

Getting the measure of diversity

Andrew B. Smith

Department of Palaeontology, The Natural History Museum, Cromwell Road, London SW7 5BD, United Kingdom. E-mail: a.smith@nhm.ac.uk

Accepted: 10 August 2002

Paleontological data have for long been paramount in providing a long-term perspective on global biodiversity. But all is not as simple and secure as it once seemed. Apparently rapid diversification events recorded in the fossil record have been challenged by new molecular data (Bromham et al. 1999; Wray 2001; reviewed in Smith and Peterson 2002), certain mass extinctions are not as well founded as was previously supposed (Smith et al. 2001; Peters and Foote 2002b), and even such a deeply cherished belief as the long-term trend of increasing diversity through the Phanerozoic is once again under question (Alroy et al. 2001; Peters and Foote 2002a). Why is the fossil record not currently providing us with reliable, clear-cut data, and what can be done to correct the situation?

The Root of the Problem

Unlike the biological fraternity, paleontologists have concentrated not on estimating absolute numbers of species that may have existed, but on using time-series data to say something relative about global diversity dynamics. Taxic counts at generic or family level have been used as a proxy for species diversity, and rises or falls that occur between time intervals then used as evidence for changing global diversity. This, however, throws up a major complication: Time-series analysis requires that the biota from each interval be uniformly sampled, or, if not uniformly sampled, that biases in sampling at least be understood so that any differences can be factored out. But this complication until recently has been ignored—at best preservation and sampling biases have simply been assumed to be random over time (Sepkoski and Raup 1986).

Two aspects of sampling need to be considered: how well we have sampled the available

rock record, and how representational that rock record is of what once existed. There is plenty of evidence that, at least for well-studied parts of the world, our knowledge of the fossil biota preserved in rocks at outcrop is really rather good (e.g., Donovan and Paul 1998; Benton et al. 2000; Adrain and Westrop 2000). It is therefore likely that we have sampled the available rock record for these groups rather efficiently, at least in North America and Europe. However, although taxonomic data may be collected from all over the world, there is an inevitable historical bias toward Northern Hemisphere sites. For example, Bleiweiss's (1998) "global database" of fossil bird localities contains 70 localities from Europe, 20 from North America, and 1 each from Ukraine, Kazakhstan, and Australia. Tectonic events in the North Atlantic-Western Tethys cannot fail to exert an undue influence on the nature of the rock record.

A more serious problem is that the rock record we have to collect from is already severely biased in complex ways before we even start sampling our first fossils. Worse still, these biases change over time.

Rock outcrop surface area is not uniform. Major transgression-regression (MTR) cycles driven by global tectonics and climate directly control the amount of rock that ends up being preserved. This is important because a species/area relationship suggests that the more rock we have to collect from, the more fossils we collect and the greater the apparent diversity. Smith (2001) and Peters and Foote (2002a) both found a striking match between diversity change and rock outcrop area using different approaches.

The record of environments is not uniform. Even where the actual surface area of rock available for collecting remains more or less constant

over time, MTR cycles can also end up changing the proportion of different environmental settings represented in the rock record. There may be an oscillation between more and less terrestrially influenced deposits affecting preservation potential, as sea level rises and falls. In some cases key environments may be completely missing for extended time intervals (Smith et al. 2001).

Preservational potential is not uniform. There may also be significant long-term changes in preservational potential driven by ocean chemistry, especially related to aragonitic faunas (e.g., Twitchett 2000; Cherns and Wright 2000).

How Do We Progress?

Clearly for time-series analysis we need to understand and account for potential biases affecting the fossil record. That requires much more information than simply first and last occurrence datums of taxa, and such an undertaking is now being actively pursued by the Paleobiology Database project (Alroy et al. 2001). Better documentation of sequence stratigraphic controls on depositional and preservational environment is also essential if we are to improve our understanding of how the rock and fossil records interrelate (e.g., Olszewski and Patzkowsky 2001).

More collecting will surely alleviate geographic bias but cannot overcome the more serious biases listed above. If rocks of the appropriate environment are not preserved and accessible at outcrop no amount of collecting will help fill the gaps. The use of rarefaction is important for removing the effect of sample-size and rock-record bias but is not without problems. Alroy et al.'s (2001) approach assumes uniformity of environmental sampling over time, yet we know that this is not the case. The Turonian in Europe, for example, will always appear to have a low-diversity fauna compared with the Cenomanian, simply because the more-onshore marine faunas appear to have been almost entirely lost from the rock record (Smith et al. 2001).

Even were all these improvements to be made, a major problem persists. If diversity and the rock record covary, as they seem to do (Smith 2001; Peters and Foote 2002b), how can

we tell whether we are looking at a biological signal, or a sampling artifact, or a complex mix of the two? Changing sea level directly alters the surface area of critically important marine zones on the continental shelf (Schaff 1996), which could be altering marine biodiversity. But it also drives sedimentary depositional cycles and the relative ratio of on-shore-offshore environments that end up being preserved in the rock record, thereby controlling sampling. Do we then have to accept that only those biodiversity trends that cannot be explained by simple rock record biases are worth pursuing? Probably not.

Cladograms are constructed independent of stratigraphic data and so, when calibrated against the observed fossil record, can provide one way to identify time intervals of relatively good or relatively poor preservation potential. Drops in diversity that coincide with increased gapiness of the fossil record and peaks of pseudoextinction¹ indicate that sampling problems exist, whereas low Lazarus taxon counts² and peaks of monophyletic clade extinction suggest a genuine biological signal. Thus cladograms can be used to differentiate between sampling artifact and biological pattern.

Although cladograms are independent of rock record bias, other phylogenetic methods that incorporate stratigraphic data into the tree-building process are not, of course, and problems concerning origination patterns may still remain. If a sampling problem removed the early history of a group (say the Late Cretaceous record of neornithine birds, for example), then even a calibrated cladogram will appear to show early branches all crowded together in an apparent "adaptive radiation" when preservational conditions markedly improved. Because early branches remain unsampled, a cladogram calibrated by using the rock record would miss these basal ghost ranges,³

¹ Pseudoextinction: the termination of a non-monophyletic group (see Smith 1994).

² Lazarus taxon: a taxon that disappears completely from the rock record for a significant period of time (see Smith 1994).

³ Ghost range: a period of time during which a taxon is predicted to have existed purely from considering its phylogenetic relationships to other taxa and their stratigraphic ranges (see Norell 1992).

and the completeness of the record would be overestimated.

It is probably time to accept that raw taxonomic counts provide only a first, crude estimate of biodiversity dynamics and occasionally may be downright misleading. Now that we are starting to understand the nature of geological biases and how global tectonics, through depositional architecture, affects the fossil record, more accurate estimates of global diversity dynamics are within our grasp.

Literature Cited

- Adrain, J. M., and S. R. Westrop. 2000. An empirical assessment of taxic paleobiology. *Science* 289:110–112.
- 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.
- Benton, M. J., M. A. Wills, and R. Hitchin. 2000. The quality of the fossil record through time. *Nature* 403:534–537.
- Bleiweiss, R. 1998. Fossil gap analysis supports early Tertiary origin of trophically diverse avian orders. *Geology* 26:323–326.
- Bromham, L., M. J. Phillips, and D. Penny. 1999. Growing up with dinosaurs: molecular dates and the mammalian radiation. *Trends in Ecology and Evolution* 14:113–118.
- Cherns, L., and V. P. Wright. 2000. Missing molluscs as evidence of large-scale, early aragonite dissolution in a Silurian sea. *Geology* 28:791–794.
- Donovan, S. K., and C. R. C. Paul. 1998. The adequacy of the fossil record. Wiley, Chichester, England.
- Norell, M. A. 1992. Taxic origin and temporal diversity: the effect of phylogeny. Pp. 88–118 in M. J. Novacek and Q. D. Wheeler, eds. *Extinction and phylogeny*. Columbia University Press, New York.
- Olszewski, T. D., and M. E. Patzkowsky. 2001. Evaluating taxonomic turnover: Pennsylvanian–Permian brachiopods and bivalves of the North American Midcontinent. *Paleobiology* 27: 646–669.
- Peters, S. E., and M. Foote. 2002a (for 2001). Biodiversity in the Phanerozoic: a reinterpretation. *Paleobiology* 27:583–601.
- . 2002b. Determinants of extinction in the fossil record. *Nature* 416:420–424.
- Schaff, A. 1996. Sea-level changes, continental shelf morphology, and global paleoecological constraints in the shallow benthic realm: a theoretical approach. *Palaeogeography, Palaeoclimatology, Palaeoecology* 121:259–271.
- Sepkoski, J. J., Jr., and D. M. Raup. 1986. Periodicity in marine extinction events. Pp. 3–36 in D. K. Elliott, ed. *Dynamics of extinction*. Wiley, New York.
- Smith, A. B. 1994. Systematics and the fossil record: documenting evolutionary patterns. Blackwell Scientific, London.
- . 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., and K. J. Peterson. 2002. Dating the time of origin of major clades: molecular clocks and the fossil record. *Annual Review of Earth and Planetary Sciences* 30:65–88.
- 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 biodiversity studies. *Paleobiology* 27:241–253.
- Twitchett, R. J. 2000. Discussion of Lazarus taxa and fossil abundance at times of biotic crisis. *Journal of the Geological Society, London* 157:511–512.
- Wray, G. A. 2001. Dating branches on the tree of life using DNA. *Genome Biology* 3:1–7.