Phanerozoic Trends in the Global Diversity of Marine Invertebrates
John Alroy, et al.
Science 321, 97 (2008);
DOI: 10.1126/science.1156963

The following resources related to this article are available online at www.sciencemag.org (this information is current as of July 7, 2008):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/cgi/content/full/321/5885/97

Supporting Online Material can be found at:
http://www.sciencemag.org/cgi/content/full/321/5885/97/DC1

A list of selected additional articles on the Science Web sites related to this article can be found at:
http://www.sciencemag.org/cgi/content/full/321/5885/97#related-content

This article cites 32 articles, 20 of which can be accessed for free:
http://www.sciencemag.org/cgi/content/full/321/5885/97#otherarticles

This article appears in the following subject collections:
Paleontology
http://www.sciencemag.org/cgi/collection/paleo

Information about obtaining reprints of this article or about obtaining permission to reproduce this article in whole or in part can be found at:
http://www.sciencemag.org/about/permissions.dtl
Phanerozoic Trends in the Global Diversity of Marine Invertebrates


It has previously been thought that there was a steep Cretaceous and Cenozoic radiation of marine invertebrates. This pattern can be replicated with a new data set of fossil occurrences representing 3.5 million specimens, but only when older analytical protocols are used. Moreover, analyses that employ sampling standardization and more robust counting methods show a modest rise in diversity with no clear trend after the mid-Cretaceous. Globally, locally, and at both high and low latitudes, diversity was less than twice as high in the Neogene as in the mid-Paleozoic. The ratio of global to local richness has changed little, and a latitudinal diversity gradient was present in the early Paleozoic.

Diversity curves showing changes through the geological record in the number of fossil marine genera or families have fueled decades of macroevolutionary research (1–10). Traditionally, these curves were based on literature compilations that recorded only the first and last appearances of taxa (3, 5). These compilations suggested that diversity rapidly rose during the Cambrian and Ordovician and then either stayed at a plateau (3) or declined erratically (5) through the Paleozoic. There was a modest rebound after the end-Permainian mass extinction, the largest (11) and most ecologically important (9, 12) of the Phanerozoic. Diversity in these curves then rose steadily and with a possibly increasing absolute rate, suggesting to some an exponential reaction (4, 6, 9, 10).

The appearance of a Paleozoic equilibrium followed by a nearly unbridled Mesozo-Cenozoic radiation has presented a puzzle: How could global diversity reach limits and then much later cast them off, rising to far higher levels than those seen during the Paleozoic? The possibility that the apparent radiation was exaggerated by secular trends in the quality and especially quantity of the preserved fossil record (8, 13–15) was first proposed more than three decades ago (16). This claim was at first put aside because sampling efforts were thought to be minor and could not be assessed without more detailed information (2, 8).

The Paleobiology Database (http://paleodb.org) makes it possible to address the problem with the use of contemporary statistical methods because it records occurrences of genera and species within particular fossil collections. An intensive data collection effort has quadrupled the database since data for two long Phanerzoic intervals were presented in 2001 (17). This initiative has focused on both filling gaps and sampling the Cenozoic at a high level (18). The data set includes 44,446 collections with individually recorded ages and geographic coordinates. The collections comprise 284,816 fossil occurrences of 18,702 genera that equate with the geological record in the number of fossil marine genera or families have fueled decades of macroevolutionary research (1–10). Traditionally, these curves were based on literature compilations that recorded only the first and last appearances of taxa (3, 5).

Sampling and counting. The amount of data per time period and, therefore, the shape of a diversity curve may vary greatly as a result of uneven preservation and sampling effort. The key advantage of collection data is that this variation can be removed by subsampling (17). A random subset of the available collections is drawn until each interval, called a sampling bin, includes the same estimated number of specimens. Genera are counted, and the procedure is repeated to obtain averages. We tallied actual specimen counts when available and otherwise estimated them using a gently curved, one-parameter empirical function that relates the logarithms of specimen and geographic coordinates. The collections comprise 284,816 fossil occurrences of 18,702 genera that equate with the geological record in the number of fossil marine genera or families have fueled decades of macroevolutionary research (1–10). Traditionally, these curves were based on literature compilations that recorded only the first and last appearances of taxa (3, 5).

local depression of curves near extinction and origination events [the Signor-Lipps effect (23)], and interpolated presences of polyphyletic ("wastebasket") genera that have artificially long ranges with large gaps connecting unrelated species. Our direct documentation of unrelated species. Our direct documentation of changes in global diversity (thin line, same as in Fig. 1) are shown for comparison.

Global diversity. The sampling-standardized diversity curve (Fig. 1) shows many key features of older curves (e.g., 5), such as the Cambrian-Ordovician radiation, 78% end-Permian extinction, and 63% end-Triassic extinction (18).

Fig. 1. Genus-level diversity of both extant and extinct marine invertebrates (metazoa less tetrapods) during the Phanerozoic, based on a sampling-standardized analysis of the Paleobiology Database. Points represent 48 temporal bins defined to be of roughly equal length (averaging 11 My) by grouping short geological stages when necessary. Vertical lines show the 95% confidence intervals based on Chernoff bounds, which are always conservative regardless of the number of genera that could be sampled or variation in their sampling probabilities (18). Data are standardized by repeatedly drawing collections from a randomly generated set of 65 publications until a quota of 16,200 specimens has been recovered in each bin. On average, 461 collections had to be drawn to reach this total. The curve shows average values found across 20 separate subsampling trials—enough to yield high precision with such large sample sizes. Ma, million years ago. Cm, Cambrian; O, Ordovician; S, Silurian; D, Devonian; C, Carboniferous; P, Permian; Tr, Triassic; J, Jurassic; K, Cretaceous; Pg, Paleogene; Ng, Neogene.

Fig. 2. Evenness estimates (thick line) for 11-My-long bins, based on Hurlbert's PIE index (27). The values are computed from the calibrated occurrence weights used in subsampling, and each is a weighted moving averages across five consecutive bins. Changes in global diversity (thin line, same as in Fig. 1) are shown for comparison.
index of evenness (27). PIE bears no necessary relation to the total number of taxa that might be sampled.

Evenness changed substantially through the Phanerozoic, if not as dramatically as sometimes suggested (29), and the greatest values are seen in the late Cretaceous and Cenozoic (Fig. 2). They imply, for example, that 200 specimens will yield about 11.3 genera in the latest Ordovician, 22.2 in the Paleocene, and 19.3 in the late Neogene. The general pattern of a long-term increase is confirmed by related (but more restricted) studies (26, 30). However, the evenness curve (Fig. 2) does not simply increase. It suggests a plateau between the Late Ordovician and Carboniferous and then a large rise through the Permian. After a weak recovery from the Permo-Triassic decline, it shows little change until a rise in the Late Cretaceous (18). All of these features are broadly consistent with the idea that local diversity reached its maximum sometime during the past 100 Myr (29) as the number of occupied niches expanded (31). However, they do not suggest a radical increase between the early Paleozoic and Cenozoic. They also do not mirror the shift in abundance distributions from simple to complex shapes at the Permo-Triassic boundary (12), because evenness dropped across this boundary instead of rising, and the Triassic figures are well within the narrow range seen throughout most of the Paleozoic.

Theoretically, if biogeographic and environmental gradients (i.e., beta diversity) do not change, global diversity should track local sampled diversity and, therefore, evenness. Furthermore, the global curve is methodologically grounded on the evenness data because the calibrated function estimates fewer specimens per collection when evenness is high, which causes more collections to be drawn. Hence, it is no surprise that after logging and differencing the global and local curves (Fig. 2), we find a significant correlation (Spearman's rank-order correlation ρ = 0.332, P = 0.023). Specific similarities include the parallel increase during the Cambro-Ordovician radiation, the joint rise during the Permian, and the drop in evenness during the severe global extinction at the Permo-Triassic boundary (9, 11, 12).

However, the curves depart from one another for long stretches of time. Evenness remained relatively high during the long late Paleozoic trough in global richness (Fig. 2). The offset may be driven by geography: The continents were widely dispersed during the early Paleozoic, late Mesozoic, and Cenozoic, and the late Paleozoic low roughly corresponds with the assembly of the supercontinent Pangaea. Thus, steeper biogeographic gradients during the late Phanerozoic may have accommodated greater global diversity (1), and the same may have been true during the early Paleozoic. Nevertheless, this hypothesis needs to be explored in more detail because the mid-Palaeozoic offset between the curves is not consistent, and nonbiogeographic factors such as onshore-offshore gradients also may have changed.

**Litudinal diversity gradients.** To ensure that the global curve captured any radiation in the tropics that might have occurred, we deliberately focused on the Neogene tropics while simultaneously establishing a baseline sampling level for the entire Phanerozoic. Counts of references in the Paleobiology Database, Sepkoski's compendium, GeoRef, and the Treatise on Invertebrate Paleontology suggest that our data sample the tropics as well as, or better than, those compilations (SOM text). Enough data are available to compare latitudinal belts within several key time intervals.

Subsampling curves for individual Neogene bins show that low-latitude diversity (within 30° palaeolatitude of the equator) was substantially higher than northern temperate zone diversity (Fig. 3A). The Neogene tropical curves overlap, and the high-latitude curves are also very similar to each other. In the Ordovician, there was little benthic habitat in the north, so we have computed low-latitude and southern temperate zone curves (Fig. 3B). They show a clear latitudinal diversity gradient for one bin but not the other. Thus, regardless of whether the modern gradient came into existence recently, at least some much older intervals did witness similar patterns. Furthermore, because there is a large difference between the Neogene and Ordovician at both high and low latitudes (Fig. 3), the moderate net increase through the Phanerozoic (Fig. 1) seems to have been a global phenomenon instead of being driven strictly by a radiation in the tropics.

**Previous data sets.** The fact that most treatments of older compilations depict a massive Cretaceous-Cenozoic radiation (1–7, 9, 10) raises the question of whether the differences are primarily methodological or primarily related to coverage of the literature. We answer this question by tabulating our data set and Sepkoski's range-based, genus-level compendium (5) in exactly the same way. Because it is impossible to either sample-standardize or take sampled-in-bin counts if only ranges are available (5), we use raw, unstandardized data and treat our genus age ranges as if they were continuously sampled.

The two data sets yield similarly shaped curves of a comparable magnitude (Fig. 4). The genus-level curves (Fig. 4) virtually overlap, and the same is true for ordinal-level data sets (SOM text). The new genus-level curve is higher than Sepkoski's throughout much of the Mesozoic but still suggests a large Cenozoic radiation, albeit smaller than that in Sepkoski's data. Our raw curve suggests a 3.74 times difference between the late Neogene and the median Ordovician, Silurian, or Devonian interval—slightly more than the factor of 3.54 seen in Sepkoski's compendium (Fig. 4). Both curves identify not just major but also minor features, such as a peak in the late Jurassic.

Thus, the dramatic differences between the standardized and conventional curves (Figs. 1 and 4) do not result from data-quality or -quantity...
problems. We instead attribute these differences to two biases that can only be removed with standardized sampling and sampled-in-bin counting of occurrence data (SOM text). First, we have exceeded the magnitude of Sepkoski's curve in the mid-Phanerozoic with a modest data set but fallen just short of it in the Cenozoic with a massive data set (18). Fully matching his curve would require us to make our sampling even more heterogeneous. Hence, the literature compiled by Sepkoski seems likely to contain a strong Cenozoic sampling bias. Second, counting each extant taxon as sampled everywhere from its last fossil appearance to the Recent [i.e., the Pull of the Recent (32)] exaggerates the artificial Tertiary and Cenozoic increase. This problem cannot affect counts of sampled taxa because they take no note of which genera are extant or which extend beyond a particular bin.

Conclusion. The new diversity curve (Fig. 1) records substantial volatility, including some potentially meaningful excursions that might have implications for evolutionary innovations, paleogeographic shifts, climate change, sea-level change, or other factors. In particular, the fact that local and global diversity have not always changed in tandem (Fig. 2) implies that compositional differences between environments or geographic regions have waxed and waned.

Regardless of whether this is true, a more general and important pattern is evident: Most of the Cenozoic radiation took place well before the Late Cretaceous and Cenozoic, and the net increase through the Phanerozoic was disproportionately small relative to the enormous amount of time that elapsed. Strong latitudinal diversity gradients as far back as the Ordovician and modest changes in local-scale sampled diversity (Fig. 2) are consistent with the suggestion that global biodiversity is constrained. Additional evidence includes rapid reboots from all of the major extinction episodes (Fig. 1) (3). Although any limit to diversity could not have been static and must have had a net increase (13, 16, 17, 33), our results cannot be reconciled with previous studies that argued for exponential long-term growth on the basis of raw, unstandardized data (1, 4, 6, 9, 10). Allowing for ecological reorganizations in the wake of mass extinction (9) or for the addition of ecological niches by means of evolutionary innovation (31, 34) does not explain how diversification could have been limited in the face of interval-to-interval changes that rivaled the entire net increase over the Phanerozoic.

Thus, we now must ask what mechanisms could have led to saturation. They may have involved the way that energy is captured from lower trophic levels. This capture could have remained roughly constant (35). Alternatively, if total energy capture did increase (36), this change may have been offset by the diversification of groups with high metabolic rates (9), making it energetically difficult for a large net radiation to occur.

References and Notes
8. A. I. Miller, Paleobiology 26, (suppl. 1), 53 (2000).
15. J. S. Crampton et al., Science 301, 358 (2003); published online 12 June 2003, 10.1126/science.1085075.
18. Materials and methods are available as supporting material on Science Online.
37. We thank R. Bambach and D. Jablonski for extensive and helpful discussions, and M. J. S. Tovez, D. Korn, K. Koverman, J. Millet, K. Nei, P. Novack-Gottshall, T. Schossleitner, C. Schweitzer, M. Uhen, P. Wall, and especially D. Hensh, S. Hicks, and U. Meldal for contributing to the data set. This work was funded by the NSF Division of Environmental Biology (grant DEB-0063893 to C.R.M., J.A., A.I.M., J. M. Hayes, and D. H. Rithmann), the NSF Sedimentary Geology and Paleontology program (grants EAR-0521425 and EAR-0545059 to J.A., EAR-0434443 to D.J.S., and EAR-0505469 to M.E.P.), the NASA Program in Exobiology and Evolutionary Biology (grants to A.I.M. and M.E.P.), and the VolkswagenStiftung (grant to W.K.). Our research was conducted while J.A. was a Center Associate at the National Center for Ecological Analysis and Synthesis, a Center funded by NSF (grant DEB-0553760), the University of California–Santa Barbara, and the state of California. This is Paleobiology Database publication number 80.