

# Setting the record straight

By analysing masses of data from fossils throughout the world, a group of palaeontologists hopes to address the big questions about the history of life on Earth. Quirin Schiermeier logs on to the Paleobiology Database.

**W**e've never had it so good — or at least that has been the prevailing view among palaeobiologists who have tried to track the history of our planet's biodiversity. On the long road from the first stirrings of multicellular life to today's shimmering diversity, untold numbers of species have fallen by the wayside. From time to time, legions of creatures have perished together in mysterious mass extinctions. But if you examine the fossil record, the evolution of new species seems generally to have had the upper hand over extinction. Like stock indices in a bull market, plots showing the diversity of life over geological time reveal a rising trend, despite occasional setbacks.

But how can we be sure that this isn't a sampling artefact? Even high-school biology students are taught that the fossil record is far from complete. Given that younger rocks are more likely to be exposed at the surface, it is possible that the apparent rise in biodiversity merely represents the greater scrutiny that has been applied to these strata. Palaeontologists have even coined a term for this source of bias: 'the pull of the recent'. Add in the confusion caused by the varied names used to describe the same organisms, and some researchers argue that attempting to assess the history of Earth's biodiversity is a fool's quest.

John Alroy, a palaeontologist at the University of California, Santa Barbara (UCSB), begs to differ. He is one of the founders of the Paleobiology Database, a project set up in 2000 with financial support from the US National Science Foundation. This freely accessible database, hosted by UCSB's National Center for Ecological Analysis and Synthesis, already holds information on more than 30,000 different fossil collections, and is still growing. Through sheer weight of numbers, and by applying various statistical tricks to account for sampling biases, Alroy and his colleagues hope to determine whether the Earth's bio-

diversity really has been on the rise — and to answer some other tricky questions.

The data are divided into individual collections, retrieved from specific locations and strata by particular palaeontologists. In addition to descriptions of specimens, the database includes information on the composition and age of the sediments in which they were found, and the fossils' state of preservation. "It is the multitude of easily retrievable information that makes it so useful," says Wolfgang Kiessling, a palaeontologist at the Museum of Natural History in Berlin, who is one of the 70 or so scientists authorized to enter information into the database. "It allows us to interpret the known fossil record in a more unbiased way, and we can now ask a whole host of new questions about how natural systems operate."

But some sceptics suspect that no amount of statistical sophistication will eliminate the uncertainty inherent in palaeontology. It may always be impossible to determine the degree to which the fossil record is 'known', they argue. And some fear that the Paleobiology Database will seduce unwary researchers into drawing erroneous conclusions. "No doubt palaeontology will benefit from more informed fossil data," says Andrew Smith, an invertebrate palaeontologist at London's Natural History Museum. "But you can easily be misled if you assume that all data are objective."

Multicellular life left little impression on the fossil record for around half a billion years. But at the onset of the Cambrian period, some 550 million years ago, oxygen and calcium had become sufficiently abundant in the oceans for the development of organisms with hard components. The result was the 'Cambrian explosion' of marine biodiversity. Our understanding of diversity trends since then owes a heavy debt to the work of John 'Jack' Sepkoski, a palaeontologist at the University of Chicago. In the 1970s, Sepkoski began to scour the palaeontological literature for information on the first and last appear-



Rock compilation: snapshots of fossil diversity are being united to reveal the bigger picture.



John Alroy hopes the database will answer crucial fossil questions.

ances of marine organisms, extrapolating the ups and downs of life in the oceans. Because individual species appear only fleetingly in the fossil record, and are often misidentified, he focused on genera — the trilobites *Paradoxides bohemicus* and *Paradoxides gracilis*, for instance, are species within the same genus, whereas *Asaphus cornutus* and *Crozonaspis struvei* belong to different genera.

## Expanding knowledge

Sepkoski's work<sup>1</sup> indicated that diversity continued to expand from the Cambrian explosion until the end of the Ordovician period, some 440 million years ago. From then on, it remained on a more or less stable plateau until the Permian-Triassic mass extinction — the most severe experienced by our planet. But Sepkoski found an upward trend in marine biodiversity after the beginning of the Triassic period, 250 million years ago, with just one significant interruption: the extinction event that put paid to the dinosaurs at the end of the Cretaceous period, some 65 million years ago (see figure, opposite). "Sepkoski has given us a clue about the dramatic things that happened in the history of life," says Alroy. "But we need more comprehensive data, and better analytical tools, to quantify his findings."

This is what the Paleobiology Database aims to provide. Because of the information included about individual collections, it is possible to correct for sampling biases in ways that

and his colleagues argued, the increase in diversity over time does seem to be real.

Alroy and his colleagues believe that the database is the key to resolving this and other controversies surrounding the history of life on Earth, such as whether the great mass extinctions really were as dramatic as has been assumed. It should also give palaeo-ecologists a better idea of whether biodiversity is controlled by environmental parameters such as climate, volcanic activity and ocean chemistry — or whether, as a theory proposed two years ago<sup>4</sup> suggests, it varies randomly.

### Share and share alike

Getting the most out of the database may require a cultural change on the part of some palaeontologists, however. Further expansion of its scope will require researchers to make their collections available for analysis. The situation in New Zealand, where palaeontologists began three decades ago to compile and publish all fossil data in an openly searchable way, under an agreement between the Geological Society of New Zealand and the New Zealand Geological Survey, provides an ideal model. In a paper published just a few weeks ago, these data were used to study bias in measurements of mollusc diversity caused by variation in the total area of exposed rocks of different ages<sup>5</sup>. But in Europe, says Kiessling, some palaeontologists still jealously guard their own collections to maintain an advantage over their rivals.

The value of the Paleobiology Database will depend on the quality, as well as the quantity, of its information. Some experts fear that quality-control issues could cause misleading results, particularly in the hands of scientists who are not experts on the organisms that they are trying to analyse. Smith, for instance, is concerned about the potential for confusion due to problems with taxonomic nomenclature. “Names may disappear, but their last occurrence in the record does not necessarily mean extinction of a species, family or genus,” he says. Despite such shortcomings, however, Smith intends to use the database and contribute to it. “But I would only work with taxa that I know,” he adds.

Alroy and his colleagues are trying to address the problem that Smith has highlighted. At a meeting later this year, they plan to set up task force to resolve inconsistencies in the database. Once this group’s work is done, enthusiasts claim, the database will be a powerful tool. “It will add a long-term perspective to many open questions,” says Kiessling. ■

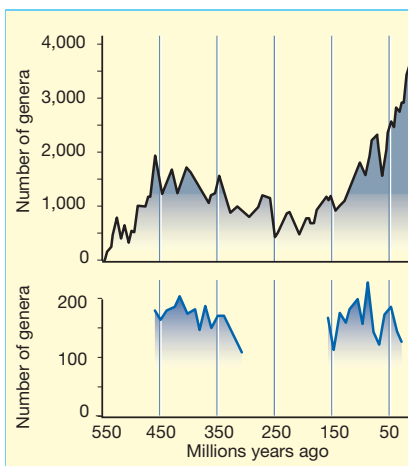
**Quirin Schiermeier is Nature’s German correspondent.**

1. Sepkoski, J. J. Jr *Paleobiology* 7, 36–53 (1981).
2. Alroy, J. *et al. Proc. Natl Acad. Sci. USA* 98, 6261–6266 (2001).
3. Jablonski, D., Roy, K., Valentine, J. W., Price, R. M. & Anderson, P. S. *Science* 300, 1133–1135 (2003).
4. Hubbell, S. P. *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton Univ. Press, Princeton, New Jersey, 2001).
5. Crampton, J. S. *et al. Science* 301, 358–360 (2003).

▶ [www.paleodb.org](http://www.paleodb.org)



SOURCE: REF. 2



**On the level: the traditional view that biodiversity has increased over time (top) is challenged by an analysis that corrects for sampling bias.**

Sepkoski, with his simple analysis of the first and last appearances of genera in the fossil record, was unable to do. For instance, within each interval being studied you can examine a fixed number of collections, in an attempt to account for variation in sampling intensity from rocks of different ages. Other corrections can be applied to account for bias in geographical coverage, and so on. Sepkoski himself realized the need to do this, and contributed to the Paleobiology Database until his death from heart failure, aged just 50, in 1999.

Initial analyses have already given some intriguing hints of discrepancies from Sepkoski’s earlier conclusions. In the first major paper<sup>2</sup> to make use of the Paleobiology Database, Alroy and 24 colleagues — including the late Sepkoski — sampled the database’s marine component, which at the time contained 8,591 collections, mainly from North America and Europe. They applied four different statistical methods to correct for variation in sampling intensity in rocks of different ages. Each gave roughly the same result, suggesting that marine biodiversity has not risen over the past 150 million years, and is at a similar level to that during the period between 450 million and 300 million years ago (see figure).

If this finding is correct, it means that Sepkoski’s conclusion about rising diversity since the Triassic is a sampling artefact. “This is a very important and surprising result,” says Kiessling. “It is the first time evidence has been found that there may be an upper threshold to biodiversity — a maximum holding capacity of the environment.” The threshold theory is controversial, however, and the picture may yet change again, as researchers consider data on other taxonomic groups or from different regions.

Indeed, a team led by palaeontologist David Jablonski of the University of Chicago has analysed the database’s entries for bivalve molluscs, concluding that the pull of the recent has been overestimated in previous studies<sup>3</sup>. For this group of animals, at least, Jablonski