Global databases will yield reliable measures of global biodiversity

John Alroy

National Center for Ecological Analysis and Synthesis, University of California, 735 State Street, Suite 300, Santa Barbara, California 93101. E-mail: alroy@nceas.ucsb.edu

For decades, paleobiologists have treated global diversity estimation as a straightforward problem (Miller 2000): count up the known higher taxa in each geological time interval, make a diversity curve, and go straight ahead to analyzing and interpreting the trends. However, global diversity curves recently have come under attack from all sides. Some researchers argue that although traditional curves are strongly affected by sampling biases (e.g., Smith 2001; Peters and Foote 2002), these biases can be corrected by assembling large, locality-level databases with detailed contextual information (Alroy et al. 2001). Others point to the large gap between true total global richness and the meager head counts the fossil record has to offer, and conclude that workers should focus exclusively on local and regional diversity (Jackson and Johnson 2001). Here I argue that although further fieldwork surely is needed, understanding global diversity in the short term remains a tractable goal—as long as we move quickly to build a discipline-wide, globally extensive paleontological database.

Relative versus total diversity: Because the fossil record always is incomplete, we cannot directly determine true *total* diversity, i.e., a series of exact, direct counts of all the species that ever existed at each point in time (e.g., Jackson and Johnson 2001). However, demonstrating an adaptive radiation or diversity crash only requires showing in relative terms that diversity was higher or lower in one time interval than the next. The same is true of other important paleontological patterns: the replacement of one taxonomic group by another (e.g., McKinney et al. 1998), the contrast between taxonomic diversity and morphological disparity (Foote 1991), diversity differences among geographic regions (Miller 1997; Jablonski 1998), and the temporal dynamics of diversification (Sepkoski 1978). All of these things can be studied by quantifying relative diversity levels, instead of waiting for centuries to inventory every species that ever was fossilized.

Sampling biases: Skeptics also may point to the enormous impact of sampling biases (Raup 1972) as an indictment against global diversity compilations. Indeed, diversity estimates are strongly affected by rock outcrop area (Raup 1976; Smith 2001; Peters and Foote 2002), taxonomic definitions (Patterson and Smith 1987; Wagner 1995), the historical accumulation of data (Sepkoski 1993; Niklas and Tiffney 1994; Alroy 2000), methods for counting diversity (Alroy et al. 2001), and the sheer amount of fossil data (Alroy 1996; Miller and Foote 1996; Alroy et al. 2001). However, the best evidence that biases are tractable is precisely the fact that sensible methods for removing sampling artifacts do indeed create large differences in the relative shapes of diversity curves. These apparently effective methods include classical rarefaction (Miller and Foote 1996), randomized subsampling of taxonomic lists (Alroy 1996; Alroy et al. 2001), extrapolation indices based on ecological theory (Nichols and Pollock 1983; Connolly and Miller 2002), and modified survivorship analysis (Foote 2002). Although many of these methods are simply too new for their properties to be well understood, the pace of theoretical discovery has increased so dramatically that optimism seems warranted.

Coupling of regional and global diversity: Some workers have argued persuasively that diversity patterns may be quite different in different biogeographic provinces (Miller 1997; Ja-

blonski 1998). If global diversity is just the sum of regional patterns and these regional patterns differ unpredictably, then one could argue that only the regional patterns matter. However, the alternative would be that regional patterns exactly reflect global trends, which if true would mean that intensive, continuous data for any one region would tell the whole story. Surely we can agree that these rival hypotheses can and should be tested by collecting data on multiple regions—i.e., global data, but global data grounded at the collection level. Thus, whatever we think of global patterns per se, we all should share the goal of compiling globally extensive data.

Estimation of regional diversity: Having everyone focus narrowly on estimating local, basin-scale species richness (e.g., Jackson and Johnson 2001) would leave methodological problems unsolved. Basin-scale data suffer from many of the same biases as global data, including changes through time in sea level (Holland 1995; Smith et al. 2001), paleolatitude (Allison and Briggs 1993), and major taphonomic signatures (Koch and Sohl 1983; Kidwell and Brenchley 1994; Schubert et al. 1997), not to mention temporal variation in total sampling intensity (e.g., Badgley and Gingerich 1988; Alroy 1996; Miller and Foote 1996). Some of these effects may be unavoidable. However, many other biases could be balanced out by methodically sampling multiple regions with distinctive megataphonomic overprints, which is exactly the kind of analysis made possible by a global, but collection-level, database.

Estimation of diversity within assemblages: One could argue that measuring diversity within a collection is tractable in a way that tallying global diversity is not, because at least the sampling pool is relatively well defined. However, at the scale of single fossil collections sampling biases become profound, because only a small fraction of the regional species pool can be represented in a tray full of specimens. Furthermore, even the most sophisticated ecological methods for extrapolating the overall species pool size are known to suffer from crippling biases (Colwell and Coddington 1994). Such handicaps are substantial even when dealing with sample sizes of thousands

that far exceed the usual specimen quotas in paleoecological studies (Koch 1978). However, the nature of sampling is such that compiling large numbers of even relatively small collections eventually must yield a usable picture of regional and even global diversity—so compiling many collections that individually are inadequate for paleoecology may prove crucial for large-scale studies.

Evenness versus richness: Changes through time in the shape (including evenness) of abundance distributions might cause sampled diversity—and even sampling-standardized diversity—to change in a way that does not reflect total diversity. Indeed, it is possible for the summed relative abundance of the few most common species in an otherwise fixed pool to change dramatically. However, this concern may be overstated. Paleoecological data sets often most closely resemble infinite, untruncated distributions such as the log series (Koch 1987), whereas naturally truncated distributions such as the log normal typically fail to appear unless sample sizes are truly heroic (Koch 1978). True total richness is undefined for distributions such as the log series, which fundamentally assume that more species always can be recovered by more sampling. Therefore, the evenness/richness distinction means nothing in such cases. Meanwhile, the log normal is governed by just two parameters that correspond to richness and a measure (''gamma'') of evenness (May 1975). In practice, gamma shows remarkably little variance across data sets, with theorists arguing not over whether this pattern is real, but rather over whether it is a mathematical artifact (May 1975) or biologically meaningful (Sugihara 1980). By elimination, therefore, the major source of variation in such abundance distributions is likely to be variation in total richness. All of this suggests that it would be more helpful to redirect theoretical attention to the interplay between alpha (= ''evenness''), beta, and gamma diversity (Sepkoski 1988).

Higher and lower taxa: Yet another objection to the quest for new species-level data is the assumption that counts of higher taxa (e.g., orders: Sepkoski 1978) are more robust to sampling artifacts; ergo, they may be better prox-

ies for species-level data than the species-level data themselves (Benton et al. 2000). However, even if we put aside concerns about sampling, biologically different patterns at different taxonomic levels are expected on theoretical grounds (Sepkoski and Kendrick 1993) and do appear in traditional synoptic compilations (Sepkoski 1997). Furthermore, if we really want to know about the relationship between higher- and lower-level patterns, we need to collect significant amounts of lower-level data—and doing so will make the argument moot.

Importance of diversity curves: Many paleontologists simply may not care too much about global diversity. However, our colleagues in such varied disciplines as astronomy, oceanography, ecology, and genetics certainly do. Key papers on diversity dynamics such as Raup and Sepkoski (1982) rank among the most highly cited contributions ever published by paleontologists. Scientists in other fields such as molecular systematics always will continue to make claims about (say) the Cambrian explosion and the Cretaceous/Tertiary boundary mass extinction. Therefore, it instead falls on us to do the best possible job of providing accurate information to the public in the form of an evolving, communally maintained database.

Benefits of databases: Many workers will always remain skeptical about the future of diversity studies. Note, however, that almost all of the preceding arguments hinge on the prospects for developing a discipline-wide database of the fossil record (Miller 2000; Alroy et al. 2001). Regardless of the diversity issue, the Paleobiology Database initiative should yield considerable benefits for almost every practicing paleontologist. First, collection-level databases will become a key tool for assessing previously published literature—regardless of whether one wants to know about just one species or fossil locality, or else an entire taxonomic group, geological formation, temporal interval, or geographic region. Second, compilations of collection-level data will provide strong justification for additional field- and specimen-based research by showing exactly where the published literature is deficient. Finally, a community-wide database will serve as a permanent repository for the hard-won data recovered by generations of paleontologists. Quantifying global diversity is a tractable, multifaceted, and scientifically important goal, and everyone stands to gain by fostering this endeavor.

Acknowledgments

Literature Cited

Allison, P. A., and D. E. G. Briggs. 1993. Paleolatitudinal sampling bias, Phanerozoic species diversity, and the end-Permian extinction. Geology 21:65–68.

Alroy, J. 1996. Constant extinction, constrained diversification, and uncoordinated stasis in North American mammals. Palaeogeography, Palaeoclimatology, Palaeoecology 127:285–311.

———. 2000. Successive approximations of diversity curves: ten more years in the library. Geology 28:1023–1026.

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.

Badgley, C., and P. D. Gingerich. 1988. Sampling and faunal turnover in early Eocene mammals. Palaeogeography, Palaeoclimatology, Palaeoecology 63:141–157.

Benton, M. J., M. A. Wills, and R. Hitchin. 2000. Quality of the fossil record through time. Nature 403:534–537.

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.

Connolly, S. R., and A. I. Miller. 2002. Joint estimation of sampling and turnover rates from fossil databases: capture-mark-recapture methods revisited. Paleobiology 27:751–767.

Foote, M. 1991. Morphological patterns of diversification: examples from trilobites. Palaeontology 34:461–485.

———. 2002. Inferring temporal patterns of preservation, origination, and extinction from taxonomic survivorship analysis. Paleobiology 27:602–630.

Holland, S. M. 1995. The stratigraphic distribution of fossils. Paleobiology 21:92–109.

Jablonski, D. 1998. Geographic variation in the molluscan recovery from the end-Cretaceous extinction. Science 279:1327–1330.

Jackson, J. B. C., and K. G. Johnson. 2001. Paleoecology: measuring past biodiversity. Science 293:2401.

Kidwell, S. M., and P. J. Brenchley. 1994. Patterns in bioclastic accumulation through the Phanerozoic: changes in input or in destruction. Geology 22:1139–1143.

Koch, C. F. 1978. Bias in the published fossil record. Paleobiology 4:367–377.

———. 1987. Prediction of sample size effects on the measured temporal and geographic distribution patterns of species. Paleobiology 13:100–107.

Koch, C. F., and N. F. Sohl. 1983. Preservational effects in paleo-ecological studies: Cretaceous mollusc examples. Paleobiology 9:26–34.

May, R. M. 1975. Patterns of species abundance and diversity. Pp. 81–120 *in* M. L. Cody and J. M. Diamond, eds. Ecology and evolution of communities. Harvard University Press, Cambridge.

McKinney, F. K., S. Lidgard, and J. J. Sepkoski Jr. 1998. Decoupled temporal patterns of evolution and ecology in two post-Paleozoic clades. Science 281:807–809.

Miller, A. I. 1997. Dissecting global diversity patterns: examples from the Ordovician Radiation. Annual Review of Ecology and Systematics 28:85–104.

———. 2000. Conversations about Phanerozoic global diversity. *In* D. H. Erwin and S. L. Wing, eds. Deep time: *Paleobiology*'s perspective. Paleobiology 26(Suppl. to No. 4):53–73.

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

Nichols, J. D., and K. H. Pollock. 1983. Estimating diversity, extinction rates, and speciation rates from fossil data using capture-recapture models. Paleobiology 9:150–163.

Niklas, K. J., and B. H. Tiffney. 1994. The quantification of plant biodiversity through time. Philosophical Transactions of the Royal Society of London B 345:35–44.

Patterson, C., and A. B. Smith. 1987. Is the periodicity of extinctions a taxonomic artifact? Nature 330:248–251.

Peters, S. E., and M. Foote. 2002 (for 2001). Biodiversity in the Phanerozoic: a reinterpretation. Paleobiology 27:584–601.

Raup, D. M. 1972. Taxonomic diversity during the Phanerozoic. Science 177:1065–1071.

———. 1976. Species diversity in the Phanerozoic: an interpretation. Paleobiology 2:289–297.

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

Schubert, J. K., D. L. Kidder, and D. H. Erwin. 1997. Silica-replaced fossils through the Phanerozoic. Geology 25:1031–1034.

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

———. 1988. Alpha, beta, or gamma: where does all the diversity go? Paleobiology 14:221–234.

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

———. 1997. Biodiversity: past, present, future. Journal of Paleontology 71:533–539.

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

Smith, A. B. 2001. Large-scale heterogeneity of the fossil record: implications for Phanerozoic biodiversity studies. Philosophical Transactions of the Royal Society of London B 356:351–367.

Smith, A. B., A. S. Gale, and N. E. A. Monks. 2001. Sea-level change and rock record bias in the Cretaceous: a problem for extinction and diversification studies. Paleobiology 27:241–253.

Sugihara, G. 1980. Minimal community structure: an explanation of species abundance patterns. American Naturalist 116:770–787.

Wagner, P. J. 1995. Diversity patterns among early gastropods: contrasting taxonomic and phylogenetic descriptions. Paleobiology 21:410–439.