ARTICLES

Phanerozoic marine diversity and the fossil record

J. John Sepkoski Jr*, Richard K. Bambach†, David M. Raup‡ & James W. Valentine§

* Department of the Geophysical Sciences, University of Chicago, 5734 South Ellis Avenue, Chicago, Illinois 60637, USA
† Department of Geological Sciences, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061, USA
‡ Field Museum of Natural History, Roosevelt Road at Lake Shore Drive, Chicago, Illinois 60605, USA
§ Department of Geological Sciences, University of California, Santa Barbara, California 93106, USA

Strong correlations between various local and global estimates of Phanerozoic marine diversity for taxa below the ordinal level indicate a single pattern of change underlying all data on fossil density. Geological time alone seems insufficient to explain all of the significant covariation among the data sets, and it is proposed that the common pattern in diversity reflects the signal from a real evolutionary phenomenon strong enough to overcome the biases inherent in the fossil record.

THE 'imperfection of the fossil record' has always been the single greatest impediment in the palaeontological study of evolution. Incomplete preservation, inadequate sampling and arbitrary taxonomy have made it difficult to discern and assess accurately large-scale evolutionary patterns in the history of life[1-3]. For the study of taxonomic diversity, these problems have led to uncertainty about both the patterns and the processes of diversification in terrestrial and especially marine ecosystems. No single data set has even been able to demonstrate conclusively whether species diversity in the oceans has increased substantially since the Palaeozoic era or whether it has remained essentially constant over the past 500 million years[4,5]. However, when all empirical estimates of Phanerozoic marine diversity are considered together, a remarkably consistent pattern emerges, suggesting that a strong evolutionary signal underlies the preservational and taxonomic noise of the known fossil record.

Phanerozoic diversity

Five major and essentially independent estimates of the diversity of lower taxa in the marine fossil record have been published during the past decade:

(A) Trace fossil diversity—Seilacher[6] plotted data on observed numbers of ichnospecies in marine facies of various ages at 36 localities from around the world.

(B) Species per Myr—Raup[7] published data on global numbers of fossil invertebrate species compiled from a '10% sampling' of the 1900–70 volumes of the *Zoological Record*.

(C) Species richness—Bambach[4], in attempt to avoid biases inherent in estimates of global fossil diversity, compiled data on the numbers of invertebrate species described from 386 marine, level-bottom communities of Cambrian to Quaternary age.

(D) Generic diversity—Raup[8], for an analysis of taxonomic survivorship, compiled data on the times of origination and extinction of all genera and subgenera listed in the pre-1978 volumes of the *Treatise on Invertebrate Paleontology*[9].

(E) Familial diversity—Sepkoski[10] published a graph showing the stage-by-stage history of the global diversity of all marine metazoan families, compiled from some 300 literature sources.

These five sets of data on Phanerozoic marine diversity (Fig. 1) have been reduced to the common stratigraphic resolution of this geological system and plotted on a normalized vertical scale. The data for trace fossil diversity may be more heterogeneous than the other four estimates; Seilacher[6] has argued that trace fossil communities in shallow and deep water are decoupled and hence have had separate histories of diversification. We have combined his data from both neritic (shelf) and flysch (continental slope and rise) facies, however, so that trace fossils as a whole can be compared with data sets on skeletal body fossils.

Correlation analysis

The extent to which all five estimates of fossil diversity are influenced by a common underlying factor can be assessed using simple correlation analysis. Table 1 lists linear correlation coefficients for all pairs of the diversity estimates. All the correlations are quite high, with 8 out of 10 ranging from 0.91 to 0.99. The only two coefficients lower than 0.90 involve the data on trace fossils. These data, as well as being heterogeneous, measure local behavioural diversity, thereby only indirectly reflecting local species diversity[6], and encompass fewer observations than the other data sets. In contrast, Raup's two data sets, which are based on the greatest numbers of observations, exhibit the highest average correlations.

If standard one-sided statistical tests are applied to the correlation coefficients in Table 1, all appear to be significantly greater than zero at the 0.05 confidence level (and all but the lowest at the 0.001 level). However, these test results cannot be considered valid. Because taxonomic diversity in any interval of geological time is partially dependent on the diversity in the preceding interval[11], the usual statistical assumptions regarding independence of samples are violated[12,13]. To avoid this difficulty, correlations independent of time must be examined.

That all five estimates of Phanerozoic diversity are strongly correlated with time (Table 2) reflects in part the necessary increase in diversity from the earliest Phanerozoic to the present. However, in no case does this correlation with time account for more than ~60% of the variance in estimated diversity. When this correlation is removed, the residual estimates of diversity remain highly intercorrelated, as indicated by both the linear and the rank correlation coefficients in Table 2. One-sided statistical tests indicate that all but the lowest one or two coefficients in either set are significantly greater than zero at the 0.05 level or better.

This last result is important when the variety of uncertainties and biases inherent in each set of data is considered. For

Table 1 Matrix of linear product–moment correlation coefficients between estimates of standing marine diversity in each Phanerozoic geological system

		A	B	C	D	E
A	Trace fossils	(9)	—				
B	Species per Myr	(10)	0.907	—			
C	Species richness	(8)	0.725	0.929	—		
D	Generic diversity	(10)	0.927	0.986	0.912	—	
E	Familial diversity	(10)	0.871	0.928	0.927	0.939	—

Numbers of systems with data are listed in parentheses. The first relative eigenvalue ($\lambda_1/5$) of this matrix is 0.925, indicating a strong overall correlation among the five diversity estimates.
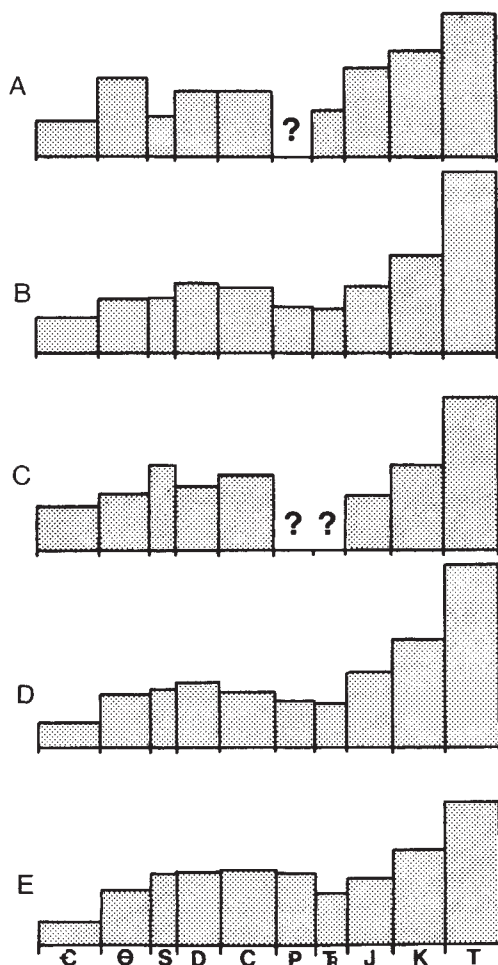
Fig. 1 Histograms displaying the relative magnitudes of marine animal diversity per geological system as indicated in five data sets published during the last decade. The geological systems are denoted with standard symbols beneath the bottom histogram. The heights of the histogram bars have been scaled so that each data set shows the same average diversity. Question marks signify systems for which adequate data are lacking. The data were compiled by the following procedures: (A) Seilacher's[6] data on trace fossils were averaged by calculating median numbers of ichnospecies per locality in each system within both neritic and flysch facies (with a more recently published[29] data point for the Jurassic included); no data were available for the Permian system. (B) Raup's[7] raw data on numbers of species per geological system were modified by first eliminating the non-marine Insecta and then normalizing the remaining data by dividing by the duration of each system. (C) Bambach's[4] data on species richness were taken directly from his Fig. 5, which lists median numbers of species in communities in 'open marine environments'; these communities, which constitute the bulk of Bambach's total data set as well as of the entire marine fossil record, comprise the only data that Bambach listed at the level of system (with adequate data for the Permian and Triassic systems lacking). (D) The data on generic diversity were reduced from Raup's[8] table of numbers of generic originations and extinctions per geological system using a modified Webb–Harper transformation[30,31]: average standing diversity in each system was calculated as the number of genera passing entirely through the system, plus half the new genera surviving beyond the system, plus half the old genera becoming extinct in the system, plus the number of genera confined to the system multiplied by a factor scaling for the system's duration. This scaling factor was computed as 14 Myr divided by the duration of the system, where 14 Myr (or half the average duration of a genus[8]) was arbitrarily chosen assuming that genera confined to single systems should have durations somewhat shorter than average. (Experimentation with numerators ranging from 10 to 20 Myr altered results by <10%.) (E) The data on familial diversity were generated by weighting the standing diversity of families in each stratigraphic stage in Sepkoski's[10] original data by the stage's duration and then averaging over each system to produce a simple weighted mean.

example, not only are the data on trace fossil diversities heterogeneous with respect to environment, but they may also be heterogeneous with respect to taxonomic level. Some ichnospecies may even represent mixtures of phyla, such as species of *Skolithos* which may have been produced by both phoronids and annelids. The data on species richness were derived from a more uniform set of open shelf environments but may still reflect considerable environmental variation; species richness in benthic communities today varies by more than an order of magnitude within analogous habitats in different climatic zones. Such variation may be confounded further in the fossil record by enrichment of local diversity by lateral mixing[14] and time averaging[15]. The compilations of global diversities of fossil species, genera, and families avoid some of these local problems but are subject to a wide range of additional biases which result from variations in the quantity and quality of the geological record, reliability and accessibility of taxonomic literature, number of active workers, and so on[3,16]. The effects of these biasing factors may vary considerably with taxonomic level. Also important is the differing taxonomic coverage of the various data sets; the familial data in Fig. 1, for example, exclude protozoans, which are incorporated into the species and generic data, but include marine vertebrates, which do not appear in the other two data sets.

Patterns

The strong similarities among the data sets lead us to conclude that all five estimates of Phanerozoic taxonomic diversity are measuring a single underlying pattern that is reflected at both local and global scales and at various taxonomic levels in the marine fossil record. Judging from the elements shared by all data sets, this common pattern is: (1) low diversity during the Cambrian period; (2) higher but not persistently increasing diversity through the post-Cambrian portion of the Palaeozoic era; (3) low diversity in the early Mesozoic era (especially the Triassic period); (4) increasing diversity through the Mesozoic leading to a Cenozoic (Tertiary) maximum.

Bambach[4] recognized essentially the same pattern in his observation of 'multiple equilibria' in his data on species richness. Differences principally involve transitional diversities: he placed the Ordovician, which is probably transitional between the Cambrian and later Palaeozoic diversities[10,11], with the Cambrian, and he included the Jurassic and Cretaceous, which are transitional between the Triassic low and the Cenozoic high, with the Palaeozoic. Similar patterns were also indirectly recognized by Raup[17] and especially Sepkoski[18,19] in analyses of sampling biases in Raup's[7] species data. Multiple regressions of species diversity on measures of sampling efficiency (specifically global sedimentary rock volume or outcrop area) and of expanse of epicontinental seas generated consistently negative residuals for Cambrian and Triassic diversities and positive residuals for Tertiary diversity; these residuals correspond to the lows and high reflected in all data sets in Fig. 1.

Table 2 Correlations removing the influence of time from the estimates of Phanerozoic marine diversity

	Time	A	B	C	D	E
A Trace fossils	−0.784	—	0.667	(0.214)	0.611	(0.444)
B Species per Myr	−0.708	0.783	—	0.571	0.867	0.644
C Species richness	−0.683	(0.352)	0.885	—	0.714	0.714
D Generic diversity	−0.771	0.805	0.978	0.888	—	0.689
E Familial diversity	−0.767	0.676	0.851	0.900	0.850	—

The column labelled 'Time' lists linear product–moment correlation coefficients between the diversity estimates and the ages of the geological systems, measured as the time from the Recent (0 Myr) to each system's midpoint. The matrix to the right lists partial linear correlations (below) and Kendall's rank correlations (above), obtained by correlating residuals from linear regressions of diversity estimates on time. Coefficients enclosed in parentheses are not significantly greater than zero at the 0.05 confidence level.

Older estimates of Phanerozoic taxonomic diversity also exhibit patterns commensurate with the data sets in Fig. 1. For example, Sepkoski's data set on familial diversity is pre-dated by comparable series-level data published by Newell[20] and Valentine[21]. These latter data sets, which were based on an older taxonomic and stratigraphical literature, still correlate highly with Sepkoski's familial data at the series level; Newell's data (with terrestrial vertebrates removed using information from Romer[22]) correlate with a magnitude of 0.934 (partial $r = 0.880$ with time removed), and Valentine's data correlate with a magnitude of 0.953 (partial $r = 0.942$). Even the oldest compilations of diversity in the fossil record bear strong resemblance to the modern estimates. Data taken from the Phanerozoic diversity curve published by Phillips[23] in 1860 possess an average correlation of 0.869, and an average partial correlation with time removed of 0.693, with the five data sets in Fig. 1. These relationships suggest that while increased systematic and stratigraphic knowledge may have enhanced the accuracy and resolution of diversity data (see, for example, ref. 24), time has not substantially altered the fundamental patterns seen in the data.

The only estimates of fossil marine diversity that do not show patterns like those in Fig. 1 involve taxonomic categories above the familial level. Ordinal data sets published by Simpson[25], Valentine[21] and Sepkoski[26] show low numbers of orders in the Cambrian and no strong trends through the later Palaeozoic, but none shows a marked increase in diversity in the post-Palaeozoic; in fact, Valentine's data show a major decline in numbers of orders following the Palaeozoic. The reasons for these differences are not entirely clear but may involve the tendency for morphologically distinct groups of animals to appear rapidly during the early phases of major evolutionary radiations[2,27]; such groups tend to become extinct rather quickly if they fail to diversify internally[10,28]. Thus, as Valentine[21] has argued, the magnitude of post-Cambrian diversity portrayed in the ordinal data sets may indicate more the variety of distinct morphological types produced during the early Palaeozoic radiations rather than the actual patterns of low-level taxonomic diversity that seem to be reflected in Fig. 1.

1. Newell, N. D. J. Paleontol. 33, 488–499 (1959).
2. Simpson, G. G. in Evolution after Darwin Vol. 1 (ed. Tax, S.) 117–180 (University of Chicago Press, 1960).
3. Raup, D. M. Science 177, 1065–1701 (1972).
4. Bambach, R. K. Paleobiology 3, 152–167 (1977).
5. Signor, P. W. III Paleobiology 4, 394–406 (1978).
6. Seilacher, A. Neues Jb. Geol. Paläont. Mh. H4, 233–245 (1974); in Patterns of Evolution (ed. Hallam, A.) 359–376 (Elsevier, Amsterdam, 1977).
7. Raup, D. M. Paleobiology 2, 279–288 (1976).
8. Raup, D. M. Paleobiology 4, 1–15 (1978).
9. Moore, R. C., Teichert C. & Robison, R. A. (eds) Treatise on Invertebrate Paleontology (Geological Society of America and University of Kansas Press, Lawrence, 1953–1979).
10. Sepkoski, J. J. Jr Paleobiology 7, 36–53 (1981).
11. Sepkoski, J. J. Jr Paleobiology 5, 222–251 (1979).
12. Yule, G. V. J. R. stat. Soc. 89, 1–64 (1926).
13. Raup, D. M. in Patterns of Evolution (ed. Hallam, A.) 59–78 (Elsevier, Amsterdam, 1977).
14. Valentine, J. W. & Mallory, B. J. Geol. 73, 683–701 (1965).
15. Peterson, C. H. Lethaia 9, 137–148 (1976).
16. Raup, D. M. Bull. Carnegie Mus. nat. Hist. 13, 85–91 (1979).
17. Raup, D. M. Paleobiology 2, 289–297 (1976).
18. Sepkoski, J. J. Jr Paleobiology 2, 298–303 (1976).
19. Flessa, K. W. & Sepkoski, J. J. Jr Paleobiology 4, 359–366 (1978).
20. Newell, N. D. Scient. Am. 208, 76–92 (1963); Geol. Soc. Am. Spec. Pap. 89, 63–91 (1967).
21. Valentine, J. W. Palaeontology 12, 684–709 (1969).
22. Romer, A. S. Vertebrate Paleontology 3rd edn (University of Chicago Press, 1966).
23. Rudwick, M. J. S. The Meaning of Fossils, 240 (Macdonald, London, 1972).
24. Grant, R. E. J. Paleontol. 54, 499–507 (1980).
25. Simpson, G. G. Brookhaven Symp. Biol. 22, 162–177 (1969).
26. Sepkoski, J. J. Jr Paleobiology 4, 223–251 (1978).
27. Valentine, J. W. Palaeontology 6, 444–450 (1980).
28. Raup, D. M. J. Paleontol. 52, 517–523 (1978).
29. Frey, R. W. & Seilacher, A. Lethaia 13, 183–208 (1980).
30. Webb, S. D. Evolution 23, 688–702 (1969).
31. Harper, C. W. Jr J. Paleontol. 49, 752–757 (1975).

Palaeontological documentation of speciation in Cenozoic molluscs from Turkana Basin

P. G. Williamson

Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts 02138, USA

A recently discovered series of mollusc faunas from the late Cenozoic of the eastern Turkana Basin constitutes one of the best documented metazoan fossil sequences. Evolutionary patterns in all lineages conform to the 'punctuated equilibrium' model; no 'gradualistic' morphological trends occur. These faunas provide the first fine-scaled palaeontological resolution of events during speciation: fundamental phenotypic transformation of both sexual and asexual taxa occurs rapidly, in comparatively large populations, and is accompanied by a significant elevation of phenotypic variance. This increase in variance reflects extreme developmental instability in the transitional populations.

THE 400 m sequence of late Cenozoic deposits east of Lake Turkana in northern Kenya, discovered[1] by Leakey in 1968, comprises the Plio-Pleistocene Kubi Algi, Koobi Fora and Guomde Formations, and the Holocene Galana Boi Beds[2]. In addition to important palaeoanthropological[1], archaeological[3] and vertebrate palaeontological[4] material, these deposits have recently yielded a uniquely well documented sequence of lacustrine mollusc faunas. The faunas have important implications for present evolutionary controversies, as they provide the first detailed palaeontological documentation of events during allopatric speciation.

Mollusc faunas are scattered throughout 1,000 km of exposures east of Lake Turkana and occur in laterally extensive lensoid accumulations 0.001–1 m thick. These accumulations have a matrix ranging from coarse silt to coarse sand grade, and are separated by finer-grained intervals devoid of molluscs. The 190 faunas reported here consist of various prosobranch, pulmonate and bivalve lineages, and represent both life and death assemblages in various shallow lacustrine and pro-deltaic settings.

Various features make this sequence particularly useful for investigating evolutionary patterns: the molluscs are well preserved and abundant, the units in which they occur are generally unconsolidated, and because most of the species lineages in the section are still extant, reasonable inferences can be drawn concerning the 'soft' biology of their fossil representatives from the Turkana Basin sequence. In biological and taxonomic terms, these mollusc lineages comprise an extremely heterogeneous assemblage (see Table 1); this heterogeneity allows useful comparisons to be made of evolutionary patterns in taxa varying widely in autecology, reproductive strategy and size. In particular, evolutionary patterns in both sexual and