using similar data collection and recording practices. They also use the same stratigraphic scheme. Therefore, the change in diversity pattern from family to genus level is more likely to be a real phenomenon rather than a result of differences in data compilation methods.

This simple model of species diversity assumes that the

change in pattern between familial and generic data is a real feature of diversity at differing taxonomic levels and that the change is systematic and of the same form and magnitude as that existing between generic and species levels. It also assumes that the taxonomic practices that dictate grouping of genera into families are similar to those governing the grouping of species into genera. Both of these assumptions require testing with empirical data sets. Bambach (1989) argued that genera and families, being assembled as groups of species with similar morphologies, will exhibit similar diversity patterns as opposed to orders, classes, and phyla, which are defined by key characters or character complexes. This suggests that the change in diversity pattern from familial level to generic can be applied from genera to species, the unit of each level being merely an increasingly large grouping of the base unit of species. If, however, the degradation in the logistic pattern evident between the familial and generic curves is a feature of poorer sampling at the generic level, then the species curve will simply model a similar degradation of sampling between genera and species data rather than any real diversity signal. The species curve presented here is a hypothesis requiring further testing.

Results

Marine Biodiversity Curves and Fitted Models

Figure 1 shows Phanerozoic global marine diversity curves plotted at ordinal, familial, and generic levels. Single exponential and logistic models are fit to all curves. Table 1 summarizes the models and fit statistics. A single logistic equation provides a good description of the data at ordinal level. At familial level, however, the model becomes inadequate since the long, ordinal-level diversity plateau from the Ordovician to the Recent is not evident, a difference which led Sepkoski (1979, 1984) to develop his coupled-logistic model, incorporating the concept of evolutionary faunas. Similarly, at generic level, a simple logistic model is a poor fit. However, the exponential model displays the opposite trend. At ordinal level, there is no fit; the same is true for the Sepkoski family data, and a poor fit is evident for the Fossil Record 2 data. However, an exponential equation describes more than 50% of the variation in the generic data. This increase in the fit of the exponential model moving down taxonomic levels is due to the deterioration of the equilibria seen in the familial and ordinal data. At generic level, the Paleozoic diversity equilibrium period is still evident, but it is much reduced when compared with the later Mesozoic expansion and difficult to discern through the greater perturbations in generic diversity data (Sepkoski 1997). The general shape of the generic curve, in particular degradation
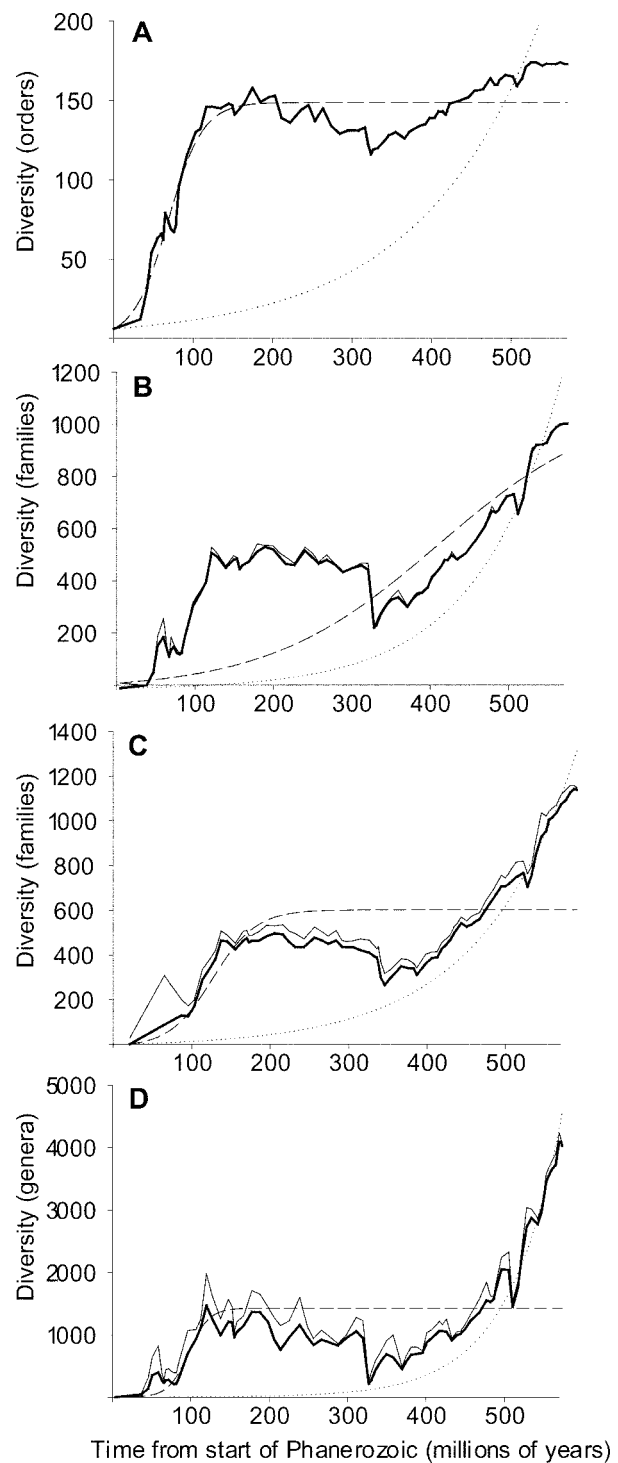

Figure 1: Phanerozoic marine diversity curves with single exponential and logistic models fitted. *A*, Orders. Data from Sepkoski (1992). *B*, Families. Data from Sepkoski (1992). *C*, Families. Data from Benton (1993). *D*, Genera. Data from Sepkoski (2002). Solid lines exclude uncertain taxa and/or singletons. Dotted lines indicate the fit of the exponential model, and dashed lines indicate the fit of the logistic model.

Table 1: Summary of data, fitted models, and fit statistics for the curves shown in figure 1

Data set, taxonomic level, and model	R^2	F	P
Sepkoski 1992:			
Order:			
Exponential	0	...	...
Logistic	.83	200	<.0001
Family:			
Exponential	0	...	...
Logistic	.33	19	<.0001
Benton 1993:			
Family:			
Exponential	.2	19	<.0001
Logistic	.24	12	<.0001
Sepkoski 2002:			
Genus:			
Exponential	.54	93	<.0001
Logistic	.21	10	<.0001

Note: R^2 = coefficient of determination, F = ratio of regression variance to residual variance, P = probability that the association between variables is incorrect.

of the Paleozoic plateau, is similar to that evident in Foote's (2000*a*) marine Phanerozoic generic curve produced using the "boundary crosser" method of summing diversity.

These results demonstrate that neither an exponential nor a single logistic model adequately explains the variation in marine diversity seen at familial and generic levels. Figure 2 shows the marine diversity curves with three sequential logistic equations fitted to the three "equilibrium" or "approaching-equilibrium" intervals of the Phanerozoic. For simplicity, only the diversity data with uncertain taxa and singletons removed have been plotted. Table 2 contains a summary of the statistical results. These best fits of the logistic model to the three sections of the diversity curves have been used to determine the free parameters of the equations. Table 3 contains the parameter values for all the logistic fits shown in figure 2. The initial diversity parameter, D_0, is in each case constrained to equal the diversity level of the system at the start of each of the three time intervals; r_0 and D_{eq} are the two free parameters derived from the model fit.

Three sequential fits of a logistic equation provide a good description of marine diversification at familial and generic level, with R^2 values ranging from 0.72 to 0.99. Both the Sepkoski (1992) and Benton (1993) data curves (fig. 3*B*, 3*C*) display an excellent fit for the post-Cambrian Paleozoic and the Meso-Cenozoic intervals, with R^2 values >0.92. The fit to the Cambrian interval of the Sepkoski (1992) curve is less convincing ($R^2 = 0.79$), but that of the Fossil Record 2 data is good ($R^2 = 0.96$). However, the Fossil Record 2 data include only five Cambrian data
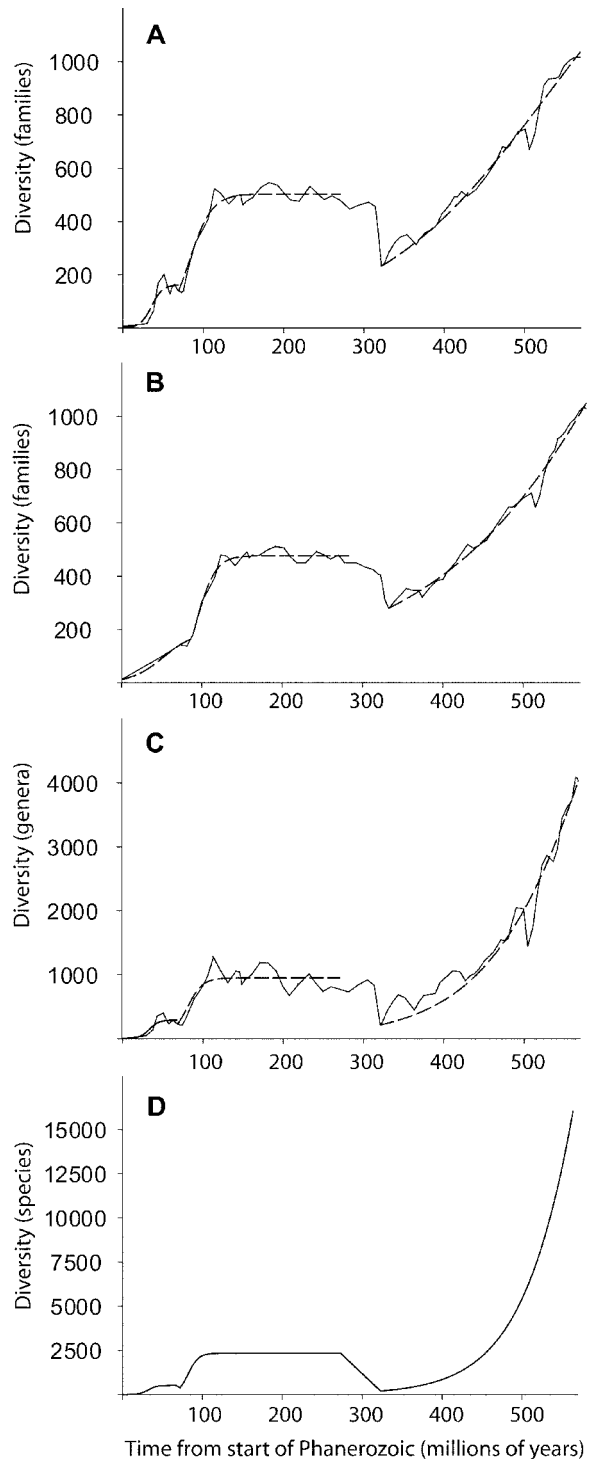

Figure 2: Empirical and model Phanerozoic marine diversity curves. Empirical curves have three sequential logistic equations fitted to the three time periods of the Phanerozoic identified in the text. *A*, Families. Data from Sepkoski (1992). *B*, Families. Data from Benton (1993). *C*, Genera. Data from Sepkoski (2002). Data excludes uncertain taxa and/ or singletons. Dashed lines indicate the fit of the equations. *D*, Species curve modeled by extrapolating the differences between the derived parameters of the familial and generic curves.

Table 2: Summary of data, fitted models, and fit statistics for empirical data curves shown in figure 2

Data set, taxonomic level, and time period of fit	No. of data points	R^2	F	P
Sepkoski 1992:				
Family:				
Cambrian	10	.79	13	.0043
Later Paleozoic	26	.95	238	<.0001
Meso-Cenozoic	42	.98	1,021	<.0001
Benton 1993:				
Family:				
Cambrian	5	.96	25	.0392
Later Paleozoic	24	.92	125	<.0001
Meso-Cenozoic	44	.99	1,675	<.0001
Sepkoski 2002:				
Genus:				
Cambrian	10	.72	9	.0118
Later Paleozoic	26	.74	33	<.0001
Meso-Cenozoic	42	.96	441	<.0001

Note: R^2 = coefficient of determination, F = ratio of regression variance to residual variance, P = probability that the association between variables is incorrect.

points, reflected in a higher P value, which indicates less confidence in the fit.

The logistic model is an adequate description of generic marine diversification (fig. 3*C*) in the Cambrian and later Paleozoic ($R^2 < 0.75$) and is an excellent fit to the Meso-Cenozoic phase ($R^2 = 0.96$). The reason for the poorer fit of the logistic model to the Paleozoic interval at generic level is due first to the greater magnitude of perturbations around the equilibrium and, second, to a gradual decline in generic diversity though the plateau period, from nearly 1,500 genera in the mid-Ordovician to just over 1,000 in the mid-Permian. Despite this, three separate logistic equations still provide a better overall description of the generic marine curve than the single exponential model (fig. 3*D*), which has an R^2 value of only 0.54. The equilibrium parameter value (D_{eq}) of the Meso-Cenozoic logistic model fit predicts that this curve will flatten out at a value of 13,000 fossilizable marine genera, compared with 4,000 known in the Recent. However, extrapolating future diversity levels from models fit to empirical data is inadvisable because of the large number of unique solutions possible for curves that are almost identical before approaching equilibrium.

A more parsimonious explanation for the diversification of life since the Permian in terms of fewer parameters required to describe the data is that expansion has been exponential and not logistic. The fit of the exponential model to the Meso-Cenozoic interval is shown in table 4. Comparing these values with those for the logistic model fit to the same time period (tables 2, 3) shows that the

exponential model provides as good a description of the data as the logistic in the case of the marine familial level curve produced by the Fossil Record 2 data. It also has almost as good a fit as the logistic model for both of the Sepkoski data sets. The diversification rate parameters (r_0 and r_d) are very similar for both models. However, the good fit of the exponential model to Meso-Cenozoic diversity data is not to say that life will continue to expand exponentially forever; as in the past, a future mass extinction may reset diversification dynamics. Alternatively, logistic slowing of diversification may indeed materialize. However, purely as a description of the empirical data, an exponential model provides as good a description of diversification since the Permian as the logistic, without the need for an additional parameter predicting future equilibrium.

A Model of Phanerozoic Marine Species Diversity

The parameters derived from the fits of the three logistic equations to the familial and generic marine diversity curves were used to calculate parameters defining a hy-

Table 3: Parameters derived from best fits of three logistic equations to the three periods of the Phanerozoic diversity curve, as shown in figure 2

Data set, taxonomic level, and time period of fit	D_0	r_0	D_{eq}
Sepkoski 1992:			
Family:			
Cambrian	1	.142	164
Later Paleozoic	141	.071	505
Meso-Cenozoic	238	.009	1,777
Benton 1993:			
Family:			
Cambrian	13	.051	192
Later Paleozoic	182	.088	474
Meso-Cenozoic	277	.006	5,218
Sepkoski 2002:			
Genus:			
Cambrian	1	.166	291
Later Paleozoic	233	.106	1,086
Meso-Cenozoic	217	.013	1.3×10^4
Model:			
Species:			
Cambrian	1	.194	516
Later Paleozoic	385	.158	2,335
Meso-Cenozoic	198	.028	95,104

Note: D_0 = initial diversity level (number of taxa), r_0 = initial diversification rate (net number of new taxa per lineage million years), D_{eq} = equilibrium diversity level (number of taxa). The modeled species-level parameters are derived by extrapolating the difference between the empirical familial and generic parameters for each period (Sepkoski data used).
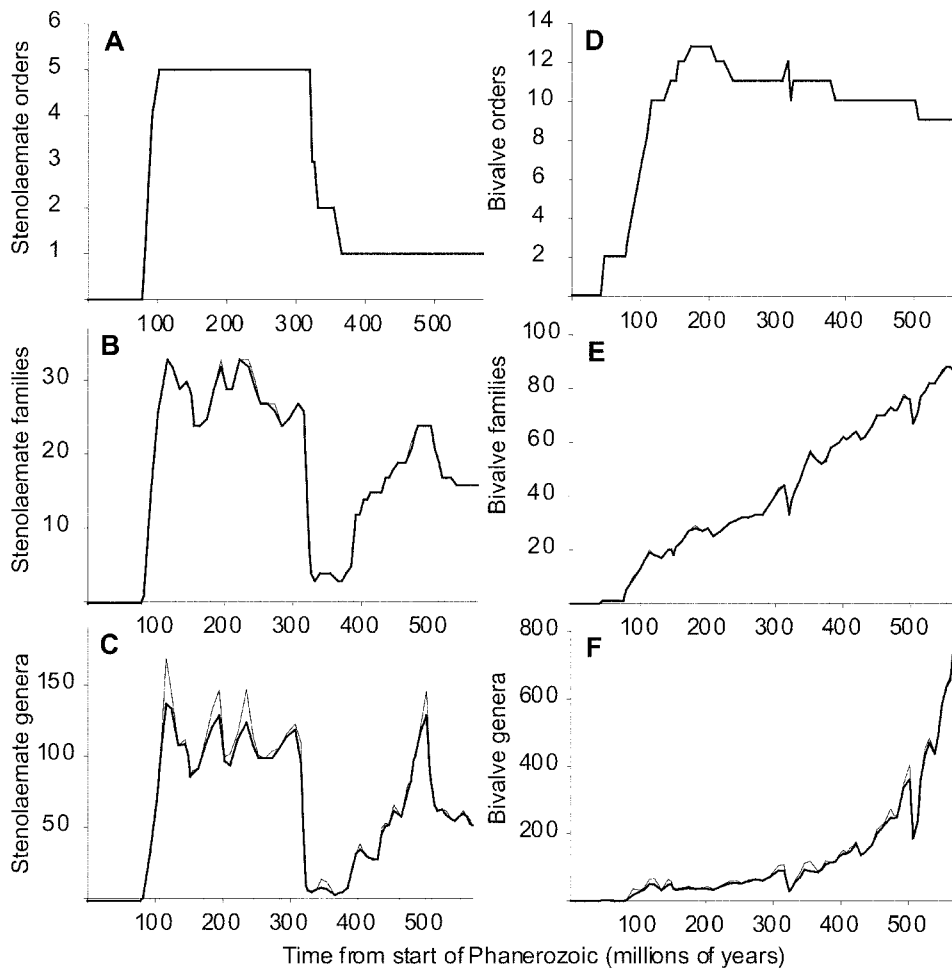
Figure 3: Global Phanerozoic diversity curves for the classes Stenolaemata and Bivalvia. *A*, Stenolaemata orders. Data from Sepkoski (1992). *B*, Stenolaemata families. Data from Sepkoski (1992). *C*, Stenolaemata genera. Data from Sepkoski (1996). *D*, Bivalvia orders. Data from Sepkoski (1992). *E*, Bivalvia families. Data from Sepkoski (1992). *F*, Bivalvia genera. Data from Sepkoski (2002). Solid lines exclude singletons.

pothetical species level curve. These model parameters are shown in table 3.

The modeled species marine diversity curve is illustrated in figure 2, with the empirical familial and generic plots for comparison. The curve (fig. 2*D*) has a diversity pattern recognizable as the Phanerozoic marine curve. However, the two equilibria apparent in the Cambrian and later Paleozoic are much reduced compared with the sharp rise in species numbers after the end-Permian extinction. This extinction event itself is more severe than those seen at generic and familial levels, with species numbers dropping from the 2,335 in the Carboniferous to 198 in the Triassic, a 91.5% drop. This is consistent with the theory that plotting diversity at higher taxonomic levels dampens extinction events (Raup 1979). Families are more likely to range through extinction events; therefore, both the percentage

and absolute numbers surviving are greater than those seen at generic and species level (cf. Jablonski and Raup 1995). The post-Permian rise in fossil species numbers is dramatic, from 198 in the Triassic to more than 17,000 in the Recent (a hypothetical future fossil sample of the present time period). Despite the long Paleozoic plateau period being evident, the sharp Meso-Cenozoic rise causes the overall form of the curve to appear exponential.

Class-Level Diversity Patterns

The effect of taxonomic level on the diversity curves of smaller groups of organisms was also investigated using the Sepkoski (1992, 2002) data sets. Diversity curves were plotted for two marine invertebrate classes at ordinal, familial, and generic levels. The resulting curves are shown

Table 4: Data and statistics for the fit of the exponential model to the Meso-Cenozoic phase of Phanerozoic marine diversity

Data set	Taxonomic level	Time period of fit	R^2	F	P	r_d
Sepkoski 1992	Family	Meso-Cenozoic	.96	876	<.0001	.006
Benton 1993	Family	Meso-Cenozoic	.99	2998	<.0001	.006
Sepkoski 2002	Genus	Meso-Cenozoic	.95	786	<.0001	.012

Note: R^2 = coefficient of determination, F = ratio of regression variance to residual variance, P = probability that the association between variables is incorrect. The value of the free parameter of the model, r_d (diversification rate), is given for each fit (units are taxa per lineage million years).

in figure 3. Those for Stenolaemata, a class with Paleozoic dominance, display a reasonably strong similarity among the plots (fig. 3A–3C), particularly through the Paleozoic era. An apparent equilibrium is an element of all three patterns, although with increasing fluctuations from ordinal to generic level. Bivalvia, a class with post-Paleozoic dominance, displays a poorer match among ordinal-, familial-, and generic-level plots (fig. 3D–3F), with each pattern appearing decoupled from the other two.

Discussion

The question of whether the convention of plotting Phanerozoic marine diversity curves at high taxonomic levels is representative of the species pattern is difficult to answer in the absence of adequate species-level data. The combination of empirical data plotting, equation fitting, and simple species diversity modeling presented here goes some way toward resolving this issue. Plots at ordinal, familial, and generic levels show that taxonomic level does have an effect on the perceived diversification pattern: the single logistic model reduces in its goodness-of-fit from ordinal to generic data, while conversely, the exponential model improves. Neither, however, provides as good a description of the marine data as three sequential logistic equations, although once again the goodness-of-fit of these equations reduces from familial to generic level. In addition, the low number of Cambrian data points and the excellent fit of the exponential model to the Meso-Cenozoic phase demonstrate that it is only the post-Cambrian Paleozoic pattern that provides convincing evidence for logistic diversification of marine life on a global scale. Courtillot and Gaudemer (1996) did not attempt to fit a Cambrian phase equation; instead, they included the Cambrian data in their Paleozoic fit of the logistic model at familial level. They also fitted multiple logistic equations to the Meso-Cenozoic data, but here both a single logistic and a single exponential equation are shown to provide an excellent fit to this section of the curve. Global marine data curves do not provide any empirical evidence suggesting that life has diversified logistically rather than ex-

ponentially since the Permian nor that marine diversity is showing signs of approaching a future equilibrium level.

Therefore, the overall trend from the plots of marine diversity at ordinal, familial, and generic levels is one of decreasing robustness of the curve's equilibrial elements moving down the taxonomic hierarchy and of an increasingly exponential form of the Phanerozoic pattern taken as a whole. The results corroborate the theory that higher taxonomic ranks become established early in the history of a group, while lower ranks gradually appear in increasing numbers as the group ages (Raup 1983), and the results also corroborate the prediction of Sepkoski (1984) that the perceived increase in familial diversity in younger geological periods underestimates the actual increase in species diversity. These trends are revealed further by the simple model of Phanerozoic marine species diversification. The resulting species diversity curve shows the logistic elements of the Cambrian and later Paleozoic significantly reduced in magnitude relative to the sharp rise in species numbers in the Meso-Cenozoic to such an extent that the overall pattern has an exponential form. This pattern could indeed be merely an extrapolation of poor sampling at generic level, where fossil genera are less likely to be sampled with increasing age of sediments. The generic data, however, have been interpreted as providing a clear biodiversity signal (Sepkoski 1996, 1997) and are considered robust enough to have been used in several recent biodiversity studies (Foote 2000a, 2000b, 2001; Kirchner and Weil 2000a, 2000b). The species model demonstrates that an expansive growth pattern can be constructed from individual logistic elements and, further, that if the equilibrial elements of the curve are degraded at species level to the same extent as is apparent from familial to generic, the overall diversification pattern of life through the Phanerozoic appears exponential, as suggested by Benton (1995, 1997, 2001) and by the species models of Signor (1985). This hypothesis requires further testing with empirical data.

The decoupling of pattern among diversity plots at differing taxonomic levels becomes more acute when smaller groups of organisms are studied. Figure 3 shows diversity plots at ordinal, familial, and generic levels for the two

marine invertebrate classes Stenolaemata and Bivalvia. The diversity patterns of bivalves again strongly demonstrate the tendency for large numbers of higher taxa to become established early in a group's history, while lower-rank taxa increase in number more slowly (Raup 1983). The equilibrium seen within the ordinal data is not evident at generic level. It is possible that bivalve genera are indeed diversifying logistically or that a future equilibrium may be attained. However, there is nothing in the empirical data itself to suggest that logistic diversification is more probable than exponential; rather, this is an assumption of macroevolutionary competition theory. Whatever the future holds for bivalves, their diversity patterns at differing taxonomic levels remain decoupled from one another. The bivalve diversity plots demonstrate that any pattern evident at ordinal, familial, or even generic levels may not be a reliable indicator of the species growth dynamics of the group. How to overcome this problem is an important consideration for paleobiodiversity analysis, particularly if higher taxa are being used as proxies for species. Correcting for sampling intensity (e.g., Alroy et al. 2001) may not be enough to reveal a correct species-level pattern if the data are analyzed at the rank of genera or families. To verify the robustness of any particular pattern, diversity must be plotted at as many taxonomic levels as the data allows and a consistent signal uncovered. In this way, we may begin to uncover the true picture of species diversification through time.

Acknowledgments

Literature Cited

Adrain, J. M., and S. R. Westrop. 2000. An empirical assessment of taxic paleobiology. Science 289:110–112.

Alroy, J. 1998. Equilibrial diversity dynamics in North American mammals. Pages 233–287 *in* M. L. McKinney and J. A. Drake, eds. Biodiversity dynamics: turnover of populations, taxa, and communities. Columbia University Press, New York.

———. 2000. New methods for quantifying macroevolutionary patterns and processes. Paleobiology 26: 707–733.

Alroy, J., C. R. Marshall, R. K. Bambach, K. Bezusko, M. Foote, F. T. Fürish, T. A. Hansen, et al. 2001. Effects of sampling standardization on estimates of Phanerozoic

marine diversification. Proceedings of the National Academy of Sciences of the USA 98:6261–6266.

Bambach, R. K. 1989. Similarities and differences in diversity patterns at different taxonomic levels using traditional (non-cladistic) groups. Geological Society of America Abstracts with Program 21:A206-A207.

Benton, M. J., ed. 1993. The fossil record 2. Chapman & Hall, London.

———. 1995. Diversification and extinction in the history of life. Science 268:52–58.

———. 1997. Models for the diversification of life. Trends in Ecology & Evolution 12:490–495.

———. 1999. The history of life: large databases in palaeontology. Pages 249–283 *in* D. A. T. Harper, ed. Numerical palaeobiology. Wiley, Chichester.

———. 2001. Biodiversity on land and in the sea. Geological Journal 36:211–230.

Bottjer, D. J., and D. Jablonski. 1988. Paleoenvironmental patterns in the evolution of post-Paleozoic benthic marine invertebrates. Palaios 3:540–560.

Bowring, S. A., J. P. Grotzinger, C. E. Isachsen, A. H. Knoll, S. M. Pelechaty, and P. Kolosov. 1993. Calibrating rates of Early Cambrian evolution. Science 261:1293–1298.

Brasier, M. D., and S. S. Sukhov. 1998. The falling amplitude of carbon isotopic oscillations through the Lower to Middle Cambrian: northern Siberia date. Canadian Journal of Earth Sciences 35:353–373.

Carr, T. R., and J. A. Kitchell. 1980. Dynamics of taxonomic diversity. Paleobiology 6:427–443.

Courtillot, V., and Y. Gaudemer. 1996. Effects of mass extinctions on biodiversity. Nature 381:146–148.

Davidek, K., E. Landing, S. A. Bowring, S. R. Westrop, A. W. A. Rushton, R. A. Fortey, and J. M. Adrain. 1998. New uppermost Cambrian U-Pb date from Avalonian Wales and age of the Cambrian-Ordovician boundary. Geological Magazine 135:305–309.

Enquist, B. J., J. P. Haskell, and B. H. Tiffney. 2002. General patterns of taxonomic and biomass partitioning in extant and fossil plant communities. Nature 419:610–613.

Erwin, D. H., J. W. Valentine, and J. J. Sepkoski, Jr. 1987. A comparative study of diversification events: the early Paleozoic versus the Mesozoic. Evolution 41:1177–1186.

Flessa, K. W., and D. Jablonski. 1985. Declining Phanerozoic background extinction rates: effect of taxonomic structure? Nature 313:216–218.

Foote, M. 2000*a*. Origination and extinction components of taxonomic diversity: general problems. Paleobiology 26(suppl.):74–102.

———. 2000*b*. Origination and extinction components of taxonomic diversity: Paleozoic and post-Paleozoic dynamics. Paleobiology 26:578–605.

———. 2001. Inferring temporal patterns of preservation,

origination, and extinction from taxonomic survivorship analysis. Paleobiology 27:602–630.

Foote, M., and D. M. Raup. 1996. Fossil preservation and the stratigraphic ranges of taxa. Paleobiology 22:121–140.

Harland, W. B., R. L. Armstrong, A. V. Cox, L. E. Craig, A. G. Smith, and D. G. Smith. 1990. A geologic time scale. Cambridge University Press, Cambridge.

Jablonski, D., and D. M. Raup. 1995. Selectivity of the end-Cretaceous marine bivalve extinctions. Science 268:389–391.

Johnson, K. G., and T. McCormick. 1999. The quantitative description of biotic changes using palaeontological databases. Pages 226–247 *in* D. T. Harper, ed. Numerical palaeobiology. Wiley, Chichester.

Kirchner, J. W., and A. Weil. 2000*a*. Correlations in fossil extinction and origination rates through geological time. Proceedings of the Royal Society of London B, Biological Sciences 267:1301–1309.

———. 2000*b*. Delayed biological recovery from extinctions throughout the fossil record. Nature 404:177–180.

Kitchell, J. A., and T. R. Carr. 1985. Nonequilibrium model of diversification: faunal turnover dynamics. Pages 277–309 *in* J. W. Valentine, ed. Phanerozoic diversity patterns: profiles in macroevolution. Princeton University Press, Princeton, N.J.

Landing, E. 1994. Precambrian-Cambrian boundary ratified and new perspective of Cambrian time. Geology 22:179–182.

Landing, E., S. A. Bowring, K. L. Davidek, S. R. Westrop, G. Geyer, and W. Heldmaier. 1998. Duration of the Early Cambrian: U-Pb ages of volcanic ashes from Avalon and Gondwana. Canadian Journal of Earth Sciences 35:329–338.

MacArthur, R. H., and E. O. Wilson. 1967. The theory of island biogeography. Princeton University Press, Princeton, N.J.

McCormick, T., and A. W. Owen. 2001. Assessing trilobite biodiversity change in the Ordovician of the British Isles. Geological Journal 36:279–290.

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

Palmer, A. R. 1998. A proposed nomenclature for stages and series for the Cambrian of Laurentia. Canadian Journal of Earth Sciences 35:323–328.

Raup, D. M. 1972. Taxonomic diversity during the Phanerozoic. Science 231:1065–1071.

———. 1979. Size of the Permo-Triassic bottleneck and its evolutionary implications. Science 206:217–218.

———. 1983. On the early origins of major biologic groups. Paleobiology 9:107–115.

———. 1995. The role of extinction in evolution. Pages 109–124 *in* W. M. Fitch and F. J. Ayala, eds. Tempo and mode in evolution: genetics and paleontology 50 years after Simpson. National Academy Press, Washington, D.C.

Raup, D. M., and J. J. Sepkoski, Jr. 1982. Mass extinctions in the marine fossil record. Science 215:1501–1503.

———. 1984. Periodicity of extinction in the geologic past. Proceedings of the National Academy of Sciences of the USA 81:801–505.

Roy, K., D. Jablonski, and J. W. Valentine. 1996. Higher taxa in biodiversity studies: patterns from eastern Pacific marine molluscs. Philosophical Transactions of the Royal Society of London B, Biological Sciences 351:1605–1613.

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

———. 1979. A kinetic model of Phanerozoic taxonomic diversity. II. Early Phanerozoic families and multiple equilibria. Paleobiology 5:222–251.

———. 1981. A factor analytic description of the Phanerozoic marine fossil record. Paleobiology 7:36–53.

———. 1984. A kinetic model of Phanerozoic taxonomic diversity. III. Post-Paleozoic families and mass extinctions. Paleobiology 10:246–267.

———. 1992. A compendium of fossil marine animal families. 2d ed. Milwaukee Public Museum Contributions in Biology and Geology 83:1–156.

———. 1993. Ten years in the library: new data confirm paleontological patterns. Paleobiology 19:43–51.

———. 1996. Patterns of Phanerozoic extinctions: a perspective from global databases. Pages 35–52 *in* O. H. Walliser, ed. Global events and event stratigraphy. Springer, Berlin.

———. 1997. Biodiversity: past, present, and future. Journal of Paleontology 7:533–539.

———. 2002. A compendium of fossil marine animal genera. Bulletins of American Paleontology 363:1–563.

Sepkoski, J. J., Jr., and D. C. Kendrick. 1993. Numerical experiments with model monophyletic and paraphyletic taxa. Paleobiology 19:168–184.

Shergold, J. H. 1995. Timescales. 1. Cambrian. Australian Geological Survey Organisation record 1995/30.

Signor, P. W. 1985. Real and apparent trends in species richness through time. Pages 129–150 *in* J. W. Valentine, ed. Phanerozoic diversity patterns: profiles in macroevolution. Princeton University Press, Princeton, N.J.

Valentine, J. W. 1969. Patterns of taxonomic and ecological structure of the shelf benthos during Phanerozoic times. Palaeontology 12:684–709.

———. 1990. The macroevolution of clade shape. Pages 128–150 *in* R. M. Ross and W. D. Allmon, eds. Causes

of evolution: a palaeontological perspective. University of Chicago Press, Chicago.

Van Valen, L. M. 1984. A resetting of Phanerozoic community evolution. Nature 307:50–52.

Walker, T. D., and J. W. Valentine. 1984. Equilibrium models of evolutionary species diversity and the number of empty niches. American Naturalist 124:887–899.

Associate Editor: Jonathan B. Losos