Rarefaction Analysis of Morphological and Taxonomic Diversity

Mike Foote

Rarefaction analysis of morphological and taxonomic diversity

Mike Foote

Abstract.—Our assessment of morphological diversity is influenced by morphological extremes and therefore depends on sample size (taxonomic richness). Rarefaction predicts the morphological diversity that would probably be observed in a sample of reduced size, thereby allowing both compensation for differences in sample size that may be strictly preservational, and analysis of diversity structure, that is, the relationship between morphological and taxonomic diversity. Middle and Late Cambrian trilobites exhibit a diversity structure characterized by many variations on a few morphological themes. In contrast, Middle and Late Ordovician trilobites occupy a larger range of morphospace per unit of species richness. Diversity structure in the Devonian is similar to that in the Middle and Late Ordovician, but the magnitude of morphological diversity is lower in the Devonian, as many fewer species are observed. For blastoids, different aspects of morphological diversity (range of morphospace occupied, number of character states possessed, and number of different regions in morphospace occupied) exhibit different relationships to taxonomic richness. In all cases Permian blastoids are characterized by a diversity structure in which morphological diversity per unit of taxonomic richness is greater than for Devonian blastoids. Changes in morphological diversity in fissiculate blastoids appear to reflect evolution of continuous variation in thecal morphology more than changes in the number of character states. Saunders and Swan's data on Namurian ammonoids illustrate some significant differences in diversity structure among stratigraphic levels, but many apparent differences in morphological diversity are consistent with the possibility that they reflect the sampling of different numbers of species from the same underlying diversity structure. Rarefaction curves are also presented for idealized increases and decreases in diversity, and these are compared to some of the observed changes in trilobites, blastoids, and ammonoids.

Mike Foote. Museum of Paleontology and Department of Geological Sciences, The University of Michigan, Ann Arbor, Michigan 48109

Introduction

The three most studied aspects of biological variety—morphological, ecological, and taxonomic diversity—are clearly not independent, but are not so redundant that one is a vicar for the others. Relatively little attention has been paid to their mutual relationships, given the attention devoted to taxonomic richness. Diversity patterns are often depicted as the trajectory of some number through time or space (e.g., number of taxa, Valentine 1969; morphological variance, Foote 1991a,b; ichnofabric index, Droser and Bottjer 1989). However, diversity has not only a magnitude but also a structure. That is, the diversity of faunas differs not only quantitatively but also qualitatively. As with the consideration of richness and evenness as measures of diversity in communities, determining explicitly the relationship between diversity measures may be more informative than the choice of one over the others.

Although ecological diversity is discussed below, this paper focuses on morphological and taxonomic diversity. The following terminology is used: *richness* refers to the number of taxa; *morphological variety* refers to the variance in form or the amount of morphological space occupied, irrespective of taxonomic richness; and *diversity* (unqualified) refers to any of the foregoing concepts. I avoid the term *disparity* to describe morphological diversity, because it has been used to refer to variation among phylum- or class-level body plans (Runnegar 1987), whereas this study focuses on morphological diversity within a body plan.

Morphological variety is commonly measured indirectly by counting the number of taxa of a rank sufficiently high that it can be

presumed to represent substantial morphological divergence (Stanley 1979: p. 65 ff.). Somewhat more directly, morphological variety may be measured by counting what are deemed major morphological themes, commonly called *morphotypes* (for recent examples, see Erwin 1990; Fortey and Owens 1990). (I use the term *morphotype* for simplicity, not to advocate the typological treatment of morphology.) However, without morphometric data the magnitude of the differences among higher taxa or morphotypes may be difficult to assess. Moreover, secular patterns are often documented with the implicit assumption that the average morphological difference among higher taxa is essentially constant through time, but this assumption is demonstrably untrue in some cases (Foote 1991a). In contrast to the categorical assignment of forms to morphotypes, several authors have documented changes in variety using a morphometric approach, thus allowing assessment of the dispersion of forms in morphological space (e.g., Ward 1980; Saunders and Swan 1984; Foote 1991a,b).

Analysis of ecological diversity has occasionally involved ordination of species on a continuum (e.g., height above or below substrate; Ausich and Bottjer 1982; Bottjer and Ausich 1986), but this approach has been used relatively little. A similar attempt at the measurement of ecospace use by means other than counting taxa is represented by analysis of the extent of bioturbation (Droser and Bottjer 1989). Using both taxonomic data and the interpretation of the ecological roles played by various taxa, Valentine's (1969) seminal paper compared patterns of diversity at different taxonomic levels. Extending this style of analysis, Bambach (1983, 1985) assigned marine classes to ecological categories, or guilds, to determine changes in the extent of ecospace use. Furthermore, he investigated patterns of species richness within guilds to determine whether the increase in species richness observed during the Phanerozoic more strongly reflects greater packing within guilds or an increase in the number of guilds. Because morphotypes serve as vicars for many of his guilds (Bambach 1983: table II), his study represents a large-scale analysis of morphological, ecological, and taxonomic diversity simultaneously.

Rarefaction can be used to investigate the joint structure of morphological and taxonomic diversity. The essence of this approach is to determine how morphological variety within a sample, as measured by the range of occupied morphological space, varies as a function of the number of species (or other units) sampled. Rarefaction facilitates comparison of both quantitative and qualitative differences in diversity and therefore may allow more resolution in diversity analyses than comparison of single measures of the magnitude of diversity. The approach will be applied to some general patterns expected to result from the constraints of branching geometry, to Cambrian through Devonian trilobites, to Ordovician through Permian blastoids, and to Namurian ammonoids. Although rarefaction is applied here only to morphological variety as a function of taxonomic richness, it should also be easily applicable to ecological and molecular diversity.

Morphological Rarefaction

Sanders (1968) presented the rarefaction method to compensate for sample-size differences in the measurement of richness in ecological studies. Raup (1975) adopted this approach to determine whether changes in the number of higher taxa in the fossil record could simply reflect differences in the number of species. In essence, these authors asked, What is the number of taxa likely to be found if the sample (of lower taxa or individual organisms) is constrained to be of a certain size? An analogous question holds for morphological variety: We observe that a taxon with a certain number of species occupies a certain amount of morphospace, but how much morphospace would it be likely to occupy if it were constrained to be represented by a smaller number of species? This question may imply that sample size (the observed number of species) does not reflect species richness well, and therefore must be corrected for. However, it may also imply that sample size reflects richness accurately, but that one aspect of the structure of diversity is to be found in the relationship between richness and

morphological variety. Two higher taxa that have substantially different numbers of species and occupy the same amount of morphospace exhibit different structures of diversity, as do two taxa that have the same number of species and occupy substantially different amounts of morphospace.

Measuring Morphological Variety.—I assume for this discussion that a sample of species or other entities has been ordinated in some morphospace. How is morphological variety to be measured? Because the sample variance provides a measure of morphological dispersion that is relatively insensitive to sample size, much concern about the effects of sample size could be avoided by the use of the variance of morphological variables, proportional to the mean squared distance among species in morphospace. If sample size is held constant, then the range and variance are positively correlated in general (Sokal and Rohlf 1981: p. 58).

The average dissimilarity among species and the density of morphospace occupation certainly are taken into account to some extent when paleontologists assess morphological variety. However, it is generally the range of forms or the amount of morphospace occupied that serves as our principal guide to morphological variety when morphology is considered explicitly, that is, when taxonomic richness is not used as a measure of morphological diversity. For example, Saunders and Swan (1984) based their conclusions regarding morphological variety in ammonoids on inspection of plots of species in principal-component space. Their discussion and interpretations suggest that they were guided largely by the size of the scatter of points, not the average distance among those points. In discussing the diversification of mammals in the Cenozoic, one commonly focuses on the evolution of "everything from bats to whales," even if most species of mammals are rats or something like them. Similarly, in assessing changes in ecological diversity, Bambach (1983) focused largely on the number of ecological strategies actually used, not the average dissimilarity among them. The use of a richness-dependent measure does not clearly represent a shortcoming in our approach to

morphological variety. Because each particular form may be of interest in its own right, not just as a member of a statistical population of forms, it is reasonable to consider the actual range of realized forms when assessing diversity. Except where explicitly noted, I hereafter use the morphological range, the difference between maximum and minimum of some morphological variable, as a measure of morphological variety.

Van Valen (1974) discusses the multivariate generalization of the univariate variance (see also Ashton and Rowell 1975), suggesting two measures: the total variance (sum of univariate variances of morphological variables) and the product of univariate variances. The latter gives some indication of the amount (volume or hypervolume) of morphospace occupied, but this quantity vanishes or nearly vanishes if any of the variances are zero or nearly zero (Van Valen 1974). Analogous measures of morphological variety can be calculated using univariate ranges rather than univariate variances. Although the problem of vanishing products is reduced by the use of the few largest principal components (Van Valen 1974), for simplicity I use the total range, that is, the sum of the ranges. The results presented are consistent with those obtained when the geometric mean range (the kth root of the product of k values) is used instead. Because the product of variances or ranges amplifies differences in morphological variety as the dimensionality of morphospace increases, use of the geometric mean may be preferable if one uses the product at all (Foote 1991b).

Rarefaction Procedure.—It is critical, as in all statistical analysis, that sampling be random. In addition, different samples to be compared should represent the aggregation of species from roughly equivalent amounts of time and space, and these species should reflect essentially consistent taxonomic treatment from sample to sample. The last two requirements may be relaxed if one is interested in investigating, say, biogeographic effects or the effects of different taxonomists' treatment of the same material (cf. Raup 1978: pp. 88–89). Because the emphasis here is on comparing the same taxon at different times, all three
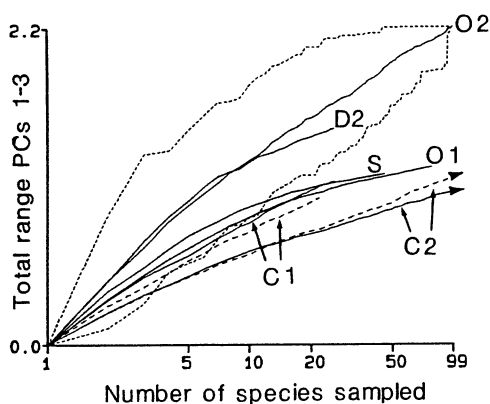
FIGURE 1. Morphological rarefaction of trilobites. N is the number of species sampled; total range is the expected sum of ranges of PCs 1–3 at the corresponding sample size. As sample size increases, each unit of increment in sample size yields a progressively smaller increment in morphological variety. Sample size is therefore plotted logarithmically. Abbreviations: C1, Early Cambrian; C2, Middle and Late Cambrian; O1, Early Ordovician; O2, Middle and Late Ordovician; S, Silurian; D2, Middle and Late Devonian. See table 1 for explanation of stratigraphic intervals. Short-dashed curves show empirically determined 90% confidence interval for O2 rarefaction curve, given total sample of 99 species. Long-dashed rarefaction curves correspond to the use of Cambrian genera (22 for Early Cambrian; 113 for Middle and Late Cambrian) rather than species as sampling units. C2 curves continue in direction of arrow; see table 1 for total range and total sample size. Rarefaction curves appear to fall into three groups. Note that Devonian, with lower morphological variety than Ordovician, may reflect simply a smaller number of species representing the same diversity structure. These and all subsequent rarefaction curves are based on the average values of 100 random samples.

conditions should hold as far as possible. The number of fossil species represented by morphological data is certainly smaller than the number that existed. Because taxa are more fully represented in some periods of time than others, direct comparison of morphological range or taxonomic richness may be difficult. However, rarefaction allows the comparison of samples that do not equally completely represent the taxon in question. Provided that the species measured represent an unbiased sample of morphology, it is inconsequential whether the limitation on sample size reflects lack of preservation, inaccessibility of material, financial or other constraints on the researcher, or some combination of these and other factors.

Rarefaction of morphological variety with

respect to number of species is illustrated with trilobites in figure 1. The data are the unstandardized scores of the first three principal components (PCs), based on the covariance matrix of the species means of 12 Fourier coefficients used to describe the outline of the cranidium (Foote 1989). These three PCs summarize roughly 88% of the total variance. Specimens were sampled randomly from large museum collections, as described previously (Foote 1991a). Data were lumped into stratigraphic intervals of approximately equal duration, based on the time scale of Harland et al. (1990) (table 1). Comparable results are obtained if different subdivisions of the Cambrian and Ordovician are used, as was done previously (Foote 1991a). Because of very low sample sizes, Early Devonian, Carboniferous, and Permian samples are not plotted. Sample size is the number of species actually measured, not the total known richness.

The rarefaction of taxonomic richness with respect to sample size can be calculated analytically, based on sampling theory (Sanders 1968; Raup 1975). Morphological variety can be rarefied analytically only if the data are well characterized, for example, multivariate normal. In general it is necessary to rarefy a sample by brute force, repeatedly pulling random subsamples of specified size. To determine the expected total range for a sample of two species, two species are drawn at random. For each morphological variable, the minimum and maximum in this subsample are determined, and the range for each variable is calculated as the absolute difference between minimum and maximum. The total range, TR, is calculated as the sum of the ranges over all the variables.

This last value represents the total range for but a single random draw of two species. Because such a single draw is not expected to yield a reliable estimate of the total range, the procedure is repeated. The two species are replaced, another random pair of species is drawn, and the total range is computed. The procedure is repeated many times, and the average total range over all the random draws of two species gives the expected, rarefied total range for a sample size of two. If there are m repeated random draws, then the dis-

TABLE 1. Time scale, number of species (N), total variance, and total range of PCs 1–3 for trilobite data. (Number of genera for Cambrian in parentheses.) Age at base of interval in millions of years before present, based on Harland et al. (1990). Duration in millions of years, also from Harland et al. (1990). (Early Cambrian, Atdabanian and Botomian; Middle and Late Cambrian, post-Botomian; Early Ordovician, Tremadocian and Arenigian; Middle and Late Ordovician, post-Arenigian; Middle and Late Devonian, post-Emsian.) Remainder of Paleozoic omitted because of very small sample size.

Stratigraphic interval	Age at base (Ma)	Duration (m.y.)	N	Total variance	Total range
Early Cambrian (trilobite-bearing)	560	24	26 (22)	.0286	1.13
Middle and Late Cambrian	536	26	191 (113)	.0157	1.20
Early Ordovician	510	34	78	.0313	1.24
Middle and Late Ordovician	476	37	99	.0691	2.21
Silurian	439	30.5	46	.0395	1.19
Middle and Late Devonian	386	23.5	26	.0576	1.50

tribution of TR_i, $i = 1, \ldots, m$, can be used to estimate confidence limits on the expectation. For simplicity, assume $m = 100$. If the TR_i are arranged from smallest to largest, then the 6th value and the 95th value provide an approximation to, respectively, the lower and upper 90% confidence limits on the expected total range for a sample of two species.

The entire procedure is then repeated for subsamples of three species, four species, and so on up to the total number of species actually sampled. All analyses presented here are based on 100 iterations of random sampling. Note, however, that there will not always be 100 or more unique ways of drawing a specified number of species from the larger sample. For example, there are only 45 ways (i.e., $10!/[2!(10 - 2)!]$) of drawing two species from a sample of ten, and there are only 50 ways (i.e., $50!/[49!(50 - 49)!]$) of drawing 49 species from a sample of 50. However, such cases are the exception rather than the rule here. I have therefore based all rarefaction curves on 100 iterations, even though the random samples drawn are sometimes redundant. In most examples presented here, the confidence limits are given for the stratigraphic interval with the largest sample size. One can check whether the (generally smaller) total morphological range of intervals with smaller samples lies within the confidence limits, thereby determining whether differences in morphological variety are consistent with simple differences in sample size. As sample size increases, the increase in morphological variety per unit increment in sample size becomes progressively smaller.

Therefore, sample size is plotted logarithmically.

Although each random sample is replaced before drawing the next subsample, the drawing of a single random sample itself is without replacement. Therefore, each species can be represented at most once in a random sample. One could also sample with replacement, which in essence effects calculation of the rarefaction curve by bootstrapping. Multiple draws of the same species in a single sample contribute no more to the morphological range of that sample than does a single draw of that species. Therefore, sampling with replacement yields rarefaction curves that are lower, that is, represent a smaller total range, than does sampling without replacement.

The rarefaction curves for trilobites appear to fall into three groups (fig. 1), roughly indicating three different structures of diversity. (1) The curve for the Middle and Late Cambrian shows the existence of many nominal species that are mere variations on a relatively limited array of morphological themes. (2) At the other extreme, both the Middle and Late Ordovician and the Middle and Late Devonian are represented by fewer species that, in total, occupy a large amount of morphospace. Whereas the Middle Ordovician is generally recognized as the time of maximal morphological variety in trilobites (Whittington 1966; Fortey and Owens 1990; Foote 1991a), the structure of morphological variety is not significantly different in the Devonian than in the Middle and Late Ordovician. The lower morphological variety in the Devonian may simply reflect the existence or sampling of a

smaller number of species. Published estimates of taxonomic diversity (Harrington 1959: fig. 107; Raup 1976: table 3) suggest that the difference in sample size between the Ordovician and Devonian accurately reflects a difference in the number of species described. Whether the number of described species in turn reflects the number that existed is another matter. (3) Early Cambrian, Early Ordovician, and Silurian trilobites are each characterized by an intermediate diversity structure.

Based on inspection of bivariate plots, I had previously interpreted the Early Cambrian as the time of minimal morphological variety among Cambrian and Ordovician trilobites (Foote 1991a). Although the total range of morphospace occupied by Early Cambrian trilobites is less than that occupied in the Middle and Late Cambrian (but barely!) (table 1), the diversity structure suggests two alternative interpretations. (1) If the difference in number of species between the Early Cambrian and the Middle and Late Cambrian reflects only sampling, preservation, taxonomic artifact, or some combination of these factors, then morphological variety may actually have been higher in the Early Cambrian. (2) If the increase in number of species reflects a true taxonomic diversification, this diversification was morphologically constrained (cf. fig. 2C below). Whether such constraint is adaptive is unclear. Use of genera rather than species as the sampling units for the Cambrian yields much the same pattern (fig. 1), suggesting that taxonomic practice may not be the principal cause of the shallow rarefaction curve for the Middle and Late Cambrian.

As mentioned above, use of the morphological variance could, in one sense, obviate the rarefaction analysis (table 1). However, the use of rarefaction curves serves at least two purposes. (1) It explicitly considers the amount of morphospace occupied, which is the aspect of a morphological distribution on which we commonly rely when assessing diversity. (2) It allows us to investigate the structure of diversity as the relationship between morphological and taxonomic diversity, rather than as a single number giving the average morphological difference be-

tween species. This is particularly important in cases where rarefaction curves cross, for example, Early Cambrian vs. Early Ordovician and Silurian. In such cases it may be difficult to state unequivocally which of several samples represents the greatest morphological variety, but one can still identify differences in how morphological variety is structured with respect to taxonomic richness.

Further Examples

General Branching Process.—Because comparative diversity analysis commonly focuses on large-scale evolutionary radiations or extinctions, it is useful to determine some of the general patterns that can be expected to result from substantial increases or decreases in richness. The suggested patterns in figure 2 are based primarily on simulations of taxonomic and morphological evolution, using an approach similar to that of Raup and Gould (1974), as subsequently modified (Foote 1991b). A simulation begins with a single lineage having an arbitrary, multidimensional morphology. At each time increment, each lineage may branch to form two daughter lineages, persist without branching, or become extinct. The prescribed probabilities for these three occurrences are stochastically constant over time, and are empirically scaled. In the event of branching, one daughter lineage retains its morphology while the other is allowed to change, with a specified probability, in each of the several morphological dimensions. Positive and negative increments in morphology are equally probable. The model uses morphological increments that are proportional to the magnitude of morphological variables (Haldane 1949), rather than being of constant size as in Raup and Gould's (1974) study. Furthermore, phyletic evolution—that is, morphological change that is independent of branching—is also permitted with a prescribed probability. Further details of the simulation procedure, as well as an analysis of the sensitivity of simulation results to parameters of the model, are given by Foote (1991b).

I emphasize that the focus here is on the geometry of branching, not the question of randomness or deviation from a null model,
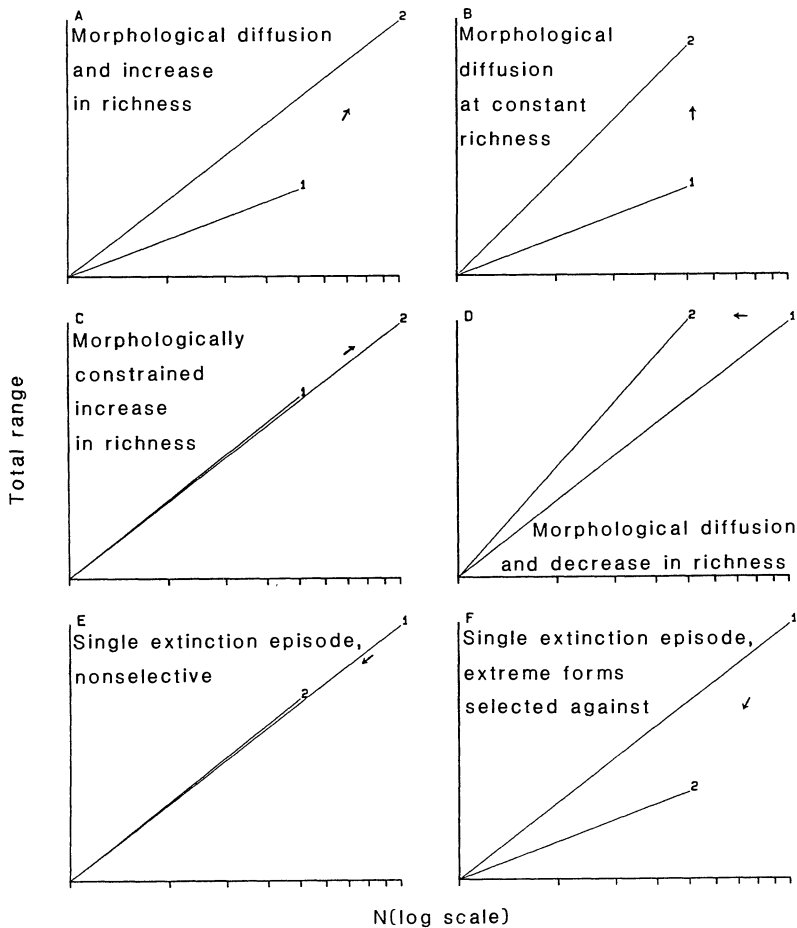
FIGURE 2. Rarefaction curves for highly idealized changes in diversity. Linearity of curves is only an approximation. In each graph, 1 and 2 indicate diversity structure before and after the diversity change, respectively. A and B, Morphological diffusion is to be expected in most clades; diffusion with constant richness cannot be readily distinguished from diffusion with a true increase in richness but a decrease in sample size. C, Morphologically constrained increase in richness cannot be readily distinguished from a simple increase in sample size. D, If extinction is morphologically nonselective and decrease in richness does not take place in a single event, morphological variety is maintained or increases despite decrease in richness. E, A single episode of nonselective extinction yields pattern identical to simple decrease in sample size. F, Single extinction episode selecting against extreme forms yields both lower richness and lower morphological variety.

and that the results are highly idealized. Furthermore, the results are not strikingly new; they are presented as a benchmark against which to compare the empirical examples presented here. Finally, whereas I present the changes in diversity structure likely to result under various evolutionary scenarios, observed changes in diversity structure that match those presented here should only be taken to suggest possible avenues for further exploration. They are by no means conclusive evidence for any particular process or event.

A clade evolving without morphological constraint can often be expected to increase in morphological variability over time, that is, to diffuse through morphospace (Fisher 1986). Whether this diffusion is associated with an increase in richness (fig. 2A) or constant richness (fig. 2B), the result is an increase in the steepness of the rarefaction curve, that is, both a quantitative change in the magnitude of morphological variety and a qualitative change in diversity structure. This change would seem to differ from the pattern expected if diversification proceeds by the early establishment of major body plans and
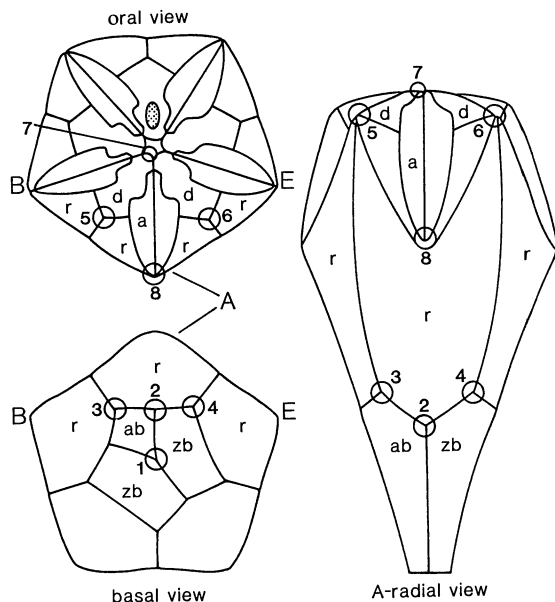
FIGURE 3. Generalized blastoid theca showing morphological landmarks. 1, conjunction of three basal plates; 2, conjunction of A-radial, E-zygous basal, and azygous basal; 3, conjunction of A-radial, B-radial, and azygous basal; 4, conjunction of E-radial, A-radial, and E-zygous basal; 5, conjunction of A-radial, B-radial, and AB deltoid; 6, conjunction of E-radial, A-radial, and EA deltoid; 7, most adoral point on AB deltoid; 8, most aboral point on A-ambulacrum. Abbreviations: a, ambulacrum; ab, azygous basal; d, deltoid; r, radial; z, zygous basal; A, B, and E refer to rays.

TABLE 2. Time scale (based on Harland et al. 1990), number of species (N), total variance and total range of PCs 1–3 for blastoid data. No morphological data are available for Pennsylvanian.

Stratigraphic interval	Age at base (Ma)	Duration (m.y.)	N	Total variance	Total range
Silurian	439	30.5	4	.0291	0.592
Devonian	408.5	46	25	.541	3.44
Mississippian	362.5	40	42	.282	3.89
Permian	290	45	13	.737	4.03

subsequent evolution at lower levels (Valentine 1969). In such a case, rarefaction curves should be steep and short initially, as a few taxa occupy much morphospace, and then become long and shallow, as diversification produces variations on these major themes. If taxonomic richness increases but newly evolved forms are constrained to be little more than minor variants of existing forms, then there will be essentially no change in diversity structure. However, the magnitude of both morphological variety and taxonomic richness will increase (fig. 2C).

If taxonomic richness decreases over some extended period of time, while speciation, extinction, and morphological evolution continue, the expected pattern is along the lines of that depicted in figure 2D. Despite the decline in richness, morphological variety can be maintained or even increase. This change is consistent with the suggestion that morphological variety tends to be concentrated later than taxonomic richness in the history

of a randomly evolving clade and that some real clades may agree with this pattern (Foote 1991b). If numerous species within a clade become extinct in a single episode, the result is somewhat different. If the species that survive are a morphologically random subset of those extant before the extinction event, the resulting pattern cannot be distinguished from a simple decrease in sample size (fig. 2E). If, in contrast, whole regions toward the periphery of morphospace are wiped out by an extinction episode, the result is more likely to be a depressed rarefaction curve (fig. 2F). Comparison of figure 2A–C with figure 2D–F demonstrates an important difference between morphological and taxonomic diversity. Both speciation and lineage extinction change taxonomic richness by the same amount, and, therefore, balance each other out. However, morphological variance may stubbornly tend to increase in a clade, even at constant richness (Fisher 1986; Gould 1988; Foote 1991b). Therefore, if extinction is effectively random with respect to morphology, a disproportionately large number of extinctions is needed to reduce morphological variety substantially.

Blastoids.—The echinoderm class Blastoidea is generally considered to have reached its acme during the Mississippian, when roughly half of all known genera are found (Waters 1988). In order to measure morphology in blastoids, morphometric variables were defined as the Cartesian coordinates of eight landmarks in three dimensions (not the distances among these landmarks), on thecae placed in a common orientation and scaled to a common size (Foote 1991b). These landmarks are primarily the junctions of sets of three thecal plates (fig. 3). The first three (unstandardized) principal components of the
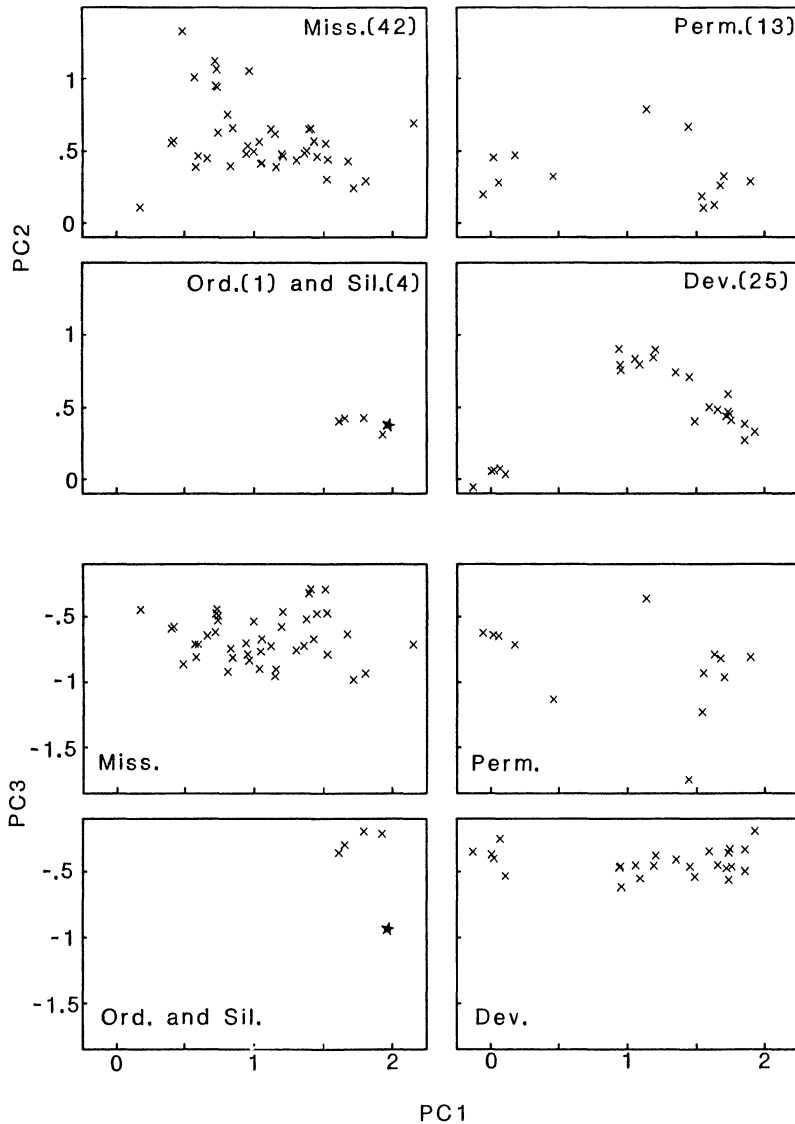
FIGURE 4. Ordination of 85 blastoid species along first three principal components, arranged by stratigraphic interval. Number of measured species in parentheses. Star gives the location of the oldest known blastoid, *Macurdablastus uniplicatus*, from the Caradocian of Tennessee (Broadhead 1984). Note increase in both morphological variety and sample size from Silurian to Devonian and Devonian to Mississippian, while transition to Permian is marked by continued increase in morphological variety despite decrease in sample size. Miss., Mississippian; Perm., Permian; Ord., Ordovician; Sil., Silurian; Dev., Devonian.

covariance matrix of these coordinates summarize roughly 94% of the total variance and are used here to measure morphological variety. The time scale (based on Harland et al. 1990), sample size, total variance, and total range are presented in table 2. No morphological data are currently available for the Pennsylvanian.

The transition from the Silurian to the Devonian appears to demonstrate an increase in both morphological variety and taxonomic richness (figs. 4, 5A), similar to that shown in figure 2A. Whereas inspection of bivariate plots might suggest greater morphological variety in the Mississippian than the Devonian (fig. 4), the diversity structures of these two intervals are difficult to distinguish (fig. 5A). However, the morphological variety of the Devonian is greatly influenced by the highly divergent species of *Nucleocrinus*, in
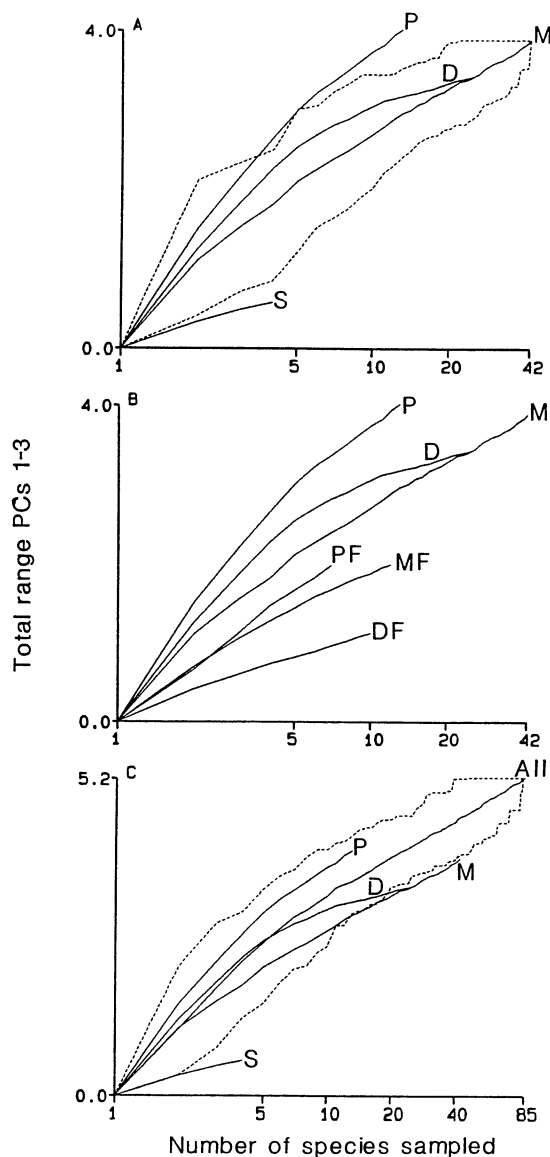
FIGURE 5. Morphological rarefaction of blastoids, based on PCs 1-3 (see fig. 4). A, Comparison of four stratigraphic intervals. Dashed curves show 90% confidence interval for Mississippian rarefaction curve. Devonian and Mississippian diversity are not significantly different, whereas Permian and Silurian blastoids occupy more morphospace and less morphospace, respectively, per unit of taxonomic richness than Mississippian or Devonian blastoids. B, Comparison of all blastoids within stratigraphic interval to fissiculates within interval. Shallower rarefaction curves indicate that fissiculate blastoids occupy less morphospace per unit of taxonomic richness than the entire class, but this is disproportionately true of Devonian blastoids, whose morphological variety is greatly influenced by the single spiraculate genus *Nucleocrinus*. C, Comparison of species within stratigraphic intervals to entire data set (curve labeled "All"). Dashed curves show 90% confidence interval for rarefaction of entire data set. Blastoids in each interval except the Perm-

the lower left of figure 4A and the upper left of figure 4B. The steep rarefaction curve for the Permian (fig. 5A) suggests a net decline in taxonomic richness accompanied by continued branching, extinction, and morphological evolution (cf. fig. 2D).

It should come as no surprise if a taxonomically restricted subset of species exhibits a shallower rarefaction curve than the entire taxon. Figure 5B shows the rarefaction of all blastoids, arranged by stratigraphic interval, compared to those of the (paraphyletic) order Fissiculata only. The Silurian is omitted because there are morphological data for only two fissiculate species. For fissiculates only, Permian blastoids still occupy morphospace more extensively per unit of taxonomic richness than Devonian or Mississippian blastoids, but note that the Devonian curve is now well below that for the Mississippian. Spiraculates, such as *Nucleocrinus*, mentioned above, appear to contribute disproportionately heavily to the morphological variety of Devonian blastoids. Of course, Silurian, Devonian, Mississippian, and Permian blastoids each occupy a smaller range of morphospace than the class as a whole. However, Permian blastoids occupy morphospace so extensively that Permian-sized random samples from the whole class would be unlikely to occupy as much morphospace as do Permian blastoids themselves (fig. 5C).

Namurian Ammonoids.—In a model analysis of morphological diversity, Saunders and Swan (1984) used coiling parameters and many other measures of form to ordinate scores of species of mid-Carboniferous ammonoids in morphological space. Dividing available species into eight successive stratigraphic intervals, they documented changes in the occupation of morphospace and related these to changes in taxonomic richness and geographic representation. In addition to determining the taxonomic representation of

←

ian occupy not only less morphospace than the entire class, but also less morphospace per unit of taxonomic richness. Abbreviations: S, Silurian; D, Devonian; M, Mississippian; P, Permian; DF, Devonian fissiculates; MF, Mississippian fissiculates; PF, Permian fissiculates.

TABLE 3. Sample size (N), total variance, and total range of PCs 1–2 for Namurian ammonoids. Intervals correspond to those used by Saunders and Swan (1984).

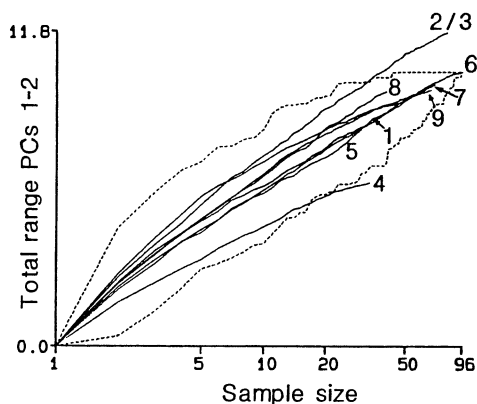| Stratigraphic interval	N	Total variance	Total range
1	46	2.75	9.08
2/3	81	2.81	11.75
4	33	1.14	6.11
5	30	1.81	8.11
6	96	2.03	10.26
7	70	1.68	9.79
8	40	2.15	9.53
9	67	2.46	9.59

FIGURE 6. Morphological rarefaction of total range of PCs 1 and 2 for Namurian ammonoids. Data digitized from Saunders and Swan (1984). Numbers 1–9 refer to Saunders and Swan's stratigraphic subdivisions. Dashed curves show 90% confidence interval for rarefaction curve of level 6. Note the relatively small variation in diversity structure among the intervals; many perceived differences in morphological variety may therefore reflect differences in sample size. Exceptions include the transition from level 2/3 to level 4, marked by a clear reduction in morphological variety per unit of taxonomic richness, and the transition from level 4 to level 5, marked by an increase in morphological variety per unit taxonomic richness. Coincidence of many rarefaction curves is consistent with Saunders and Swan's (1984) interpretation that morphological distribution reflects functional constraint.

several morphotypes, the authors assessed morphological variety by the amount of morphospace occupied, as judged from principal-component plots. Therefore, rarefaction analysis should allow further insight into the morphological variety of Namurian ammonoids.

Saunders and Swan's plots were digitized to obtain the first and second principal-component scores, and these data were rarefied with respect to sample size. The authors represented some species by more than one point, and they also approximated the position in morphospace of some existing specimens that were not measured. In order to keep the rarefaction analysis comparable to the authors' presentation, I have retained, rather than averaged, multiple specimens, and have retained specimens whose ordination in morphospace is approximated. Sample sizes, total variance, and total range are given in table 3.

Although there is apparently less difference in diversity structure among successive Namurian ammonoid samples than among trilobite or blastoid samples, the rarefaction curves for ammonoids (fig. 6) do illustrate at least two of Saunders and Swan's conclusions regarding diversity changes. (1) The authors note a "drastic reduction in morphological diversity" (Saunders and Swan 1984: p. 211) in the transition from level 2/3 to level 4. This transition reflects a change from the steepest to the shallowest rarefaction curve; that is, from one extreme of diversity structure to the other (cf. fig. 2F). (2) Comparing level 8 to level 7, the authors state that "the range in diversity in level 8 is similar to that in the preceding level, [but] morphospace is

more sparsely occupied, reflecting a decline in taxonomic diversity within major morphotypes" (Saunders and Swan 1984: p. 215). This change is reflected in the steepening of the rarefaction curve from level 7 to 8, with both curves reaching roughly the same eventual height (cf. fig. 2D).

Despite the general agreement between Saunders and Swan's conclusions and the implications of rarefaction analysis, there are at least two instances in which the rarefaction analysis allows a modified interpretation of diversity patterns. (1) The authors state that morphological diversity in level 5 is roughly the same as in level 4. This observation is based not just on the distribution of forms per se, but also on the representation of morphotypes. However, because ammonoids in level 5 are less centrally concentrated (i.e., more dispersed) than in the previous level (Saunders and Swan 1984: fig. 16), the rarefaction curve for level 5 is steeper and may suggest a morphological diversification (cf. fig. 2A,B). (2) The authors identify level 6 as

exhibiting "a sharp increase in morphological diversity and . . . a striking morphological radiation" compared to the previous levels (Saunders and Swan 1984: p. 213). However, level 6 is characterized not only by the second largest range in morphology (after level 2/3), but also by the greatest sample size (table 3). The rarefaction curve for level 5 is essentially an interpolation of that for level 6, suggesting no striking difference in diversity structure. The transition may reflect simply a change in sample size or a highly constrained diversification (cf. fig. 2C).

The foregoing comments reflect a different perspective on Saunders and Swan's analysis, not a critique of their conclusions. Rarefaction analysis suggests that there is relatively little variation in ammonoid diversity structure within the Namurian, although the magnitude of diversity certainly varies appreciably. The stability of diversity structure in Namurian ammonoids would appear consistent with the authors' interpretation that the morphological distribution reflects functional constraint (Saunders and Swan 1984; Swan and Saunders 1987), although other sources of constraint are not ruled out.

Our perception of morphological variety is so strongly tied to morphological extremes that it must be tempered with the recognition that many striking differences in morphological variety may reflect only different numbers of species representing the same underlying diversity structure. Whether such differences reflect artifacts of preservation or sampling on the one hand, or true changes in species richness on the other, is a question that requires further analysis and detailed knowledge of the evolutionary history of the group, such as Saunders and Swan present.

Further Extensions

The foregoing analyses consider only a single aspect of morphological variety, albeit one that is generally important in our interpretations. Other factors that we consider when assessing morphological variety—namely, the number of character states exhibited by an array of species and the number of different regions occupied in morphospace—also depend on sample size. Therefore, it may be

fruitful to consider them explicitly with respect to taxonomic richness. Because not all data distributions are rectangular or ellipsoidal, other measures of morphological diversity may also be useful in some cases. Such measures include the maximum pairwise distance among points in morphospace, and the ranges along major and minor axes of each subset of data, rather than along the PCs of the total distribution. Rarefaction of these measures is not discussed here but easily extends from the procedure presented above.

Discrete Character Data.—For a given number of coded characters, a group of species with a greater number of different character states would, all else being equal, be considered to exhibit greater morphological variety. If the number of characters is fixed, then the number of different character states clearly depends on the number of taxa sampled and should be rarefied. Rarefaction of discrete character data is illustrated in figure 7. As part of their comprehensive revision of the order Fissiculata, Breimer and Macurda (1972: tables I, II) presented character states for 19 characters in 38 genera of fissiculate blastoids. There are a total of 59 character states. Most of the characters are present in all genera and are represented by different states; only a few characters are scored as present or absent. Rarefaction proceeds exactly as for continuous morphometric data with the exception that, instead of the morphological range for each variable and the total range summed over all variables, the number of different character states for a particular character and the total number of character states summed over all characters are determined for each random draw. Rarefaction of the total number of character states against number of genera (fig. 7) reveals less variation in diversity structure among stratigraphic intervals than observed with rarefaction of the total range (fig. 5). Thus, morphological diversification in the Fissiculata appears to have involved evolution of continuous morphometric variation more than the addition of new character states. Note that these character states are among the features used to differentiate genera. Therefore, the two axes of the rarefaction plots are probably constrained to be more highly cor-

related than they are with continuous morphometric data. Because a different measure of morphological variety is used here, these curves should not be compared directly to those in figure 2.

Number of Occupied Regions in Morphospace.— The diversity structure of Devonian blastoids partly reflects the influence of the single genus *Nucleocrinus*; there is a large region of vacant morphospace between this genus and the remaining cluster of species (fig. 4). One may be interested in the number of subregions within morphological space actually occupied, rather than simply the total range, because the latter may be greatly affected by single extremes. If the morphospace is divided, via a gridwork, into cells of equal volume or hypervolume, one can tally whether or not a particular region of morphospace is occupied. The number of occupied cells provides a measure of morphological variety, but it clearly depends on sample size as well as the number of cells into which the space is divided. If the space is divided too finely, the number of occupied cells will exactly equal the number of species sampled, and all morphological distributions will appear to have the same diversity structure. On the other hand, if the space is too coarsely divided, nearly any morphological distribution will occupy the same number of cells. There is no a priori "correct" number of cells into which to divide the space, but one can experiment to determine the effect of using different numbers of cells. Counting the number of occupied cells is analogous to tallying the number of morphotypes without regard for the morphological differences among them.

The total, occupied region of the three-dimensional principal-component space for blastoids is divided into 27, 64, and 125 cells (corresponding, respectively, to three-, four-, and five-fold divisions of each dimension), and the number of occupied cells is rarefied with respect to the number of species sampled (fig. 8). Rarefaction proceeds exactly as for the morphological range, except that it is the number of occupied cells in morphospace, rather than the total range, that is determined for each random draw. The Permian curve lies either below or above that for the Mis-
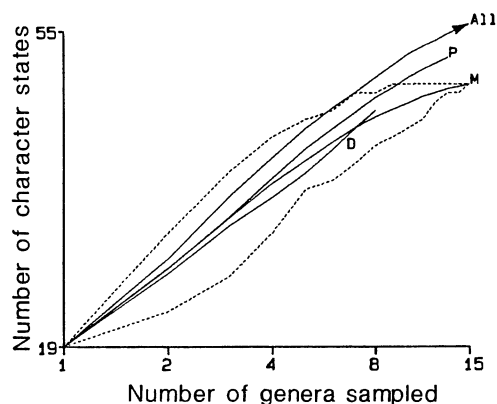

FIGURE 7. Rarefaction of number of character states against number of genera (N) of fissiculate blastoids. Data from Breimer and Macurda (1972). Abbreviations as in figure 5. (Silurian omitted, because there are only two genera, both with the same character states.) Dashed curves show 90% confidence interval for Mississippian rarefaction curve. Curve labeled "All" is for all fissiculates and continues in direction of arrow, ending at 38 genera and 59 character states. Similarity among these rarefaction curves, contrasted with greater differences among those in figure 5B, suggests that changes in morphological variety in the Fissiculata involved evolution of continuous morphometric variation more than changes in the number of character states.

sissippian, depending on the number of cells into which the morphospace is divided. Therefore, the diversity structures of these two stratigraphic intervals, as measured here, cannot be distinguished. On the other hand, as would be expected from the morphological distribution in figure 4, the Devonian curve falls well below those of the Mississippian and Permian. Thus, diversity structure as measured here appears to change significantly in the evolutionary history of the Blastoidea. These rarefaction curves should not be compared directly to those in figure 2.

Discussion

Any analysis of morphological variety depends on the number and kind of morphological characters used. In the preceding examples, the choices have been justified (explicitly or implicitly) by the various authors on grounds of taxonomic utility (Breimer and Macurda 1972), the ability to capture the major features of morphology and morphological variability (Saunders and Swan 1984; Foote 1991a,b), or agreement between
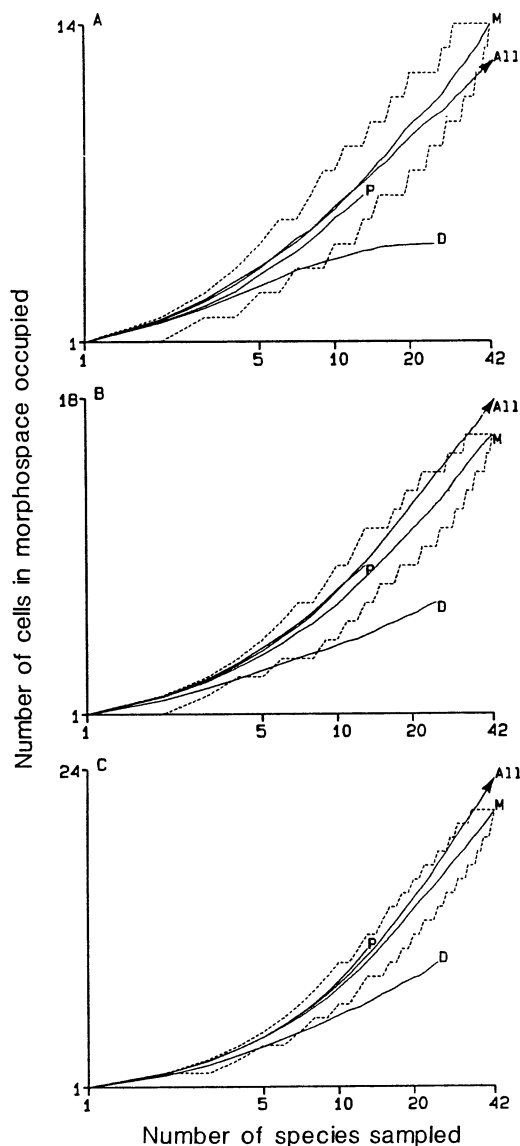
FIGURE 8. Rarefaction of number of occupied regions of equal volume (cells) in blastoid morphospace. Abbreviations and confidence limits as in figure 5. Data as in figure 4. Silurian omitted because of small sample size. A, Three-fold division of each PC; B, four-fold division; C, five-fold division. Curves labeled "All" are for entire sample and continue in direction of arrow, ending at 85 species and 16 (A), 24 (B), and 33 (C) cells. In contrast to figures 5 and 7, this analysis suggests that diversity structures of the Permian and Mississippian are not easily distinguished, whereas those of the Devonian and Mississippian are.

results obtained with the chosen measurement of morphology and results obtained in other ways (Foote 1991a). It should also be noted that the foregoing results based on cer-

tain numbers of principal components are broadly similar to those obtained when different numbers of dimensions are used.

As ecological and molecular diversity are also expected to depend on the number of species sampled, rarefaction of these aspects of diversity should also be fruitful. For example, Jackson and McKinney (1990: p. 174) mention the possibility that the observed increase in the number of life habits of clams since the Paleozoic may simply reflect the fact that the number of species of clams has increased, although they do not come out in favor of this possibility. They then suggest that potential bias in temporal comparisons may be circumvented by using similar sample sizes.

As Raup (1975) emphasized, purely preservational changes in sample size represent but one potential bias in the assessment of diversity. For example, in both the ammonoids and blastoids, differences in geographic representation are also involved. In the latter case, it was previously shown that geographic distribution plays a secondary role, after geologic age, in determining morphological variety (Foote 1991b). In results not presented here, geographically restricted samples of blastoids yield rarefaction curves that do not differ appreciably in shape from those of global samples.

Explicit consideration of the relationship between sample size and morphological variety may reflect the desire either to correct for sample size or to investigate the structure of diversity (or both). Coincidence of rarefaction curves (e.g., fig. 6) does not imply that observed differences in morphological variety are not biologically meaningful. Rather, it implies that the observed differences may be the result of drawing different numbers of species from a distribution with the same diversity structure. Since it has been suggested that change in morphological variability within a clade may be a common pattern in the history of life (Gould 1988), it is important to determine in any particular case whether such a pattern might result from a change in the number of species sampled. Rarefaction can correct for differences in sample size, but it is a separate question whether

such differences represent artifacts of preservation or sampling, or true changes in richness. If a true increase in richness is not accompanied by a change in diversity structure, then some constraint on the evolution of new forms may be indicated, such as the functional constraints suggested by Saunders and Swan (1984; Swan and Saunders 1987).

It is often useful to have a single measure of diversity that can allow comparisons among different taxonomic groups or different periods in the evolution of a group. An appropriate choice of diversity measure is not always clear. In the case of the blastoids, generic richness, morphological variance, and the three assessments of diversity structure presented above all suggest somewhat different patterns (Foote 1991b). I previously focused on morphological variance, because it could easily be compared with the expectation of a null branching model (Foote 1991b). However, it may generally be preferable not to rely on a single measure of diversity or diversity structure. Neither the morphological range, number of character states, nor number of occupied regions in morphospace is the "right" or "wrong" measure of variety. Each reflects a different aspect of morphological variety in much the same way as richness and evenness measure different aspects of diversity in communities. It is the explicit comparison among the different aspects of diversity that may be of most interest. For example, it is noteworthy that differences in morphological variety in fissiculate blastoids may reflect evolution of continuous variation in thecal shape more than changes in the number of distinguishable states of anatomical characters.

The form of rarefaction curves is often clearly related to an intuitive perception of diversity that includes both richness and morphological variety. Because large-scale studies of morphological variety have not been carried out for most groups, the examples that follow are somewhat speculative. Shallow, short curves represent taxa that are generally not considered very diverse in any respect (e.g., Monoplacophora within Mollusca). Long, shallow curves represent groups that are commonly considered diverse because of their high richness, but whose subtaxa are variations on a limited number of morphological themes (e.g., Ptychopariida within Trilobita). Long, steep curves represent groups that are generally considered very diverse, containing both many subtaxa and great morphological variety (e.g., Gymnolaemata within Bryozoa). Short, steep curves represent groups that are not commonly considered diverse because of their low richness, but which occupy much morphospace per unit of taxonomic richness (e.g., Homalozoa within Echinodermata). The diversity of taxa that lie between these extremes may be more difficult to assess. This difficulty reflects the lack of a single measure encompassing all meaningful aspects of diversity. Assessment of diversity is further complicated by the ambiguity with which one may interpret observed richness or sample size.

Because biological diversity can be measured in so many ways, none of which has clear preeminence, it is important to define diversity explicitly in any study. Even when it is clear that taxonomic richness is of interest, it matters whether higher-level or lower-level richness is the focus. In documenting morphological diversity, one may concentrate on the number of body plans, the differences among them, the variation within them, or some other measure. Ecological diversity may be measured by the number of guilds, the number of species within them, or ordination along a continuum representing some dimensions of adaptive space. I have presented one way in which two principal aspects of diversity—morphological variety and taxonomic richness—can be considered jointly. It is tempting to rank rarefaction curves in ascending order of steepness and say that these reflect a meaningful ranking from least to most diverse. Although such a linear ranking would sometimes be informative, it would often defeat one of the very purposes of rarefaction, namely, the simultaneous consideration of different aspects of diversity.

Acknowledgments

the U.S. National Museum; F. d'Escrivan and R. C. Eng at the Museum of Comparative Zoology (Harvard University); and R. D. White at the Yale Peabody Museum. B. Miljour drafted figure 3. B. E. Bodenbender and D. C. Fisher provided technical help with blastoid measurement. I thank D. C. Fisher, P. D. Gingerich, S. J. Gould, D. B. Macurda, Jr., D. M. Raup, J. J. Sepkoski, Jr., and R. E. Tabachnick for discussion. W. D. Allmon, P. D. Gingerich, D. J. Miller, and a reviewer made helpful comments on the manuscript. Acknowledgment is made to the Donors of the Petroleum Research Fund, administered by the American Chemical Society, for partial support of this research.

Literature Cited

Ashton, J. H., and A. J. Rowell. 1975. Environmental stability and species proliferation in Late Cambrian trilobite faunas: a test of the niche-variation hypothesis. Paleobiology 1:161–174.

Ausich, W. I., and D. J. Bottjer. 1982. Tiering in suspension feeding communities on soft substrata throughout the Phanerozoic. Science (Washington, D.C.) 216:173–174.

Bambach, R. K. 1983. Ecospace utilization and guilds in marine communities through the Phanerozoic. Pp. 719–746 in M.J.S. Tevesz and P. M. McCall, eds. Biotic interactions in Recent and fossil benthic communities. Plenum, New York.

———. 1985. Classes and adaptive variety: the ecology of diversification in marine faunas through the Phanerozoic. Pp. 191–243 in J. W. Valentine, ed. Phanerozoic diversity patterns: profiles in macroevolution. Princeton University Press and Pacific Division, American Association for the Advancement of Science, Princeton, New Jersey and San Francisco.

Bottjer, D. J., and W. I. Ausich. 1986. Phanerozoic development of tiering in soft substrata suspension-feeding communities. Paleobiology 12:400–420.

Breimer, A., and D. B. Macurda, Jr. 1972. The phylogeny of the fissiculate blastoids. Verhandelingen der Koninklijke Nederlandse Akademie van Wetenschappen, Afdeling Natuurkunde, Erste Reeks 26:1–390.

Broadhead, T. W. 1984. Macurdablastus, a middle Ordovician blastoid from the southern Appalachians. University of Kansas Paleontological Contributions, Paper 110:1–10.

Droser, M. L., and D. J. Bottjer. 1989. Ordovician increase in extent and depth of bioturbation: implications for understanding early Paleozoic ecospace utilization. Geology 17:850–852.

Erwin, D. H. 1990. Carboniferous-Triassic gastropod diversity patterns and the Permo-Triassic mass extinction. Paleobiology 16:187–203.

Fisher, D. C. 1986. Progress in organismal design. Pp. 99–117 in D. M. Raup and D. Jablonski, eds. Patterns and processes in the history of life. Springer, Berlin.

Foote, M. 1989. Perimeter-based Fourier analysis: a new morphometric method applied to the trilobite cranidium. Journal of Paleontology 63:880–885.

———. 1991a. Morphologic patterns of diversification: examples from trilobites. Palaeontology 34:461–485.

———. 1991b. Morphological and taxonomic diversity in a clade's history: the blastoid record and stochastic simulations. Contributions from the Museum of Paleontolgy, University of Michigan 28:101–140.

Fortey, R. A., and R. M. Owens. 1990. Trilobites. Pp. 121–142 in K. J. McNamara, ed. Evolutionary trends. University of Arizona Press, Tucson.

Gould, S. J. 1988. Trends as changes in variance: a new slant on progress and directionality in evolution. Journal of Paleontology 62:319–329.

Haldane, J.B.S. 1949. Suggestions as to the quantitative measurement of rates of evolution. Evolution 3:51–56.

Harland, W. B., R. L. Armstrong, A. V. Cox, L. E. Craig, A. G. Smith, and D. G. Smith. 1990. A geologic time scale 1989. Cambridge University Press, New York.

Harrington, H. J. 1959. Classification. Pp. O145–O170 in R. C. Moore, ed. Treatise on invertebrate paleontology, Part O, Arthropoda 1. The Geological Society of America and The University of Kansas Press, Boulder, Colorado and Lawrence, Kansas.

Jackson, J.B.C., and F. K. McKinney. 1990. Ecological processes and progressive macroevolution of marine clonal benthos. Pp. 173–209 in R. M. Ross and W. D. Allmon, eds. Causes of evolution: a paleontological perspective. University of Chicago Press, Chicago.

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

———. 1976. Species diversity in the Phanerozoic: a tabulation. Paleobiology 2:279–288.

———. 1978. Rarefaction. Pp. 80–89 in D. M. Raup and T.J.M. Schopf. Species as particles in space and time. Proceedings of a workshop held at the U.S. National Museum, Smithsonian Institution, Washington, D.C., June 5–16, 1978.

Raup, D. M., and S. J. Gould. 1974. Stochastic simulation and evolution of morphology—towards a nomothetic paleontology. Systematic Zoology 23:305–322.

Runnegar, B. 1987. Rates and modes of evolution in the Mollusca. Pp. 39–60 in K.S.W. Campbell and M. F. Day, eds. Rates of evolution. Allen and Unwin, London.

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

Saunders, W. B., and A.R.H. Swan. 1984. Morphology and morphologic diversity of mid-Carboniferous (Namurian) ammonoids in time and space. Paleobiology 10:195–228.

Sokal, R. R., and F. J. Rohlf. 1981. Biometry. Second edition. W. H. Freeman, San Francisco.

Stanley, S. M. 1979. Macroevolution: pattern and process. W. H. Freeman, San Francisco.

Swan, A.R.H., and W. B. Saunders. 1987. Function and shape in late Paleozoic (mid-Carboniferous) ammonoids. Paleobiology 13:297–311.

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

Van Valen, L. 1974. Multivariate structural statistics in natural history. Journal of Theoretical Biology 45: 235–247.

Ward, P. 1980. Comparative shell shape distributions in Jurassic-Cretaceous ammonites and Jurassic-Tertiary nautilids. Paleobiology 6:32–43.

Waters, J. A. 1988. The evolutionary palaeoecology of the Blastoidea. Pp. 215–233 in C.R.C. Paul and A. B. Smith, eds. Echinoderm phylogeny and evolutionary biology. Clarendon, Oxford.

Whittington, H. B. 1966. Phylogeny and distribution of Ordovician trilobites. Journal of Paleontology 40:696–737.