8

The History of Life: Large Databases in Palaeontology

Michael J. Benton

For almost every topic discussed in the following pages the data are insufficient. The student who attempts interpretations under these circumstances does so in the face of certainty that some of his conclusions will be rejected. It is, however, pusillanimous to avoid making our best efforts today because they may appear inadequate tomorrow. Indeed, there would be no tomorrow for science if this common attitude were universal. Facts are useless to science unless they are understood.   (Simpson, 1944, p. xviii)

Fossils have one unique attribute – a long distribution in time. Evolutionary biologists, and the general public, expect palaeontologists to tell them how life originated, how life diversified from presumably one species to the tens of millions today, how major new adaptations or life modes became established, the nature and cause of clade radiations, how many mass extinctions there were, how many smaller extinctions occurred and the effects of those extinction events. Palaeontological data on rates of extinctions and recoveries will be critical in determining just how human activities affect the present-day biota.

Only palaeontologists can answer these big questions, and big questions often require big databases and appropriate statistical analyses. Yet, until the 1980s, only a handful of palaeontologists worked with data sets on large parts of the history of life and they sought to document the history of animal life in the sea in particular. Their discovery of the pattern of diversification of life, and of patterns of extinction events, brought great acclaim.

The purpose of this chapter is to consider the strengths and pitfalls of statistical testing of data bases on the large-scale evolution of life. Some key

questions are considered, the timing of major extinction events (periodic or non-periodic), the nature of radiations (adaptive/competitive or unrestrained expansion) and the diversification of life (additive, logistic or exponential).

TIME

The time dimension in palaeontology is useful at several levels of analysis. In basin-wide studies of faunal evolution (Chapter 7), the evolution of individual species may be tracked at time scales with a resolution to hundreds, or even tens, of thousands of years. On a global scale, time resolution to the stage or substage (durations, 1–10 My) is readily achievable and for certain well-documented groups resolution can improve to the precision of zones (durations, 0.25–2 My). These levels of precision can be improved no doubt, but they are perfectly adequate for many of the larger questions in evolution.

SIMPLICITY

The key scientific questions are simple, and they can frequently be framed in such a way that they are amenable to simple statistical testing. It is pointless to try to resolve a major question in evolution head-on: Why did the dinosaurs die out? Why did skeletons appear at the beginning of the Cambrian? Why is life so diverse? Good experimental design requires that a question is framed in such a way that two or more hypotheses may be tested by a decisive statistical treatment of the data.

Why did the dinosaurs die out? This question may be reframed in terms of particular published hypotheses. First try: did dinosaurs die out because they were too stupid (Fremlin, 1979)? The analyst would have to find an index of stupidity and then compare the dinosaurs and other victims of the Cretaceous–Tertiary boundary (K/T) event with the survivors. Perhaps a stupidity index could be based on a measure of relative brain size? This can be measured from the fossils, although with some inherent error. The argument becomes difficult, however, because relative brain size alone is not a strict measure of intelligence since the brain contains parts that deal with sensory systems. Birds have relatively large brains, but much of that is to control their excellent eyesight. Perhaps the measure of stupidity based on relative brain size would be so problematic that the hypothesis must be set aside as untestable.

Second try: did the dinosaurs die out because they were so big? Perhaps this is amenable to testing. The proposal could be legitimately reframed to read: During the K/T event there was selection based on body size among

terrestrial vertebrates, and the big ones disappeared. The analyst would then carry out an exhaustive survey of all vertebrate species immediately before and after the event and assess their body sizes, ideally as estimated body weights (which, unfortunately, brings in a level of uncertainty). The hypothesis can then be further tightened up to read, 'dinosaurs and other terrestrial vertebrates over the weight of 25 kg were more likely to die out during the K/T event than those weighing less than 25 kg'. The null hypothesis is that both size classes suffered equally during the K/T event and a chi-squared test could indicate whether that null hypothesis is rejected or not. This test has yet to be done (Archibald, 1996).

A related example shows the simplicity of large-scale statistical analyses in evolutionary palaeontology. Jablonski and Raup (1995) tackled one of the key questions concerning mass extinctions: is there any evidence for selectivity of victims and survivors? They used a large data set of 3472 occurrences of 347 genera of marine bivalves from the Maastrichtian (latest Cretaceous) worldwide, and coded each genus for a variety of ecological attributes (Figure 8.1). They applied simple statistical tests to comparisons of the characteristics of victims and survivors. The null expectation was that there was no difference in the characters of survivors and victims, and the tests were designed to find significant deviations from that null expectation.

Jablonski and Raup (1995) found that bivalves that survived the K/T event were not very different from the victims. Characters such as body size, diet,
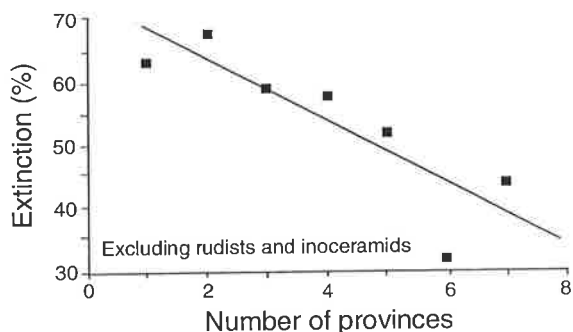
Figure 8.1 *Statistical testing in palaeontology: does the geographic distribution of a genus affect its chances of survival during a mass extinction event? Here, a plot of the breadth of geographic distribution of genera of latest Cretaceous bivalves (shown by provinces, identifiable regions with different faunas) shows that the more widely distributed a genus is, the better were its chances of survival through the K/T event. Genera restricted to one or two provinces showed extinction rates of 60–70%, while those present in six or seven provinces had extinction rates of 30–45%. Good fit of the points to a straight line suggests there is a strong relationship. Based on data in Jablonski and Raup (1995)*

bathymetric position, relative breadth of bathymetric range and life position (burrowers versus non-burrowers) conferred no advantages. The only eco-logical attribute that showed a statistically significant ($p < 0.05$) divergence from the null expectation was the geographic distribution of a genus (Figure 8.1): widely-distributed genera survived better through the rigours of the K/T event than endemic genera. This is a simple and elegant demonstration, but it is important to take note of criticisms of such studies.

THE CRITICS

The Questions

There are problems with large databases and statistical testing, but the existence of a few problems should not induce palaeontologists to ignore their greatest asset. Palaeontologists should speak to other evolutionary biologists, and not simply shrug their shoulders and say, 'Well, we haven't enough data yet' or 'That question is much too complex to answer.' Of course, no one can give the final explanation for the Cambrian explosion of skeletonized life in the sea, the huge diversity of insects, the demise of the dinosaurs, or the nature of the diversification of life through time. However, these issues have been tackled recently by the use of large databases, and at least some patterns and explanations can be ruled out, and others given some support.

Many readers will recall palaeontological meetings at which fun was made of the work of statistical palaeontologists. Some of the criticisms were based on valid concerns, but others were founded on a pessimistic view of the potential of palaeontology. The key concerns can be answered now, and younger palaeontologists are happy to explore a variety of approaches in their work.

The critics fall into several categories (I exaggerate horribly here, but perhaps there is a grain of truth in these caricatures):

1. *The proofreader.* This critic discovers errors in the database, perhaps some incorrect age assignments, or some taxa omitted. The proofreader believes that he/she has found serious flaws that are sufficient to damn the whole project.
2. *The trades unionist.* This critic is keen to preserve his/her patch, to insist on precise job definitions and the maintenance of restrictive practices. He/she has developed profound expertise in the widgetoids of the Lower Perm-ian, and he/she cannot bear to think of a statistical analyst using data second-hand: 'How can this person analyse the patterns of diversification of this group when I have spent a lifetime studying it?'

3. *The luddite.* In every field of endeavour, everyone feels threatened by new ideas and new technologies. The computer-hater cannot bear the idea of rotating vast data sets and reducing them to simple graphs. 'How is it done? My old professor never did this kind of thing, and it is a danger to the ethics of our profession.'

4. *The utopian.* This critic suggests that some day palaeontologists *will* have adequate knowledge to answer all kinds of large evolutionary questions, but that day has not yet come. The common cry of this bird is, 'We don't have enough data.'

The Answers: Errors in Databases do not Generally Affect Results

The criticism of the *proofreader* could be fatal. If it turns out that databases are riddled with errors and that they change in profound ways with each new fossil-collecting trip, with each revision of the taxonomy of a group, and with each re-evaluation of stratigraphy, then the whole statistical enterprise would be shown to be deeply flawed.

However, the proofreader's criticism has been answered before (Raup, 1991). A statistical approach automatically assumes that databases are incomplete and imperfect. After all, the standard statistical tests seek a 95% pass rate (i.e. $p < 0.05$), and most other scientists regard this as adequate for biological systems. The imperfections may be real, where there are gaps in the known record of a taxon caused by non-preservation or non-collection, or the imperfections may be artefacts of the literature, where errors have been made in identification or dating either by the primary researcher or by the data compiler. Hence, the discovery of errors here and there in a database rarely invalidates the conclusions drawn from statistical analyses of that database. The proofreader has to demonstrate a case where the correction of errors in a palaeontological database has altered the conclusions based on statistical testing.

It might even be suggested that poor data could generate an artificial signal, and if that were the case, then the whole analytical enterprise would be sunk. Hoffman (1985b) and Patterson and Smith (1987), for example, suggested that the apparent pattern of periodicity in extinction rates over the past 250 My identified by Raup and Sepkoski (1984) was generated by erroneous data. This case is actually more subtle (see below). Raup and Sepkoski (1986) and Sepkoski (1989) have made it clear that it is very unlikely that an organized pattern could be generated by errors in the data ('noise'). Statistically speaking, this is as likely an event as the production of the complete works of Shakespeare (or even one act, or one scene) by the famous team of monkeys armed with typewriters.

It is sometimes possible to predict an artificial pattern that might be

generated by poor data, and to test for it. Jablonski (1988), for example, confirmed that there is a real relationship between geographic area occupied and the longevity of Cretaceous molluscan species. Critics suggested that this was simply a sampling artefact: the geographically widespread species were just found more often and seemed to have longer durations than more localized species. Jablonski (1988) partitioned the Cretaceous mollusc species into clades with different preservation potentials. The prediction from sampling bias was that the best-preserved clades should show the weakest relation between duration and range, and clades with low preservation potential should show the strongest relation. This was not found, and the relationship between duration and range was confirmed as real.

Experience now suggests that large databases in palaeontology are remarkably robust for many purposes. Maxwell and Benton (1990) showed that lists of families of tetrapods compiled in 1900, 1933, 1945, 1966 and 1987 had changed enormously, with a doubling of numbers. In addition, known stratigraphic ranges had been much altered, with range extensions as a result of the collection of new fossils, and range contractions as a result of re-study of the critical early and late occurrences of specimens. Nevertheless, the overall pattern of familial diversification of tetrapods remained the same, based on data from 1900 to 1987, and the timing and relative magnitudes of extinction events were the same (Figure 8.2a). Sepkoski (1993) also found no change in patterns of diversification and extinction events for marine animals, based on a comparison of changes in his database from 1982 to 1992 (Figure 8.2b). The only serious difference was that extinction events had become sharper, mainly the result of new finds closer to the extinction horizons.

In a more subtle study, Benton and Storrs (1994, 1996) found that one of the main changes in databases has been to fill up predicted gaps in the fossil record, gaps of two kinds: (1) those within stratigraphic ranges, and (2) those that occur *before* the first-known fossil representative. By definition, sister groups diverged from each other at a single time, and yet two such close relatives rarely have oldest fossil representatives of the same age. The difference between the ages of oldest fossil representatives of sister groups is the minimum implied gap (MIG) or ghost range (Norell and Novacek, 1992; Benton, 1994, 1995b). The MIG represents a known gap in the fossil record, not an imagined gap, providing the sister-group relationships have been determined correctly, and providing the oldest known fossil representatives of each sister group have been correctly assigned taxonomically and correctly dated. Benton and Storrs (1994, 1996) found that the fossil record of tetrapods between 1967 and 1993 had improved by 5%; in other words, new fossils have been filling up the ghost ranges (Figure 8.3).

This study has added meat to the argument by Raup *et al.* (1973) that many patterns seen as interpretable in the fossil record could have resulted from
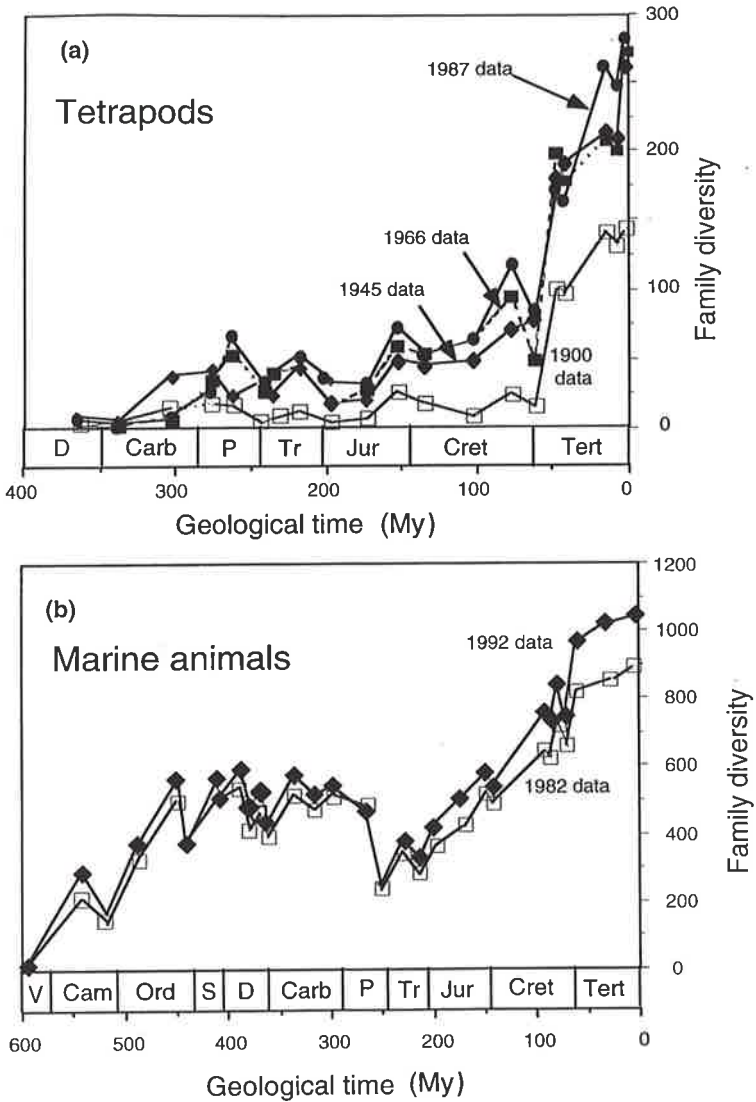
Figure 8.2 *Comparisons of the patterns of diversification of non-marine tetrapods (a) and marine animals (b) based on data sets of different vintage. The diversity patterns are broadly similar, showing gradual increases in diversity through time, interrupted by various declines, corresponding to mass extinction events (Late Cambrian, Late Ordovician, Late Devonian, end-Permian, Late Triassic, end-Cretaceous and Late Eocene). The main change, in both cases, is that overall diversity has increased. Abbreviations: V, Vendian; Cam, Cambrian; Ord, Ordovician; S, Silurian; D, Devonian; Carb, Carboniferous; P, Permian; Tr, Triassic; Jur, Jurassic; Cret, Cretaceous; Tert, Tertiary. Based on data in Maxwell and Benton (1990) and Sepkoski (1993)*
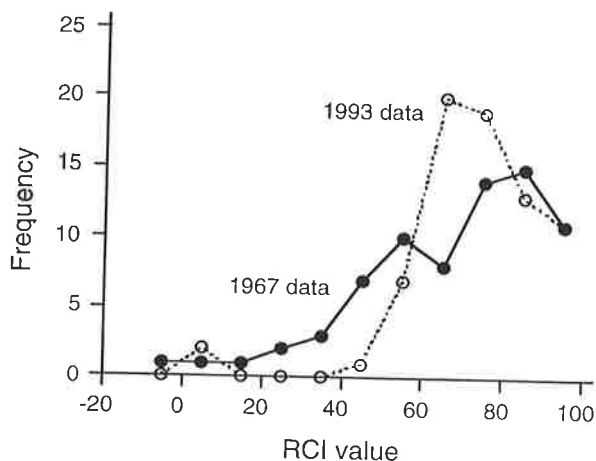
Figure 8.3 *Relative improvement in fossil record quality from 1967 (Harland et al., 1967) to 1993 (Benton, 1993). During these 26 years, gaps in the record were filled, and there is a clear shift in the distribution of RCI (Relative Completeness Index) values to the right from 1967 to 1993, indicating improvement in palaeontological knowledge (significant shift at $p < 0.05$; t-test and non-parametric signs and Wilcoxon signed ranks tests). Based on data in Benton and Storrs, 1994*

stochastic processes, and unbiased errors in data are stochastic. Likewise, Raup and Sepkoski (1986; Raup, 1991) suggested that sloppiness in big data sets actually strengthens the case for repeated or periodic patterns. Their argument was simply that if a signal is detected, that signal must be pretty powerful if it is to shine through the mess of inaccuracies and omissions in a global-scale database. They predicted that the signal would strengthen as the database was improved and corrected, since erroneous noise would be filtered out. Benton and Storrs (1994, 1996) presented independent evidence that the database on stratigraphic ranges of families of fossil tetrapods *has* improved in the past 25 years (Figure 8.3). This idea has to be extended to tests on other databases, and tests of whether evolutionary patterns *do* improve in terms of their statistical certainty as data are cleaned up.

The Answers: Recycling of Data is a Good Thing

The *trades unionist* can also be answered. The suggestion that data analysts should be intimately familiar with the taxonomy and stratigraphy of the materials they are studying is initially appealing. Ideally, perhaps, palaeobiologists should always collect their own primary data. This would mean that they would know the strengths and limitations of their data and

they would be better able to code the ranges of variation of certain attributes and the error in dating precision. In many cases, statistical analysts do collect their own data, and their work is generally enhanced by that means. However, this is not always the case, and the use of second-hand data has to be defended.

Four responses may be given. The first is the key one.

1. Some databases have to be too large-scale for any one person, or even team of people, to collect everything they need first-hand, and to maintain a level of consistency in data coding (Raup, 1991). If the question concerns the radiation of the insects, the extinction of the dinosaurs and the Cambrian explosion, it is quite impossible to gather together the necessary global data from specimens and outcrops. This is the justification behind publication of major compilations of data on the fossil record such as Harland *et al.* (1967), Sepkoski (1982, 1992), and Benton (1993). It is feasible, and usually essential, to collect first-hand for studies of single sedimentary basins, or single taxonomic groups. If data analysts had to collect all their information from specimens and outcrops, many key questions in palaeobiology would be ruled out. And our biological and geological colleagues would be astounded at our limited vision!

2. A general response is that the scientific literature exists and, on the whole, it is a high-quality vehicle for recording data, and its prime purpose is to transmit data to other scientists who may wish to use it. Scientists publish papers to tell the world about their work and, generally, they record their data accurately and in recognizable ways that allow comparison between publications by different scientists. Why not use the literature?

3. A key practical point is that analysts should make their data sets available for open scrutiny. This openness is essential so that others may test the assertions that have been made, a procedure precisely analogous with the requirement of experimental laboratory sciences that a scientist should make it possible for others to repeat his/her experiments. Until recently, the best way to make new data sets available was to publish them on paper in journals. However, editors do not always welcome huge data matrices that will drive most of their readers mad with eye fatigue. (A solution may be to use the World Wide Web, where data may be lodged and then accessed by anyone. An advantage over paper versions is that others may download the data directly into their own computers, hence saving time and avoiding the introduction of copying errors. A risk is that the proud owner of the Web site will continually tweak and improve the data set so that later re-runs will not precisely replicate what was done with an earlier generation of the database. See 'Methods' below.) If data sets have already been published, the requirement of openness is readily met.

4. A positive defence of using second-hand data is that there is merit in the practice. The data were accumulated in the first place by one or more palaeontologists for certain purposes, but the analyst often wishes to use them to test quite different ideas. This means that the critic cannot then argue that the analyst is guilty of collecting data in a biased way to favour a particular hypothesis.

The case for re-using published data can be made forcibly by referring to the work of Raup and Crick (1981). They re-analysed some classic records of bed-by-bed occurrences of the ammonite *Kosmoceras* in the Callovian of eastern England, compiled in the 1920s by Ronald Brinkmann. Brinkmann collected ammonites assiduously in the brickpits around Peterborough, and he recorded every occurrence of some 3000 specimens to the nearest centimetre. His purposes were mainly biostratigraphic and taxonomic. Raup and Crick (1981) re-analysed the data to determine short-term patterns of evolution. They could argue that Brinkmann's records were unbiased by any considerations of ammonite evolutionary patterns: he had no axes to grind about phyletic gradualism or punctuated equilibrium. In addition, the quality of the data was superb and represented many person-years of work, which would be foolish to repeat. In any case, the work could no longer be repeated, since most of the brickpits have since been filled with domestic refuse from London.

The Answers: Now is the Right Time

The *luddite* and the *utopian* need only a brief response. Clearly scientists must take advantage of new techniques if these allow new insights and new ways of solving the big questions. Palaeontologists have embraced the use of computers for many purposes, and manipulation of large databases is a perfect application.

Now is the right time for every question about the history of life to be tackled. The utopian is overly optimistic. We will never have all the data, so why procrastinate?

George Gaylord Simpson, perhaps the founder of analytical palaeobiology, expressed precisely this view in the introduction to his *Tempo and Mode in Evolution* (Simpson, 1944, p. xviii): 'For almost every topic discussed in the following pages the data are insufficient. The student who attempts interpretations under these circumstances does so in the face of certainty that some of his conclusions will be rejected. It is, however, pusillanimous to avoid making our best efforts today because they may appear inadequate tomorrow. Indeed, there would be no tomorrow for science if this common attitude were universal. Facts are useless to science unless they are understood.'

Equally, there is no need for critics to set up barriers to defend their patches against the analytical palaeobiologist. Unsupported, but authoritative, assertions that some part of the fossil record is inadequate for studies of mass extinctions, or that some other part cannot offer information on rates of evolution, are unhelpful. *All* parts of the fossil record are available for investigation, and it is for the analyst to defend the appropriateness of the data and the analytical technique. The critic may respond on those specific claims, and seek to demonstrate why the data do not support the interpretations made by the analyst. There is no place for bluster and bombast.
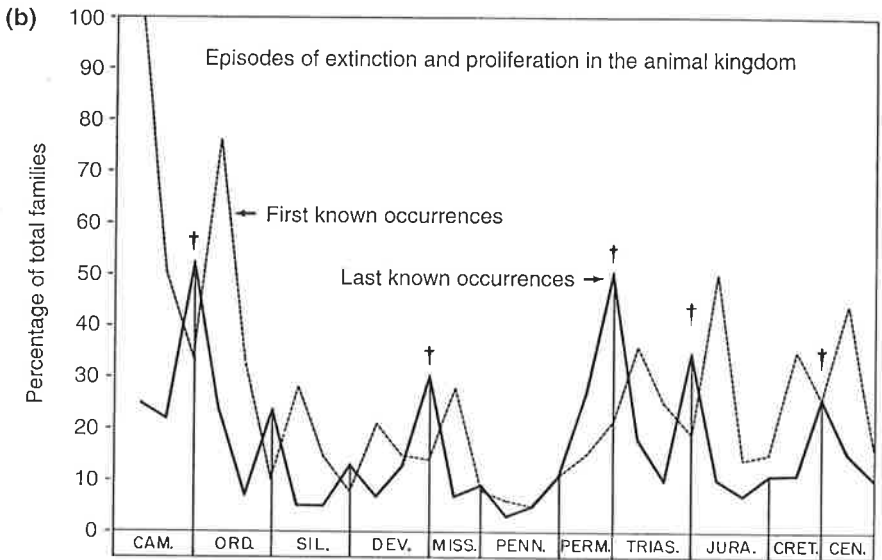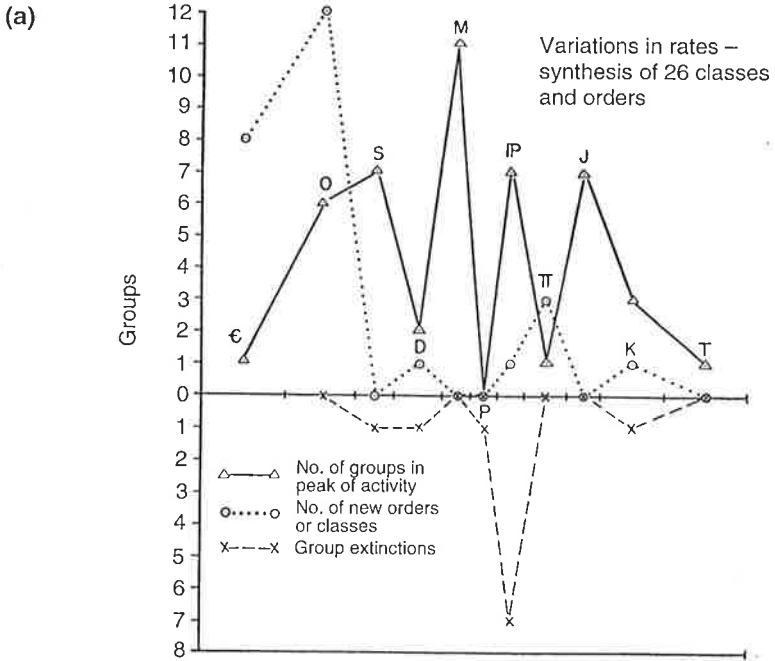
HISTORY

Palaeontological work on diversification through time had a promising early start, but was then virtually abandoned. John Phillips, the noted English geologist and palaeontologist, computed a database of all known fossil species, and produced the world's first diversity curve of all life through time (Phillips, 1860). This work stood in isolation for 80 years. Then, after this long gap, George Gaylord Simpson began his proselytizing efforts for analytical palaeobiology (Simpson, 1944). Simpson showed how some key questions concerning trends and rates of evolution could be tackled with simple and elegant quantitative assessments.

Simpson's book was one of four (Dobzhansky, 1937; Mayr, 1942; Huxley, 1942) that were critical in directing evolutionists back to Darwin and away from some of the metaphysical debates of the 1920s. He wrote at a time when biologists and palaeontologists interested in evolution were drawing together and making a self-conscious effort to unite the findings of genetics, ecology, systematics and palaeontology into what is now termed the modern synthesis, or neoDarwinism.

The young Turks of this time were stimulated to design new research programmes that crossed traditional boundaries. A positive result of this effort to combine palaeontology and biology into a seamless whole was the symposium volume edited by Jepsen *et al.* (1949), in which Stanley Westoll published his classic paper on rates of evolution in lungfishes based on the history of acquisition of characters and Bryan Patterson calculated rates of evolution of taeniodont mammals. Other palaeontologists reviewed the fossil record and geological time in terms of evolutionary rates.

The first tentative efforts to survey the nature of the history of life in a quantitative way were made in the early 1950s. One of the first attempts, by Sloss (1950), presented plots of diversity through time for species of various skeletonized planktonic and benthic marine organisms. This was followed by publication of work by several authors in a special issue of *Journal of Paleontology* in 1952. Here, Henbest (1952) presented data on patterns of

(a)

Variations in rates – synthesis of 26 classes and orders

Groups

△———△ No. of groups in peak of activity
o·····o No. of new orders or classes
x———x Group extinctions

(b)

Episodes of extinction and proliferation in the animal kingdom

Percentage of total families

← First known occurrences

Last known occurrences →

CAM. | ORD. | SIL. | DEV. | MISS. | PENN. | PERM. | TRIAS. | JURA. | CRET. | CEN.

diversification and extinction of foraminifera, Cooper and Williams (1952) on brachiopods and Moore (1952) on crinoids. Camp (1952) investigated major faunal replacements among vertebrates, and identified cases of outright extinction, competitive replacement, and expansion. Simpson (1952) presented the first plots of patterns of origination of orders, families, and genera of vertebrates through the Phanerozoic, but he found little evidence to tie these to major physical changes on the Earth.

The most remarkable paper in the 1952 volume was by Newell (1952), who analysed huge databases on patterns of diversification among invertebrates. With the enthusiasm of youth, he collected information on the stratigraphic occurrences of 9000 genera recorded to period level, and he presented plots of generic diversity of graptolites, ostracodes, foraminifera, bryozoans, corals, brachiopods, echinoderms, ammonoids, nautiloids, trilobites and fishes. He identified sets of clades that diversified in similar ways (corals and crinoids; brachiopods, bryozoans and ostracodes; foraminifera and echinoids), presaging the factor-analytic efforts of Flessa and Imbrie (1973) and Sepkoski (1981). He also combined the information from all the groups sampled to try to identify times of global-scale diversification and extinction (Figure 8.4a). He noted three phases of peak diversification (Ordovician, Mississippian, Jurassic) and he identified the profound end-Permian mass extinction, and lesser phases of extinction in the Silurian, Devonian and Cretaceous. He tied much of the pattern of diversification and extinction to physical causes, perhaps changes in sea level.

Much of this pioneering work of the early 1950s was ignored, and Norman Newell (1963, 1967) continued as virtually a lone voice through the 1950s and 1960s. He was able to further expand and improve his database on marine invertebrates by extracting numerical data from the new *Treatise on Invertebrate Paleontology* (Moore *et al.*, 1953–1992), which gave comprehensive generic-level summaries of most groups. In his 1963 paper, Newell presented a plot of diversification of 2500 families through the Phanerozoic, now recorded at epoch level. This was refined in his 1967 paper, in which he also gave clear evidence for mass extinctions at the end of the Cambrian, Devonian, Permian, Triassic and Cretaceous periods (Figure 8.4b). Independent attempts to document the diversity of well-skeletonized marine invertebrates were also presented by Müller (1961) and Gregor (1968).

The beginning of the modern era is probably marked by four independent efforts from 1967 and 1969, Newell's (1967) paper, Bretsky's (1968, 1969) summaries of marine invertebrate assemblages through time, *The Fossil*

Figure 8.4   (opposite) *Two classic studies of the fossil record of marine invertebrates from the 1950s and 1960s: patterns of origination and extinction through the Phanerozoic, according to data available in 1952 (a) and 1967 (b). Based on data in Newell, 1952 (a) and Newell, 1967 (b)*

Record (Harland *et al.*, 1967), and Jim Valentine's (1969) survey of diversifica-
tion through time of the well-preserved skeletonized fauna of marine shelves
(foraminifera, sponges, corals, bryozoans, brachiopods, molluscs, arthropods,
and echinoderms). *The Fossil Record* was the first published comprehensive
database designed specifically for studies of the nature of the history of life. In
it, Cutbill and Funnell (1967) provided exhaustive analyses of the data, and
plots of diversification, origination and extinction for each phylum and for all
life, presented for the first time at the stratigraphic level of the stage.

Of these pieces of work in the late 1960s, the synthesis by Valentine (1969),
and subsequent commentaries (Valentine, 1970, 1973, 1974; Valentine and
Moores, 1972), finally marked the beginning of a wider interest in the pattern
of the history of life based on large-scale compilations of data. Valentine
extracted data on the stratigraphic distributions of phyla, classes, orders and
families of the commoner groups of marine invertebrates from *The Fossil
Record* and from the *Treatise on Invertebrate Paleontology*, and he discovered a
variety of patterns indicating rising diversity through time, but with numer-
ous set-backs (Figure 8.5a). Valentine focused on the well-documented
groups in order to achieve the best-supported plot of diversification through
the Phanerozoic. He assumed that the less well-preserved groups probably
followed similar patterns of diversification, so his graphs would act as a
proxy for the diversification of life as a whole.
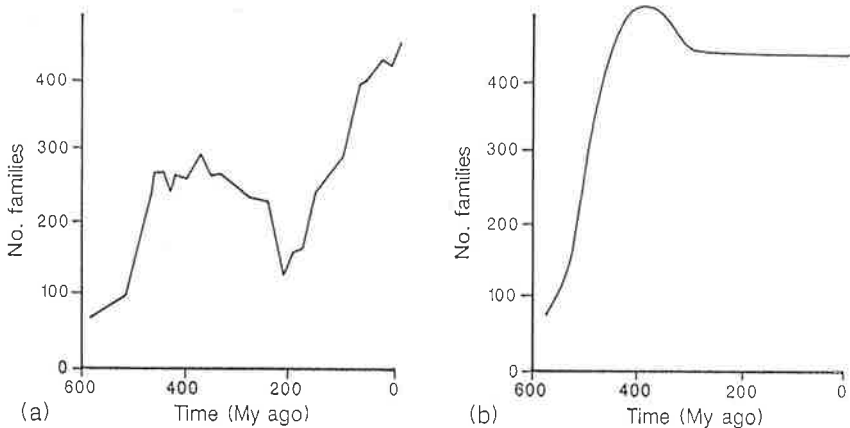
Figure 8.5   *Comparison of the empirical (a) and bias-simulation models (b) for diversifi-
cation of well-skeletonized marine invertebrates through the Phanerozoic. The empirical
pattern (a) is a literal reading of changes in diversity of families, and the bias-simulation
model (b) is a theoretical construct that purports to show the true pattern of diversification
after corrections for the poorer Palaeozoic fossil record and lower levels of study of such
materials. Based on data in Valentine, 1969 (a) and Raup, 1972 (b)*

Raup (1972) brought the work to a wider audience, and focused the debate precisely on questions of data-base quality. He compared Valentine's (1969) literal reading of the fossil record, in which the empirical pattern was accepted as close to reality, with a bias-simulation model which took account of statistically systematic bias in the fossil record (Figure 8.5b). As Raup (1972) suggested in his abstract, 'The increase in the number of marine species since the Paleozoic may be more apparent than real.' He suggested that the empirical pattern discovered by Valentine (1969), and indeed by Müller (1961), Newell (1963, 1967), Cutbill and Funnell (1967) and Gregor (1968), documented the gaps in the fossil record rather than the true diversities. He suggested that the low diversity levels of the Palaeozoic could as readily be attributed to a relative shortage of rocks, poorer preservation of fossils, less investigation by palaeontologists and other factors.

Raup's (1972) paper set the agenda for analytical palaeontologists for the next decades. Efforts were made to estimate the true scale of biases of rock volume, available rock area, palaeontologist interest and the like. In the end, the jury came out in favour of the empirical model: a variety of semi-independent data sets on the history of life yielded consistent patterns (Sepkoski *et al.*, 1981), and modelling to compensate for sampling bias (Signor, 1982) produced patterns that were closer to the empirical pattern (Figure 8.5a) than the bias-simulation pattern (Figure 8.5b).

The discovery, between 1972 and 1982, that a literal reading of the fossil record produced results on global diversity patterns that were probably close to the truth was helpful for palaeontologists. This finding legitimized the programme of data gathering and testing for patterns of large-scale evolution, and it gave palaeontologists confidence in their work when presenting it to the world at large. Most informed non-palaeontologists accept that fossil record data are of good quality. If anything, their faith is stronger than that of many palaeontologists: witness the rapid acceptance of Raup and Sepkoski's (1984) discovery of a periodic pattern of extinction events by many astrophysicists and cosmologists, but its automatic rejection by nearly all palaeontologists!

Several major programmes of analysis were pursued in the 1970s and 1980s, and it is possible to highlight only one or two, in the fields of diversification and extinction. Sepkoski (1978, 1979, 1984) engaged in a programme to explain the empirical pattern of diversification of marine animals (Figure 8.6a). He argued that equilibrium modelling from ecology was an appropriate tool, and he presented evidence that the irregular pattern of increase in global diversity of marine animals could be explained by an interplay of logistic patterns of diversification (slow lead-up, rapid increase, maintenance of dynamic equilibrium) and mass extinction. On the other hand, Gould and Calloway (1980), Benton (1983, 1987a), Hoffman (1985a, 1989), and others, argued that much of the history of life was non-competitive expansion,
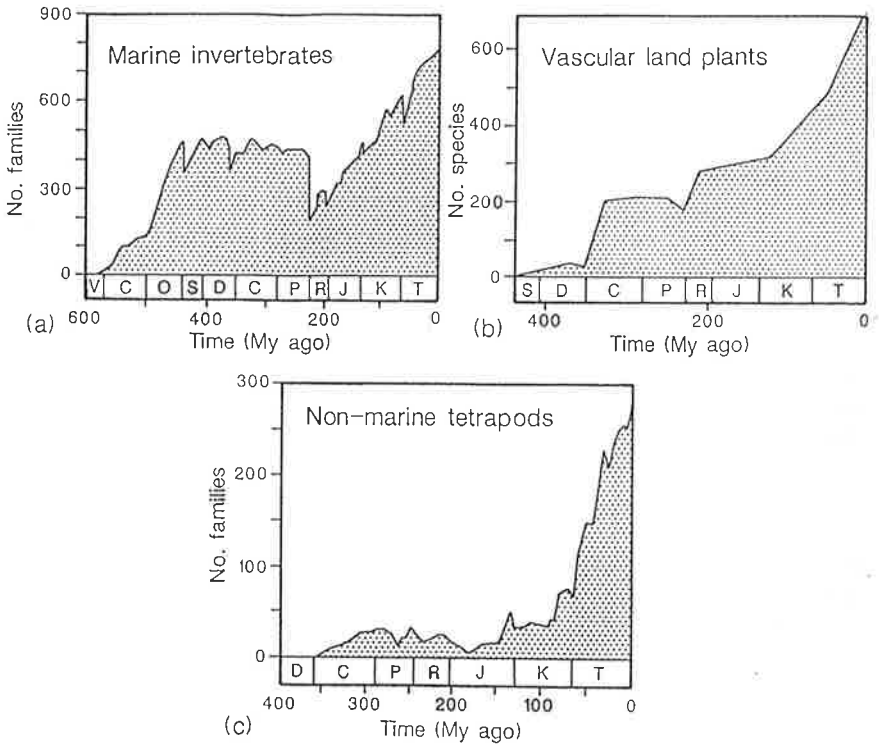
Figure 8.6   *Diversification of marine animal families (a), vascular land plant species (b), and continental tetrapod families (c) through the Phanerozoic. Based on data in Sepkoski, 1984 (a), Niklas et al., 1983 (b), and Benton, 1985 (c)*

sometimes triggered by major extinctions. Bambach (1977, 1985) investigated changes within marine faunas, and found evidence for expansion of species diversity through time which was related to an expansion of ecospace occupied. Niklas *et al.* (1980, 1983) presented data on the diversification of vascular land plants (Figure 8.6b) and Benton (1985, 1987b, 1989) did the same for continental tetrapods (Figure 8.6c).

Broad-scale studies of extinction events included Raup and Sepkoski's (1982) attempt to find a statistical test that would distinguish mass extinctions from background extinction. This proved impossible, but Jablonski (1986) identified patterns of selectivity during the K/T mass extinction that differed from those operating in normal background times. Raup and Sepkoski (1984, 1986) also found evidence for a periodicity of 26 My in the post-Palaeozoic record of major extinction events. This opened up a marvellous interdisciplinary debate on predictable extraterrestrial impacts on the Earth, and what caused them.

These, and many more themes, were explored in the 1970s and 1980s. It would be impossible to attempt to document the huge array of related studies that have been published since 1990, a heady mix of large-scale empirical work, statistical testing and mathematical modelling. Some multi-author review volumes provide a flavour: Kauffman and Walliser (1990), McNamara (1990), Ross and Allmon (1990), Taylor and Larwood (1990), Gilinsky and Signor (1991) and Jablonski *et al.* (1996). Three broad themes are selected for discussion, however, each controversial, and each of consider-able importance in understanding the nature of the history of life. First, there is a brief discussion of some basic techniques in manipulating large databases in palaeontology.

METHODS

Large-scale analytical work in palaeobiology can be pursued in a variety of ways. Early work was done almost entirely by hand: information was extrac-ted from published works, entered on cards, sorted and totalled by hand. Rates of origination and extinction were calculated by hand and graphs were also plotted manually. Computers are now, of course, routinely used for these tasks, sometimes mainframe computers, but more commonly, personal computers.

There is still a great deal of hand work in collecting the data, in checking published records of family, genus and species names, and of localities and ages of occurrences. At the core of most large-scale work in palaeobiology is a standard spreadsheet package, such as Microsoft Excel. A typical spread-sheet might consist of 100 rows, one for each of the standard stratigraphic stages, and a number of columns, one for each measure (e.g. duration of stage [in My], number of families present, number of families originating, number of families becoming extinct, total extinction rate, per-taxon extinction rate, and so on). It is possible to extract simple statistics within the spreadsheet package (e.g. total, mean, standard deviation for each measure). It is also possible to calculate whole columns of figures almost instantly: for example, extinction rates can be calculated by entering a simple formula, such as 'divide column 4 by column 2'.

It is possible to plot a variety of graphs within a spreadsheet package like Excel, but at this stage it is sometimes better to transfer the data to a graph-plotting package or a statistics package. It is easy to export Excel spreadsheet data (saved as 'text') to most other software. A graph-plotting package like Kaleidagraph is quick and easy to use, and it allows a great deal of editing so that publication-quality graphs may be plotted. Statistics pack-ages like SYSTAT or PALSTAT offer most of the tests that are required. Business statistics packages are of less value since they do not offer all the

tests a palaeontologist requires. In many cases, especially where some measure is plotted against geological time, nonparametric, or even time-series, tests are appropriate (see Chapter 7).

A final step may be to publish the data set in some way. This may be done either in the form of a book (e.g. Harland *et al.*, 1967; Sepkoski, 1982, 1992; Benton, 1993) or a computer disk (accompanying Sepkoski, 1992). The book at least provides a permanent reference, but the user has to extract and key in data. Sepkoski's disk is an advance in that it saves retyping, but disks must often be translated in some way to make them compatible with the computer and the software available to the researcher.

The ideal solution may now be to make the data available on the World Wide Web (WWW). The WWW provides an important new tool for data transfer among researchers. Raw data sets may be lodged on the home computer of the researcher, and made available for downloading by any interested person. The advantages of the system are that: (1) it provides an open form of scientific communication to enable testing of published hypotheses; (2) the researcher has control of the arrangement of his/her own data sets, and can find out who is using the material; (3) data transfer is instant, with no need to send disks through the post, and it allows easy transfer between widely different computer systems and software; (4) copying errors are minimized since the data set may be cut and pasted directly from the spreadsheet in which the data were assembled during the research phase into the web site, and then to any other person's computer; (5) this is a way to make data available that journals are often reluctant to publish.

The Palaeobiology Group in Bristol has opened a series of research-related databases, including a *The Fossil Record 2* site and a site for cladistic data sets and a site for tests of cladograms vs. stratigraphy (http:// www.palaeo.gly.bris.ac.uk/services.html). Data may be either downloaded automatically, or cut-and-pasted from the WWW site directly into the reader's own software. At the moment, these sites are compiled in the html language, which is relatively straightforward. Most html compilers act rather like glorified word processing packages. It is easy to cut and paste great masses of text from other documents, and they require only a rapid pass through to insert some specific html codes, and they are then ready for use.

Further development of the WWW for data exchange is possible. *The Fossil Record 2* site is linked to a search engine that allows enquirers to search the whole database for a particular family, order or phylum. Further keyword searches might be possible in the future. More elaborate facilities are available at the *Plant Fossil Record* site (http://www.uel.ac.uk/palaeo/) where the user may search by plant name, geological age or geographic location. In addition, the raw records may be plotted on palaeogeographic maps, and other calculations may be carried out. These more advanced capabilities are programmed in Java language, which is much more difficult than html.

The possibilities for manipulation of large databases in palaeontology are expanding all the time. A key advance has been the growth in storage capacity and speed of personal computers, and matching improvements in software. These capabilities will no doubt improve further. Perhaps, the development of WWW sites for palaeontological researchers will add a further dimension, not only in making other people's data readily available, but in providing remote workspaces for specialized calculations and plots. These technological advances have assisted palaeontologists in their exploration of some major questions in macroevolution, three of which are now discussed.
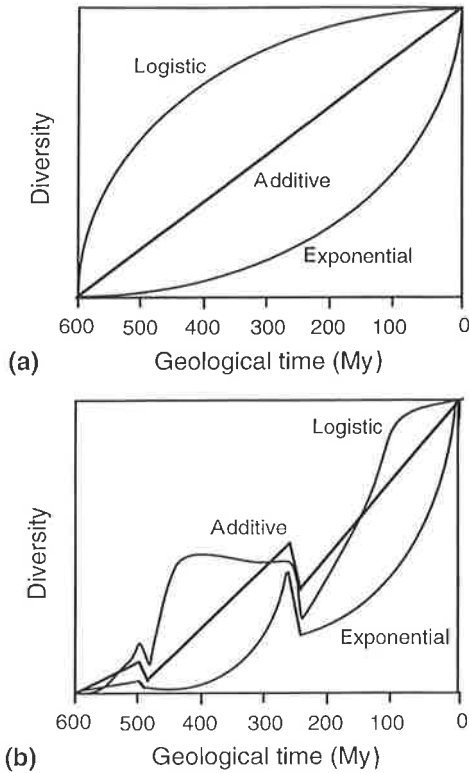
THE DIVERSIFICATION OF LIFE

There are many different routes that life might have followed in diversifying from a single species some time around 3500 My ago to the present total, estimated as 10–50 million species (May, 1990). These can be represented as a straight line, an exponential curve, and a logistic curve, first as ideal uninterrupted models (Figure 8.7a), and second with some mass extinctions superimposed (Figure 8.7b).

The straight-line model represents additive increase, simply the addition of a fixed number of new species in each unit of time. (The increase is, in this example, and the others, a net increase, that is, true increase minus extinctions.) In terms of an evolutionary branching model, additive increase would mean that, through time, speciation rates have declined, or that extinction rates have increased regularly at a rate sufficient to mop up the excess speciations. Such a model has been proposed by Walker (1985) and Hoffman and Fenster (1986).

The exponential model is more understandable in terms of a branching model of evolution. If speciation and extinction rates remain roughly constant, then there will be regular doubling of diversity within fixed units of time. This model was proposed first by Cailleux (1950, 1954) and, in passing, by Benton (1995a), based on a compilation of data on the history of families of organisms through time (Benton, 1993). Further evidence has been found for such an exponential pattern based on line-fitting calculations (Hewzulla *et al.*, 1999).

The logistic curve model involves one or more classic S-shaped curves, each consisting of a long period of slow diversity increase, a rapid rise and then a plateau. Such a model implies the existence of global equilibrium diversity levels, corresponding to the plateau(x). The logistic model has gained considerable credibility, and two different logistic patterns have been found by curve-fitting tests on data on familial diversity. In studies of the diversification of marine animal families, Sepkoski (1978, 1979, 1984) found evidence for a short plateau in the Cambrian (*c.* 40 My), and a longer one

Figure 8.7   *Theoretical models for the diversification of life, (a) in the absence of major perturbation, and (b) with two mass extinctions superimposed. In each case, the upper curve is the logistic or equilibrium model, the middle curve is the additive or straight-line model, and the lower curve is the exponential model. Based on Benton, 1997*

from the Ordovician to the Permian (*c.* 250 My). In an analysis of all marine and continental families, based on a different data set (Benton, 1993), Courtillot and Gaudemer (1996) identified one plateau lasting for most of the Palaeozoic (*c.* 300 My), and the beginning of a plateau phase in the Late Miocene or Pliocene, some 5–10 My ago. Logistic curves were also applied to large databases on marine invertebrate diversity by Carr and Kitchell (1980), and on vascular land plant diversity by Knoll *et al.* (1984). Kitchell and Carr (1985) used coupled logistic curves.

It is hard to provide a clear test of which of these kinds of curves fits the data (Figure 8.8) best. In all cases, investigators accept that the curve fits are not perfect, since the patterns of generally increasing diversity are offset by many drops in diversity, some associated with major mass extinctions, others with extinction events of more local scale, or affecting only certain taxa.

When such perturbations are excluded, proponents of the exponential and logistic models claim to have found curves that fit the empirical data well. However, different data sets (Sepkoski, 1992; Benton, 1993) give subtly different results, although they agree in the broad-scale features, such as the timing and relative magnitudes of mass extinctions and smaller extinction events, and in the relative rates of diversification.

Large-scale plots of the diversification of life seemingly cannot yet distinguish between patterns of unfettered expansion (exponential curves) and those of long-term steady-state conditions (logistic curves). This is an important problem to resolve, since it goes to the heart of our understanding of evolution: do species evolve within a tight straitjacket imposed by their interactions with other organisms (the equilibrium view), or has much of evolution been limited only by the capacity of organisms to enter new ecospace (the expansionist view)?

THE NATURE OF RADIATIONS

Another long-term debate in analytical palaeobiology has been the nature of clade radiations. The classic assumption is that such radiations are 'adaptive', in other words, driven by the acquisition of a new character that gives the radiating group abilities that are superior to those of any competitors. The radiation may take place in the absence of competitors (expansion), or it may be at the expense of some other competing group (competitive). Two alternative approaches have been explored. First, Raup *et al.* (1973) suggested that clade radiations could arise randomly, not driven by any deterministic cause, whether competition with another taxon or expansion associated with a specific new adaptation or opportunity. This stochastic viewpoint (see also Hoffman and Fenster, 1986; Hoffman, 1989) excludes the classic competitive model, as well as any expansion model that depends on a broad-scale deterministic process of diversification. Investigation of individual case studies has suggested (Gould and Calloway, 1980; Benton, 1983, 1987a, 1991; Jablonski, 1986) that many that had been identified as competitive replacement radiations turn out to have been expansions into empty ecospace. However, there is strong support for the importance of competitive interactions between clades (e.g. Rosenzweig, 1995). How can these widely different viewpoints be resolved?

One approach to assessing the relative importance of competition and expansion would be to multiply the numbers of case studies by engaging in an extensive census of all cases of postulated biotic replacement. The *census* approach has been used, for example, in evolutionary ecology, in attempts to quantify the role of competition in evolution: Connell (1983), and others, summarized all published accounts of competition in nature, and assessed

(a)

(b)

(c)
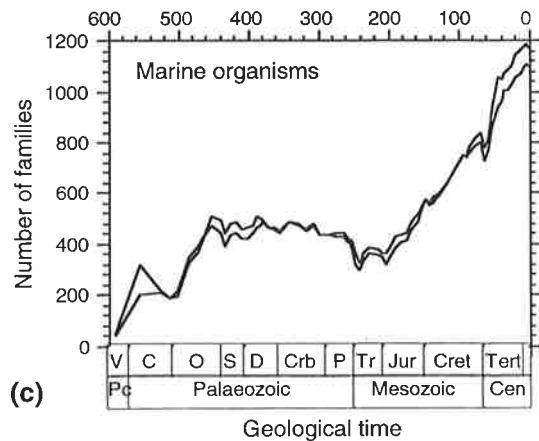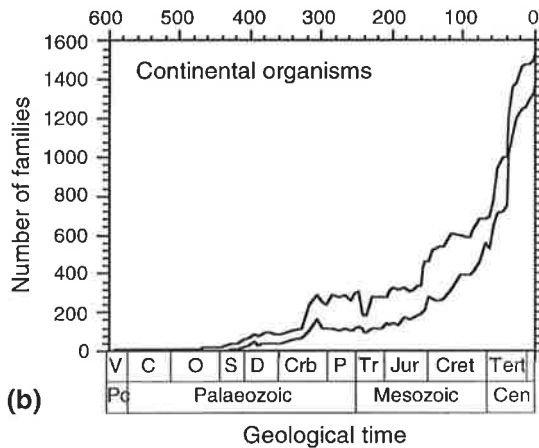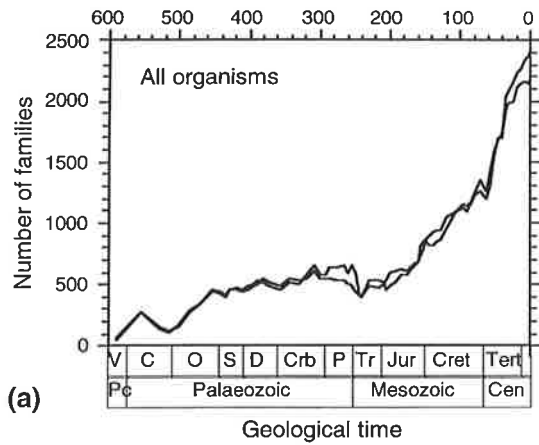
the results of such competition in each case. The census approach can offer suggestive evidence about the prevalence of one or other model in nature, but it cannot provide an absolute measure. This is because censuses are dependent on counts of the available observations, and these might reflect the reality of nature, or they might be biased by the skill and energy of the protagonists on one side of the debate.

These problems can be overcome by a modification to the census method. If the study is restricted to one clade, it should be possible to scrutinize all cases, and to score them. This then provides a complete census, and does not depend on what has been published, and what has not. In addition, if the scoring is done by one person, some measure of consistency is introduced. Such a *comprehensive census* has been carried out recently on the Tetrapoda (Benton, 1996a,b), and it suggests that expansion has been much more prevalent than competition in biotic replacements.

The clade Tetrapoda is suitable for such a comprehensive evolutionary ecological census since its members are relatively completely known today, the species are relatively well defined and ecologies are relatively well known. Fossil taxa are also relatively easy to identify and their ecologies may be determined readily in a general way. The fossil record of vertebrates is also adequate (Maxwell and Benton, 1990; Norell and Novacek, 1992; Benton, 1994, 1995b; Benton and Storrs, 1994; Benton and Hitchin, 1996): the order of appearance of taxa in the fossil record matches their order as predicted from cladograms and molecular phylogenies, and new collecting efforts rarely turn up any surprises.

Tetrapods (amphibians, reptiles, birds and mammals) arose in the Devonian Period, perhaps 380 My ago, and the Tetrapoda is a well-characterized clade. The group has diversified since the Devonian from, presumably, a single species to 23 500 species today (Wilson, 1992), and from a single initial

Figure 8.8   (opposite) *Patterns of the diversification of life through time, plotted for all organisms (a), continental organisms (b) and marine organisms (c), in terms of changes in numbers of families extant per stratigraphic stage. In each graph, a maximum and minimum is shown, based on a combination of stratigraphic and habitat-preference information. The minimum measure includes only families recorded as definitely present within each stratigraphic stage, or as definitely spanning that stage, and only families designated as restricted solely to the marine or continental realm. The maximum measure includes also all doubtful stratigraphic attributions of families, and all equivocal and shared habitat designations. The sum of minimum measures for continental and marine organisms is equal to the minimum measure for all taxa together, but the sums of maximum measures do not equal the maximum measure for all taxa, because families with equivocal environmental assignments, and those which occur in both marine and continental settings are counted as both marine and continental. Abbreviations: C, Cambrian; Cen, Cenozoic; Crb, Carboniferous; Cret, Cretaceous; D, Devonian; Jur, Jurassic; O, Ordovician; P, Permian; Pc, Precambrian; S, Silurian; Tert, Tertiary; Tr, Triassic; V, Vendian. Based on data in Benton, 1995a*
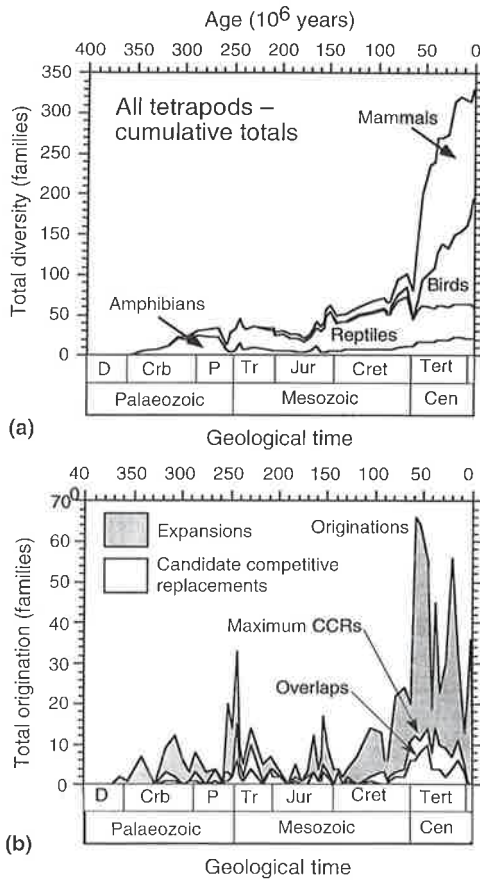
Figure 8.9 *Diversification of tetrapod families through time (a), and their styles of origination (b). Data are from* The Fossil Record 2 *(Benton, 1993) and are based on all 840 non-singleton families of tetrapods. A total of 194 singleton families were excluded, those that are currently based on single specimens or single species from single localities. In (b), all originations of tetrapods of all habitats are shown, as well as overlaps and of maximum candidate competitive replacements (CCRs). CCRs were identified by comparison of pairs of families. First, stratigraphic range charts were plotted for each combination of body size, diet, and habitat, and maximum geographic ranges of each family were noted. Then, the point of origin of each of the 840 families was scrutinized to determine whether it was a CCR or an expansion. CCR cases were subdivided into overlaps, where the stratigraphic range of the family overlapped another family, and situations where the family apparently originated at the precise time of extinction of another (gap 0), or after a gap of one or two (gap 1, 2) stratigraphic stages. Overlaps give evidence that the two families could have encountered each other, while the gap 0, 1, and 2 cases allow for possible incompleteness of the fossil record. CCRs are plotted as overlaps and maximum CCRs (the sum of all overlaps and gap 0, 1, and 2 cases). Abbreviations as for Figure 8.8. Based on data in Benton, 1996a,b*

family in the Devonian to 360 families today (Benton, 1993). The empirical curve for the diversification of tetrapod families (Figure 8.9a) looks more like an exponential curve than a logistic curve. Indeed, there is little evidence for a plateau anywhere, unless the bending-over of the upper cumulative curve in the Miocene, 20 My ago, indicates the beginning of a modern global steady-state diversity level. The accuracy of the empirical curve cannot be proved and it is possible that true diversities in the first half of the graph were much higher than is indicated here. There is no evidence, however, for a major bias in the tetrapod fossil record, whereby more ancient families are seriously under-represented.

Benton (1996a,b) found that most tetrapod families could not have orig-inated after competition with another family, and that at least seven-eighths (or at the lowest, three-quarters) of families must have originated by expan-sions. This had been hinted at by a previous study (Benton, 1990) in which it was shown that the earliest tetrapods were fish-eaters that lived in and around shallow waters. Jumps in familial diversity were associated with the addition of new guilds to existing ecosystems. New diets were acquired, such as insects, tetrapods, leaves, seeds, fruit and molluscs, and new habitats were entered, such as fully terrestrial, arboreal, subterranean, marine and aerial. In all such cases, the new families were exploring previously untap-ped ecospace and the familial origins must be interpreted as expansions.

The new quantified study of tetrapod familial originations (Benton, 1996a,b) showed that 13% of familial origins *could* be (but need not be) explained by competitive interaction with a pre-existing family on a reason-able estimate. Even using a maximal estimate of possible competitive interac-tions, where it was assumed that the fossil record was extraordinarily incom-plete, the proportion of familial originations that were candidate competitive replacements (CCRs) rose to 26%. The highest proportions of CCRs were in fully terrestrial habitats (19–36%), with lowest figures for freshwater habitats (2–11%) and for arboreal habitats (0%).

If this test is considered to have been a valid one, then competition in clade origination has to be assigned a minor role. Tetrapod evolution has shown continued expansion over the past 380 My, with little evidence of steady states or inter-clade competition. The result may be peculiar to tetrapods or to vertebrates. Equivalent tests have not yet been performed for other groups of organisms, but it seems unlikely that their modes of evolution would differ significantly from those of tetrapods.

TIMING OF MAJOR EXTINCTION EVENTS

An important assertion was made above, that large-scale databases, particu-larly those listing familial distributions in time, may be modified substan-
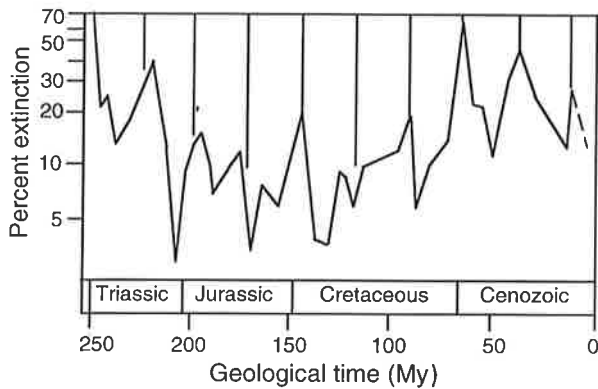
Figure 8.10 *Periodic pattern in the record of familial extinctions of marine animals in the past 250 My. Peaks of high extinction intensity appear to follow a 26 My periodic cycle. Based on data in Raup and Sepkoski 1984*

tially by new research, but macroevolutionary conclusions based upon them are robust. There is one possible exception. Raup and Sepkoski (1984, 1986) reported that they had found evidence that extinctions among marine animals occurred periodically, every 26 My over the past 250 My (Figure 8.10). This finding was based on Sepkoski's (1982) data compilation, and confirmed from the revised edition (Sepkoski, 1992). Some critics suggested that the database itself was seriously flawed in various ways, and that these flaws had contributed to the false identification of apparently periodic peaks of high extinction rate. The criticisms focused on a number of issues.

Time scale employed

Hallam (1984) and Hoffman (1985b) pointed out that the periodic pattern depended on the time scale that was used: changes in age dates for stage boundaries affected the relative timing of the events. There are currently several global standards for geological time available, each of which differs in some age assignments. Sepkoski (1989) tested the timing of extinction events with four then current time scales, and he found evidence for a 26 My periodicity for all, although the Harland *et al.* (1982) time scale provided the best fit. Use of the new Harland *et al.* (1989) time scale eliminates periodicity from the earlier Jurassic events (Sepkoski, 1996a). However, until one of these schemes can be shown to be more accurate than the others, a decisive test of the chronometric criticism is impossible.

The nature of stratigraphic divisions

Hoffman (1985b) argued that it was inevitable that one in four stages would

be characterized as an extinction event, since the extinction rate could only increase or decrease between any sequential pair of stages. The probability of one or other eventuality is 0.5, and the probability of an increase followed by a decline would then be 0.25. An extinction event is defined by an increase followed by a decline, and this would occur in one in four stages, all other things being equal. However, there is no predictable requirement that the peaks should occur in every fourth stage, just once every four, so this argument for a stochastic pattern of periodicity fails (Sepkoski, 1989).

Periodicity based on paraphyletic taxa

Patterson and Smith (1987) argued that, for echinoderms and fishes at least, the periodic pattern of extinction was an artefact of ill-defined family-level taxa. These authors found that 55% of Sepkoski's (1982) families of echinoderms and fishes were not clades, but monotypic, paraphyletic or even polyphyletic groupings. The majority were paraphyletic taxa, in other words, families which did not include all descendants of their basal taxon. The truncation of a paraphyletic family corresponds either to 1, a time when one or more species in the family are regarded as so different from their predecessors that the new part of the lineage is given a different name, or 2, a time when the family was nearly, but not quite, driven to extinction (i.e. many species disappeared, but not all). Paraphyletic families then are defined by pseudoextinctions, times of evolutionary transformation or partial extinction. Patterson and Smith (1987) found that the periodic signal for extinctions among fishes and echinoderms resided in the faulty data, and that the 45% of monophyletic families showed no such periodic signal. They argued then that the periodic signal came from the statistical noise in the data set, and not from the small proportion of robust data.

There are still many issues to be explored here. A better comparison of paraphyletic and monophyletic families of fishes and echinoderms would have been to generate a new data set of purely monophyletic families. This has now been done (with the usual caveat that of course the listings are merely a best effort, and they could change) by Simms *et al.* (1993), Patterson (1993) and others. Tests have yet to be carried out for periodicity in the 'all-cladistic' data sets.

A further question concerns the level of analysis. Patterson and Smith (1987) argued that their discovery that the periodic signal resided in the 55% of mainly paraphyletic familial data disproved the notion of periodic patterns of extinction in the real world. However, in a positive interpretation, paraphyletic families can be said to mimic species-level extinction better than cladistically-defined monophyletic families (Sepkoski, 1989), especially when the fossil record is incomplete (Sepkoski and Kendrick, 1993). Perhaps Raup and Sepkoski (1984, 1986) were finding a species-level effect by the use

of paraphyletic familial data? If this were the case, it could be argued that paraphyletic taxa are preferable since they document species-level behaviour that evolutionists really wish to understand better than monophyletic families.

On the other hand, too many paraphyletic taxa might falsify the signal. It could be argued that Raup and Sepkoski's (1984, 1986) analysis was self-fulfilling, since the rates of familial extinction were enhanced at certain times since true extinctions and pseudoextinctions were combined. A few insignificant species-level extinctions during a known extinction event could feed through to a pseudoextinction of a paraphyletic family, which would in turn acquire much greater significance when it is combined with other familial data. Hypothetically, some extinction peaks could be based entirely on familial pseudoextinctions. It is possible to imagine a situation where the concurrent loss of 10 species out of 1000 (0.1% loss) scales up to the loss of 10 paraphyletic families out of 50 (20% loss). A low level of background extinction becomes a mass extinction! This has not been the case in the post-Palaeozoic at least, since generic-level data confirm the periodic signal (Sepkoski, 1989, 1990, 1996a).

Local effects

Perhaps some of the less robust extinction peaks in Raup and Sepkoski's (1984, 1986) analysis are founded on local events that have been scaled up by monographic effects. If two or three more of the 10 peaks turn out to be artefacts, then the periodicity argument must falter.

Perhaps the three extinction peaks in the Jurassic are not real. Hallam (1986) suggested that the Early Jurassic (Pliensbachian) and end-Jurassic (Tithonian) extinction peaks do not represent global events, but were restricted to NW Europe. Indeed, these two supposedly global events could have been magnified from some small-scale events in the Jurassic rocks of Yorkshire and southern Germany. The ammonites and other marine fossils of Yorkshire and Dorset in England, and the Stuttgart area of SW Germany, have been collected in vast abundance for two centuries, and they are heavily documented. The fully marine rocks of the Lias and of the Tithonian in both England and Germany are followed by marginal marine and continental sediments of the Middle Jurassic and Early Cretaceous respectively. Could these two, out of 10, supposedly periodic extinctions simply reflect facies changes in England and Germany? Little and Benton (1995) found that the Lower Jurassic extinction event was real and global, but much smaller than the well-known mass extinctions, and it appears to have been smeared over several stratigraphic zones of the Toarcian stage. Here is a further case where improvements in the database are perhaps producing real changes in interpretations. The Middle Jurassic peak was not identified by Raup and Sep-

koski (1984), but Sepkoski (1990, 1996a) has found a small enhancement in extinction intensity in the Callovian, on the basis of generic data. This event has yet to be confirmed by local and regional studies.

Alternative data compilations

Some alternative databases have confirmed periodicity, while others have not. Sepkoski's unpublished generic-level database on marine animals, consisting of (1996) 33 180 genera, provides a semi-independent test of patterns of diversification and extinction. It shows (Sepkoski, 1996a) patterns reminiscent of those revealed by familial data. Extinction intensities are, as expected, generally higher than for families, and all 10 predicted post-Palaeozoic extinction peaks were revealed, together with two smaller peaks, in the Carnian and Pliocene.

Benton's (1995a) work, however, does not support periodicity. His semi-independent analysis of diversification and extinction in the history of life is based on an alternative database (Benton, 1993), which included largely monophyletic families for certain phyla, where this was possible. This analysis confirmed the existence of some of the better-known extinction events in Raup and Sepkoski's (1984, 1986) time series: Late Permian (Kazanian–Tatarian), end-Triassic (Rhaetian) and end-Cretaceous, as well as three smaller events in the end-Jurassic (Tithonian), mid-Cretaceous (Aptian? and Cenomanian) and Late Eocene (Priabonian). The other three or four events postulated by Raup and Sepkoski (1984, 1986), and essential for any hypothesis of periodicity, were not evident in the new study. Perhaps, in this case, aspects of the databases, Sepkoski (1982, 1992) versus Benton (1993), have affected the results.

CONCLUSIONS

Palaeontology is perhaps unique among the sciences in that there has been opposition to certain kinds of data testing. Physicists are not afraid to ask big questions, nor are biologists, chemists, geochemists and geophysicists. This negative attitude from within the palaeontological profession may have been grafted on by a world of 'hard' scientists who have in the past sneered at palaeontologists as 'mere stamp collectors', in Ernest Rutherford's famous phrase. If he had been right in his contemptuous attitude, perhaps palaeontologists should retire gracefully and humbly to their museum garrets. Fortunately, Rutherford was hopelessly wrong: all Nature is not physics, and he should have known better than to sustain such an antiquated reductionist view of the world.

Palaeobiologists have tackled some of the large questions about evolution

in an increasingly quantitative way. Such studies began in the 1940s and 1950s, and they were championed by palaeobiologists of the stature of George Gaylord Simpson and Norman Newell. However, their work was largely ignored, and the majority of palaeobiologists saw little merit in such speculative exercises. This delayed the development of palaeobiology as an interdisciplinary science by at least 20 years. Large database work began again in the late 1960s and 1970s: growing acceptance of the usefulness of such work was matched by increasing expectations from scientists in other fields who sought palaeobiological evidence on evolution.

A glance through the pages of *Science* and *Nature* shows that there is a demand for palaeontological data on a variety of topics. Other scientists seek precise quantitative information on:

1. The diversification of life through time (What was the pattern? Has diversification involved addition of new ecospace, or subdivision of existing niches? Is there a global equilibrium for diversity?)
2. The narrative story of the history of life (Is the narrative broadly correct? What are the dates of major branching events in phylogeny for the calibration of molecular clocks? When did humans arise? Is the fossil record good enough, in whole or in part? Can the variable quality of the fossil record be predicted? Do cladograms give an accurate view of phylogeny? What about molecular trees?)
3. Biotic replacements (What happens when previously isolated faunas come into contact? What happens when faunas are split apart by major plate movements? Do new groups arise mainly by displacing pre-existing groups, or by entering new ecospace?)
4. Mass extinctions (Are mass extinctions in a class of their own, or is there a continuum of scales of extinction event from small to large? Is there any ecological or geographic selectivity during extinction events and mass extinctions? Do major extinction events occur periodically or not? How long did mass extinctions take? How did life recover after major extinction events?)
5. Speciation (How do new species arise? How long does speciation take? How do species evolve after speciation? How long do species exist?)
6. Palaeobiology (How did ancient organisms live? How did the largest land animals and largest flying animals operate? Has the range of adaptations within groups increased or diminished through time?)

Can palaeobiologists answer these questions? Perhaps the larger numbers of researchers, the increasing sophistication of the databases, and developments in software have led to some firm solutions. No. Despite the increase in effort devoted to such questions, none of them can be answered firmly. However, at least the questions are now being asked and, more hopefully,

palaeobiologists are not ashamed to try to answer them. These are indeed exciting times to be involved in numerical palaeobiology!

ACKNOWLEDGEMENTS

REFERENCES

Archibald, J.D., 1996, *Dinosaur Extinction and the End of an Era*. Columbia University Press, New York, 237 pp.

Bambach, R.K., 1977, Species richness in marine benthic habitats through the Phanerozoic, *Paleobiology*, **3**: 152–167.

Bambach, R.K., 1985, Classes and adaptive variety: the ecology of diversification in marine faunas through the Phanerozoic. In J.W. Valentine (ed.), *Phanerozoic Diversity Patterns: Profiles in Macroevolution*. Princeton University Press, Princeton, NJ, pp. 191–253.

Benton, M.J., 1983, Dinosaur success in the Triassic: a noncompetitive ecological model, *Quarterly Review of Biology*, **58**: 29–55.

Benton, M.J., 1985, Mass extinction among non-marine tetrapods, *Nature*, **316**: 811–814.

Benton, M.J., 1987a, Progress and competition in macroevolution, *Biological Reviews*, **62**: 305–338.

Benton, M.J., 1987b, Mass extinctions among families of non-marine tetrapods: the data, *Mémoires de la Société Géologique de France*, **150**: 21–32.

Benton, M.J., 1989, Mass extinctions among tetrapods and the quality of the fossil record, *Philosophical Transactions of the Royal Society of London, Series B*, **325**: 369–386.

Benton, M.J., 1990, The causes of the diversification of life. In P.D. Taylor and G.P. Larwood (eds), *Major Evolutionary Radiations*. Systematics Association Special Volume 42, Clarendon Press, Oxford, pp. 409–430.

Benton, M.J., 1991, Extinction, biotic replacements, and clade interactions. In E.C. Dudley (ed.), *The Unity of Evolutionary Biology*. Dioscorides Press, Portland, OR, pp. 89–102.

Benton, M.J., 1993, *The Fossil Record 2*. Chapman and Hall, London, 839 pp.

Benton, M.J., 1994, Palaeontological data, and identifying mass extinctions. *Trends in Ecology and Evolution*, **9**: 181–185.

Benton, M.J., 1995a, Diversification and extinction in the history of life, *Science*, **268**: 52–58.

Benton, M.J., 1995b, Testing the time axis of phylogenies, *Philosophical Transactions of the Royal Society, Series B*, **348**: 5–10.

Benton, M.J., 1996a, On the nonprevalence of competitive replacement in the evolution of tetrapods. In D. Jablonski, D.H. Erwin and J.H. Lipps (eds), *Evolutionary Paleobiology*. University of Chicago Press, Chicago, pp. 185–210.

Benton, M.J., 1996b, Testing the roles of competition and expansion in tetrapod evolution. *Proceedings of the Royal Society of London, Series B*, **263**: 641–646.

Benton, M.J., 1997, Models for the diversification of life, *Trends in Ecology and Evolution*, **12**: 490–495.

Benton, M.J. and Hitchin, R., 1996, Testing the quality of the fossil record by groups and by major habitats, *Historical Biology*, **12**: 111–157.

Benton, M.J. and Storrs, G.W., 1994, Testing the quality of the fossil record: paleontological knowledge is improving, *Geology*, **22**: 111–114.

Benton, M.J. and Storrs, G.W., 1996, Diversity in the past: comparing cladistic phylogenies and stratigraphy. In M.E. Hochberg, J. Clobert and R. Barbault (eds), *Aspects of the Genesis and Maintenance of Biological Diversity*. Oxford University Press, Oxford, pp. 19–40.

Bretsky, A.J., 1968, Evolution of Paleozoic marine invertebrate communities, *Science*, **159**: 1231–1233.

Bretsky, A.J., 1969, Evolution of Paleozoic benthic marine invertebrate communities, *Palaeogeography, Palaeoclimatology, Palaeoecology*, **6**: 45–59.

Cailleux, A., 1950, Progression géometrique du nombre des espèces et vie en expansion, *Comptes rendus de la Société Géologique de France*, **13**: 222–224.

Cailleux, A., 1954, How many species? *Evolution*, **8**: 83–84.

Camp, C.L., 1952, Geological boundaries in relation to faunal changes and diastrophism, *Journal of Paleontology*, **26**: 353–358.

Carr, T.R. and Kitchell, J.A., 1980, Dynamics of taxonomic diversity, *Paleobiology*, **6**: 427–443.

Connell, J.H., 1983, On the prevalence and relative importance of interspecific competition: evidence from field experiments, *American Naturalist*, **122**: 661–696.

Cooper, G.A. and Williams, A., 1952, Significance of the stratigraphic distribution of brachiopods, *Journal of Paleontology*, **26**: 326–337.

Courtillot, V. and Gaudemer, Y., 1996, Effects of mass extinctions on biodiversity, *Nature*, **381**: 146–148.

Cutbill, J.L. and Funnell, B.M., 1967, Numerical analysis of the fossil record. In W.B. Harland, C.H. Holland, M.R. House, N.F. Hughes, A.B. Reynolds, M.J.S. Rudwick, G.E. Satterthwaite, L.B.H. Tarlo and E.C. Willey (eds), *The Fossil Record.* Geological Society of London, London, pp. 791–820.

Dobzhansky, T., 1937, *Genetics and the Origin of Species*. Columbia University Press, New York.

Flessa, K.W. and Imbrie, J., 1973, Evolutionary pulsations: evidence from Phanerozoic diversity patterns. In D.H. Tarling and S.K. Runcorn (eds), *Implications of Continental Drift to the Earth Sciences.* Academic Press, London, pp. 245–285.

Fremlin, J., 1979, Dinosaur death: the unconscious factor, *New Scientist*, **81**: 250–251.

Gilinsky, N.L. and Signor, P.W. (eds), 1991, *Analytical Paleobiology.* University of Tennessee, Knoxville, TN, 216 pp.

Gould, S.J. and Calloway, C.B., 1980, Clams and brachiopods – ships that pass in the night, *Paleobiology*, **6**: 383–396.

Gregor, C.B., 1968, The rate of denudation of post-Algonkian time, *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen*, **71**: 22–30.

Hallam, A., 1984, The causes of mass extinctions, *Nature*, **308**: 686–687.

Hallam, A., 1986, The Pliensbachian and Tithonian extinction events, *Nature*, **319**: 765–768.

Harland, W.B., Cox, A.V., Llewellyn, P.G., Pickton, C.A.G., Smith, A.G. and Walters, R.R., 1982, *A Geologic Time Scale.* Cambridge University Press, Cambridge, 131 pp.

Harland, W.B., Holland, C.H., House, M.R., Hughes, N.F., Reynolds, A.B., Rudwick,

M.J.S., Satterthwaite, G.E., Tarlo, L.B.H. and Willey, E.C., 1967, *The Fossil Record: A Symposium with Documentation.* Geological Society of London, London, 827 pp.

Henbest, L.G., 1952, Significance of evolutionary explosions for diastrophic division of earth history – introduction to the symposium, *Journal of Paleontology*, **26**: 299–318.

Hewzulla, D., Boulter, M.C., Benton, M.J. and Halley, J.M., 1999, Evolutionary patterns from mass originations and mass extinctions, *Philosophical Transactions of the Royal Society of London, Series B*, (in press).

Hoffman, A., 1985a, Biotic diversification in the Phanerozoic: diversity independence, *Palaeontology*, **28**: 387–391.

Hoffman, A., 1985b, Patterns of family extinction depend on definition and geological timescale, *Nature*, **315**: 659–662.

Hoffman, A., 1989, *Arguments on Evolution.* Oxford University Press, New York.

Hoffman, A. and Fenster, E.J., 1986, Randomness and diversification in the Phanerozoic: a simulation, *Palaeontology*, **29**: 655–663.

Huxley, J.S., 1942, *Evolution: the modern synthesis.* Chatto & Windus, London.

Jablonski, D., 1986, Causes and consequences of mass extinctions: a comparative approach. In D.K. Elliott (ed.), *Dynamics of Extinction.* John Wiley, New York, pp. 183–229.

Jablonski, D., 1988, Estimates of species durations, *Science*, **238**: 360–363.

Jablonski, D., Erwin, D.H. and Lipps, J.H. (eds), 1996, *Evolutionary Paleobiology.* Chicago University Press, Chicago, 484 pp.

Jablonski, D. and Raup, D.M., 1995, Selectivity of end-Cretaceous marine bivalve extinctions, *Science*, **268**: 389–391.

Jepsen, G.L., Simpson, G.G. and Mayr, E. (eds), 1949, *Genetics, Paleontology, and Evolution.* Princeton University Press, Princeton, NJ, 474 pp.

Kauffman, E.G. and Walliser, O.H. (eds), 1990, *Extinction Events in Earth History.* Springer, Berlin, 432 pp.

Kitchell, J.A. and Carr, T.R., 1985, Nonequilibrium model of diversification: faunal turnover dynamics. In J.W. Valentine (ed.), *Phanerozoic Diversity Patterns.* Princeton University Press, Princeton, NJ, pp. 277–309.

Knoll, A.H., Niklas, K.J., Gensel, P.G. and Tiffney, B.H., 1984, Character diversification and patterns of evolution in early vascular plants, *Paleobiology*, **10**: 34–47.

Little, C.T.S. and Benton, M.J., 1995, Early Jurassic mass extinction: a global long-term event, *Geology*, **23**: 495–498.

Maxwell, W.D. and Benton, M.J., 1990, Historical tests of the absolute completeness of the fossil record of tetrapods, *Paleobiology*, **16**: 322–335.

May, R.M., 1990, How many species? *Philosophical Transactions of the Royal Society, Series B*, **330**; 293–304.

Mayr, E., 1942, *Systematics and the Origin of Species.* Columbia University Press, New York.

McNamara, K.J. (ed.), 1990, *Evolutionary Trends.* Belknap, London, 368 pp.

Moore, R.C., 1952, Evolutionary rates among crinoids, *Journal of Paleontology*, **26**: 338–352.

Moore, R.C., Teichert, C., Robison, R.A. and Kaesler, R.L., 1953–1992, *Treatise on Invertebrate Paleontology.* Geological Society of America and University of Kansas Press, Lawrence, KS.

Müller, A.H., 1961, *Grossabläufe der Stammesgeschichte- Erscheinungen und Probleme.* Gustav Fischer, Jena.

Newell, N.D., 1952, Periodicity in invertebrate evolution, *Journal of Paleontology*, **26**, 371–385.

Newell, N.D., 1963, Crises in the history of life. *Scientific American*, **208**: 76–92.

Newell, N.D., 1967, Revolutions in the history of life, *Special Paper of the Geological Society of America*, **89**: 63–91.

Niklas, K.J., Tiffney, B.H. and Knoll, A.H., 1980, Apparent changes in the diversity of fossil plants; a preliminary assessment, *Evolutionary Biology*, **12**: 1–89.

Niklas, K.J., Tiffney, B.H. and Knoll, A.H., 1983, Patterns in vascular land plant diversification. *Nature*, **303**: 614–616.

Norell, M.A. and Novacek, M.J., 1992, The fossil record and evolution: comparing cladistic and paleontologic evidence for vertebrate history, *Science*, **255**: 1690–1693.

Patterson, C., 1993, Osteichthyes: Teleostei. In M.J. Benton (ed.), *The Fossil Record 2*. Chapman and Hall, London, pp. 621–663.

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

Phillips J., 1860, *Life on Earth: its origin and succession*. Macmillan, London.

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

Raup, D.M., 1991, The future of analytical paleobiology. In N.L. Gilinsky and P.M. Signor (eds), *Analytical Paleobiology. Short Courses in Paleontology*, No. 4, The Paleontological Society, Knoxville, TN, pp. 207–216.

Raup, D.M. and Crick, R.E., 1981, Evolution of single characters in the Jurassic ammonite *Kosmoceras*, *Paleobiology*, **7**: 200–215.

Raup, D.M., Gould, S.J., Schopf, T.J.M. and Simberloff, D.S., 1973, Stochastic models of phylogeny and the evolution of diversity, *Journal of Geology*, **81**: 525–542.

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

Raup, D.M. and Sepkoski, J.J. Jr, 1984, Periodicity of extinctions in the geologic past, *Proceedings of the National Academy of Sciences of the USA*, **81**: 801–805.

Raup, D.M. and Sepkoski, J.J. Jr, 1986, Periodic extinctions of families and genera, *Science*, **231**: 833–836.

Rosenzweig, M.L., 1995, *Species Diversity in Space and Time*. Cambridge University Press, Cambridge, 436 pp.

Ross, R.M. and Allmon, W.D. (eds), 1990, *Causes of Evolution: a paleontological perspective*. University of Chicago Press, Chicago, 479 pp.

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

Sepkoski, J.J. Jr, 1979, A kinetic model of Phanerozoic taxonomic diversity. II. Early Phanerozoic families and multiple equilibria, *Paleobiology*, **5**: 222–251.

Sepkoski, J.J. Jr, 1981, A factor analytic description of the Phanerozoic marine fossil record, *Paleobiology*, **7**: 36–53.

Sepkoski, J.J. Jr, 1982, A compendium of fossil marine families, *Milwaukee Public Museum, Contributions in Biology and Geology*, **51**: 1–125.

Sepkoski, J.J. Jr, 1984, A kinetic model of Phanerozoic taxonomic diversity. III. Post-Paleozoic families and mass extinctions, *Paleobiology*, **10**: 246–267.

Sepkoski, J.J. Jr, 1989, Periodicity in extinction and the problem of catastrophism in the history of life, *Journal of the Geological Society of London*, **146**: 7–19.

Sepkoski, J.J. Jr, 1990, Evolutionary Faunas. In D.E.G. Briggs and P.R. Crowther (eds) *Palaeobiology: a synthesis*. Blackwell Scientific, Oxford, pp. 37–41.

Sepkoski, J.J. Jr, 1992, A compendium of fossil marine animal families, 2nd edn. *Milwaukee Public Museum, Contributions in Biology and Geology*, **83**: 1–156.

Sepkoski, J.J. Jr, 1993, Ten years in the library: how changes in taxonomic databases affect perception of macroevolutionary pattern, *Paleobiology*, **19**: 43–51.

Sepkoski, J.J. Jr, 1996a, Patterns of Phanerozoic extinction: a perspective from global databases. In O.H. Walliser (ed.), *Global Events and Event Stratigraphy.* Springer, Berlin, pp. 35–51.

Sepkoski, J.J. Jr, 1996b, Competition in macroevolution: the double wedge revisited. In D. Jablonski, D.H. Erwin and J.H. Lipps (eds), *Evolutionary Paleobiology.* University of Chicago Press, Chicago, pp. 211–255.

Sepkoski, J.J. Jr, Bambach, R.K., Raup, D.M. and Valentine, J.W., 1981, Phanerozoic marine diversity and the fossil record, *Nature,* **293**: 435–437.

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

Sharpton, V.L. and Ward, P.D. (eds), 1990, *Global Catastrophes in Earth History: an interdisciplinary conference on impacts, volcanism, and mass mortality.* Geological Society of America, Boulder, CO, 631 pp.

Signor, P.W., 1982, Species richness in the Phanerozoic: compensating for sampling bias, *Geology,* **10**: 625–628.

Simms, M.J., Gale, A.S., Gillilland, P., Rose, E.P.F. and Sevastopulo, G.D., 1993, Echinodermata. In M.J. Benton (ed.), *The Fossil Record 2.* Chapman and Hall, London, pp. 491–528.

Simpson, G.G., 1944, *Tempo and Mode in Evolution.* Columbia University Press, New York, 237 pp.

Simpson, G.G., 1952, Periodicity in vertebrate evolution, *Journal of Paleontology,* **26**: 359–370.

Sloss, L.J., 1950, Rates of evolution, *Journal of Paleontology,* **24**: 131–139.

Taylor, P.D. and Larwood, G.P. (eds), 1990, *Major Evolutionary Radiations.* Clarendon Press, Oxford, 437 pp.

Valentine, J.W., 1969, Patterns of taxonomic and ecological structure of the shelf benthos during Phanerozoic time, *Palaeontology,* **12**: 684–709.

Valentine, J.W., 1970, How many marine invertebrate fossil species? A new approximation, *Journal of Paleontology,* **44**: 410–415.

Valentine, J.W., 1973, Phanerozoic taxonomic diversity: a test of alternate models, *Science,* **180**: 1078–1079.

Valentine, J.W., 1974, Temporal bias in extinction among taxonomic categories, *Journal of Paleontology,* **48**: 549–552.

Valentine, J.W. and Moores, E.M., 1972, Global tectonics and the fossil record, *Journal of Geology,* **80**: 167–184.

Walker, T.D., 1985, Diversification functions and the rate of taxonomic evolution. In J.W. Valentine (ed.), *Phanerozoic Diversity Patterns.* Princeton University Press, Princeton, NJ, pp. 311–334.

Wilson, E.O., 1992, *The Diversity of Life.* Harvard University Press, Cambridge, MA, 406 pp.