

New data for old questions

Kenneth G. Johnson

Department of Invertebrate Paleontology, Natural History Museum of Los Angeles County, 900 Exposition Boulevard, Los Angeles, California 90007. E-mail: KJohnson@nhm.org

Accepted: 25 July 2002

Global biodiversity dynamics result in large-scale patterns of change in the variety of life on earth. One deceptively simple measure of global biodiversity is the total count of species or supraspecific taxa that have been described from each subinterval of the Phanerozoic record. These tallies usually are obtained from data compilations based on extensive literature searches supplemented with examination of collections in museums (Sepkoski 1993; Alroy et al. 2001). Such measures of global biodiversity dynamics are probably not meaningless, but it is difficult to understand what they mean. The number of taxa alive on Earth at any one time is a function of innumerable ecological factors acting at the full range of temporal and spatial scales (Willis and Whittaker 2002). Changes in global richness might result from increasing local richness as local communities become more complex (Valentine 1969; Bambach 1977). At the regional scale, changing oceanographic conditions, climate, and the relative positions of continents and shallow seas might alter global taxonomic diversity by changing the size and arrangement of biogeographic units (Valentine 1970). Therefore, the best hope for understanding global diversity patterns is to collect information on local assemblages. These hierarchically organized data can then be integrated to test alternate hypothetical mechanisms of change. Data compiled as the endpoints of stratigraphic or geographic ranges are useless for this purpose. This has been known at least since Bambach (1977) explored patterns of change in alpha diversity of level-bottom marine communities.

Compiling lists of local assemblages is also the key to assessing the effects of an imperfect record. The actual distribution of fossils is strongly biased by nonrandom differences in

the probability of fossil preservation among environments, taxa, and time intervals (Raup 1972). Moreover, the subset of this record that has actually been studied is also strongly biased. For example, many more fossils have been collected from North America and Western Europe than from the rest of the world, and the few monographs that exist for whole regions are likely to be unrepresentative samples of the real fossil record. In the Neogene, generally only the most well-preserved faunas are chosen for systematic study, so faunas preserved in fine-grained siliciclastic settings have received more attention than the recrystallized or moldic fossils that can be recovered from typical shallow-water carbonates.

If properly designed, synthetic databases that are compiled from a variety of sources (for example Alroy et al. 2001) can help identify undersampled sites and regions for further collecting, but we already know that our knowledge of the Cenozoic Tropics is inadequate (Jackson and Johnson 2001). Numbers of molluscan genera and subgenera collected by the Panama Paleontology Project exceed totals from previous studies of Neogene mollusks, and similar numbers result whenever adequate sampling is attempted (Jackson et al. 1999). As collections improve, we now realize that the best-described regional faunas unfortunately have proved to be the lowest-diversity faunas. Collections exceeding 12,000 fossils were required to recover the high generic richness reported from the late Miocene Gatun Formation of Panama (Woodring 1957–1982), but greater richness was discovered in collections one-third to one-half as large from the diverse Pliocene deposits of the Bocas del Toro Basin (Jackson et al. 1999). This chronic undersampling is not limited to studies of mollusks. Previous synthesis of the strati-

graphic distribution of reef corals suggested that Pliocene Caribbean reef-coral diversity was similar to extant diversity (Frost 1977). But we now know from extensive new collecting that the extant Caribbean coral fauna is a depauperate relict of the Pliocene diversity (Budd et al. 1999; Jackson and Johnson 2000).

The biotic history of the tropical Indo-Pacific center of extant global diversity is even less known than that of the western Atlantic, although some studies suggest that the regional marine biota has been very diverse at least since the Pliocene (Veron and Kelley 1988). Certainly, diversity remains markedly underestimated for those few rich Neogene mollusk faunas of Indonesia that have been studied, and existing museum collections are inadequate to solve this problem (J. Todd personal communication February 2002). This is significant because strong latitudinal gradients are likely to have existed since at least the Middle Miocene (Barron and Baldauf 1989), so estimates of global diversity in the Cenozoic are extremely sensitive to the quality of data from tropical biotas. So far, next to nothing is known of the Cenozoic marine biotic history of the Tropics outside North America.

This inadequacy of biodiversity sampling is obviously just as much a problem for the present as the past. Most benthic marine organisms have extremely patchy distributions, so well-designed and intensive sampling is required to extract useful information. The Montrouzier Expedition Party to New Caledonia (Bouchet et al. 2002) sampled only 42 stations in a range of marine habitats on a 294-km² site outside the peak Indo-Pacific coral diversity hotspot. Approximately 400 person-hours of sampling revealed 2738 species of mollusks, and cumulative curves suggest that with additional sampling the count will level off at 3900 species. This is many times more than ever recorded in a similar area anywhere, with one-fifth of the taxa represented by a single specimen and one-third of the taxa recorded from a single station. More exciting is that another site 200 km distant shares just 36% of the species. Similar patterns are likely to be discovered in Cenozoic deposits of the Indo-Pacific, and those sites must be sampled with a similar philosophy. Paleobiologists will nev-

er understand the history of Cenozoic global diversity without new data from the region.

We can use new collections to take advantage of the statistical tools that ecologists have long applied to the study of extant biodiversity (Ricklefs 1990; Colwell and Coddington 1994; Hayek and Buzas 1997). Knowledge of local abundances of organisms allows us to determine sampling completeness (Koch 1987) and to better dissect patterns of biodiversity, because some aspects of biotic transformation are not registered in simple counts of taxa. For example, the regional richness of mollusks did not change significantly during a major ecological reorganization of trophic structure in the southwest Caribbean during Plio-Pleistocene time (Todd et al. 2002). These changes were confined to large shifts in the relative proportion of molluscan predators and their prey. Without abundance data, we would have missed this important transition in the history of the Caribbean biota.

New collections are essential because abundance data cannot be recovered from most museum collections. These collections are archives of the characters used to define taxa, so duplicate, fragmentary, or poorly preserved specimens might not be kept. The end result is an overemphasis on rare taxa and few useful abundance data. In general, museum collections were not compiled to address ecological questions and they are not up to the task. Ecological censuses require extremely large collections, especially if taxa cannot be identified in the field, and this is the case for the majority of Cenozoic tropical marine invertebrates. Existing collections must play a subordinate role of helping to identify potential sampling sites and as taxonomic reference collections.

We know that the global biota has responded to long-term changes in earth systems in a complex way, and that understanding these responses should be a primary target for paleobiology. Fossils are the primary data source for this project, but we have only begun to assemble collections for the task, and much more remains to be discovered. Compilations of existing data that were collected to address other problems are a step in the right direction, but we already have overwhelming evi-

dence that the existing data bank is severely biased, at least for the Cenozoic. Existing museum collections will be useful, but they are not sufficient to document large-scale patterns of biotic change or to test whether any observed patterns are the result of incomplete or biased sampling. In short, paleobiology needs more raw data before it runs out of gas.

Acknowledgments

This essay is a result of many long discussions with A. F. Budd, J. B. C. Jackson and J. A. Todd, and it was improved by the comments of S. Wing and R. Bambach. This work was supported by the National Science Foundation (EAR-9909485).

Literature Cited

- Alroy, J., et al. 2001. Effects of sampling standardization on estimates of Phanerozoic marine diversification. *Proceedings of the National Academy of Sciences USA* 98:6261–6266.
- Bambach, R. K. 1977. Species richness in marine benthic habitats through the Phanerozoic. *Paleobiology* 3:152–167.
- Barron, J. A., and J. G. Baldauf. 1989. Tertiary cooling steps and paleoproductivity as reflected by diatoms and biosiliceous sediments. Pp. 341–354 in W. H. Berger, V. S. Smetacek, and G. Wefer, eds. *Productivity of the ocean: Present and Past*, Wiley, New York.
- Bouchet, P., P. Lozouet, P. Maestrati, and V. Heros. 2002. Assessing the magnitude of molluscan species richness in tropical marine environments: exceptionally high numbers of molluscs at a New Caledonia site. *Biological Journal of the Linnean Society* 75:421–436.
- Budd, A. F., K. G. Johnson, T. A. Stemmann, and B. H. Tompkins. 1999. Pliocene to Pleistocene reef coral assemblages in the Limon Group of Costa Rica. In L. S. Collins and A. G. Coates, eds. *A paleobiotic survey of Caribbean faunas from the Neogene of the Isthmus of Panama*. *Bulletins of American Paleontology* 357:119–158.
- Colwell, R. K., and J. A. Coddington. 1994. Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions of the Royal Society of London B* 345:101–118.
- Frost, S. H. 1977. Cenozoic reef systems of the Caribbean—prospects for paleoecological synthesis. In S. H. Frost, M. P. Weiss, and J. B. Saunders, eds. *Reefs and related carbonates: ecology and sedimentology*. *American Association of Petroleum Geologists Studies in Geology* 4:93–110.
- Hayek, L.-A. C., and M. A. Buzas. 1997. Surveying natural populations. Columbia University Press, New York.
- Jackson, J. B. C., and K. G. Johnson. 2000. Life in the last few million years. In D. H. Erwin and S. L. Wing, eds. *Deep time: Paleobiology's perspective*. *Paleobiology* 26(Suppl. to No. 4): 221–235.
- . 2001. Measuring past biodiversity. *Science* 293:2401–2404.
- Jackson, J. B. C., J. A. Todd, H. Fortunato, and P. Jung. 1999. Diversity and assemblages of Neogene Caribbean Mollusca of lower Central America. In L. S. Collins and A. G. Coates, eds. *A paleobiotic survey of Caribbean faunas from the Neogene of the Isthmus of Panama*. *Bulletins of American Paleontology* 357:193–230.
- Koch, C. F. 1987. Prediction of sample size effects on the measured temporal and geographic distribution patterns of species. *Paleobiology* 13:100–107.
- Raup, D. M. 1972. Taxonomic diversity during the Phanerozoic. *Science* 177:1065–1071.
- Ricklefs, R. E. 1990. *Ecology*, 3d ed. W. H. Freeman, New York.
- Sepkoski, J. J., Jr. 1993. Ten years in the library: new data confirm paleontological patterns. *Paleobiology* 19:43–51.
- Todd, J. A., J. B. C. Jackson, K. G. Johnson, H. M. Fortunato, A. Heitz, M. Alvarez, and P. Jung. 2002. The ecology of extinction: molluscan feeding and faunal turnover in the Caribbean Neogene. *Proceedings of the Royal Society of London B* 269: 571–577.
- Valentine, J. W. 1969. Patterns of taxonomic and ecological structure of the shelf benthos during Phanerozoic time. *Palaeontology* 12:684–709.
- . 1970. How many marine invertebrate fossil species? A new approximation. *Journal of Paleontology* 44:410–415.
- Veron, J. E. N., and R. Kelley. 1988. Species stability in reef corals of Papua New Guinea and the Indo-Pacific. *Association of Australasian Palaeontologists Memoir* 69.
- Willis, K. J., and R. J. Whittaker. 2002. Species diversity—scale matters. *Science* 295:1245–1248.
- Woodring, W. P. 1957–1982. *Geology and paleontology of the Canal Zone and adjoining parts of Panama: geology and description of Tertiary mollusks*. U. S. Geological Survey Professional Papers 306A–E, 307.