Bicentennial Review

Marine diversity through the Phanerozoic: problems and prospects

ANDREW B. SMITH

Department of Palaeontology, The Natural History Museum, Cromwell Road, London SW7 5BD, UK
(e-mail: a.smith@nhm.ac.uk)

Abstract: The fossil record provides direct evidence of how diversity has changed over time, but cannot be taken at face value. Diversity curves constructed from counting taxa in the rock record are seriously biased by unevenness of geographical and stratigraphical sampling effort, inequality in the rock record available for sampling, and inconsistent taxonomic practice. Sample standardization removes some bias, but does not overcome more general incompleteness problems. Modelling that accounts for potential biases is a newer approach but needs accurate estimates of rock record and consistent taxonomic data. Uncertainty remains over whether the steep rise in diversity over the last 100 Ma is real or reflects sampling bias. The repeated rise and fall of marine diversity over time correlates closely with the areal extent of sedimentary deposits and independent estimates of the quality of the fossil record, implying a common driving factor, namely tectonically mediated sea-level change. However, whether changes in diversity are primarily biological in origin, or reflect sampling artefact, remains contentious. There is a distinct possibility that many of the apparent rises and falls in diversity over the Phanerozoic, including most of the ‘mass extinctions’, arise either partially or entirely from rock record bias.

Ever since Phillips (1860) first sketched out the history of diversity, palaeontologists have been turning to the fossil record for insight into how life on Earth has changed over time. Whereas Phillips based his sketch on an intuitive knowledge of the British rock record, later workers, starting with Newell (1952, 1967) and Raup (1972), have employed a quantitative approach, counting the numbers of taxa described from each time interval. A major landmark was reached with the publication of Sepkoski’s marine family database and diversity curve (Sepkoski 1982). This summarized the distribution of nearly 3500 families through 72 time intervals, and provided the first detailed picture of marine diversity through the Phanerozoic (Fig. 1b). Later refinement of the family-level curve (Sepkoski 1993) and replication of the data gathering by a consortium of specialists (Benton 1993) made little discernible impact on the shape of the Phanerozoic marine diversity curve (Benton 1995; Fig. 1b). Meanwhile, Sepkoski had begun an even more monumental task, compiling data on the distribution of marine genera. His compendium of genus ranges was eventually published posthumously (Sepkoski 2002; Fig. 1a), although results using these data began appearing from 1986 onwards (Raup & Sepkoski 1986; Sepkoski 1996, 1997). Thus for 140 years palaeontologists have striven to establish the pattern of diversity change through the Phanerozoic by documenting the stratigraphical occurrence of fossils in the rock record with ever-increasing precision. However, precision is not the same as accuracy and these estimates are ultimately only as reliable as the rock record from which the fossils are extracted. The first doubts concerning the reliability of diversity curves based on a direct reading of the fossil record surfaced when Raup (1976) began exploring the quality of the rock record through time. He gathered data on the volume of sedimentary rock preserved at Epoch level and noted that the Phanerozoic diversity curve seemed to follow the rise in the volume of sedimentary rocks that have been preserved. He concluded that the rise in diversity towards the present day need not be real but might simply be an artefact of a steadily improving rock record. This created an initial flurry of concern, and set one of the key research agendas for the next decade, but in 1981 the problem was ‘solved’ when Sepkoski et al. (1981) showed that the same Phanerozoic diversity curve was recovered from a variety of approaches (species-level data, genus-level data, family-level data, trace fossil data, and comparative community diversity). Amazingly, Sepkoski et al. never addressed whether all five sets of data could have been distorted by the same basic rock and sampling biases.

As a result of the Sepkoski et al. (1981) paper, taxon counting became generally accepted as the preferred method for documenting Phanerozoic diversity patterns, and potential problems were given little consideration over the next decade. However, by the mid-1990s researchers were again starting to think about potential biases and their effect on the fossil record. The pioneering work by Holland (1995) highlighted the important control sequence stratigraphic cycles had on fossil sampling patterns, and the careful analysis by Miller & Foote (1996) of Ordovician diversity patterns threw the spotlight on problems of sampling bias. However, for the most part, such problems were seen as fine tuning or regional, not ones that would seriously affect the overall shape of the Phanerozoic diversity curve.

In 2001, however, the reliability of a Phanerozoic marine diversity curve based on raw taxonomic counts came under concerted attack. Smith (2001) and Peters & Foote (2001), using different metrics, highlighted the existence of a major systematic bias in the outcrop area of sedimentary rocks through time. Both showed that the rock record varied over time and found a strong and significant correlation between the amount of rock at outcrop and sampled diversity. Alroy et al. (2001) were also concerned about the effect of collection biases on the Phanerozoic diversity
Fig. 1. Phanerozoic diversity curves derived from counting the number of taxa present in each stage, with range interpolation. (a) Genus-level diversity, from Sepkoski (2002). (b) Family-level diversity from Benton (1995) and Sepkoski (1997).

What is wrong with the way marine diversity has been estimated in the past?

Prior to 2001 Phanerozoic diversity curves were constructed from a simple count of numbers of taxa recorded in any given time interval (usually the 72–77 stage-level intervals of Sepkoski (1982) and Benton (1993)). Compilations at any taxonomic level can be used to construct diversity curves, but Robeck et al. (2000) demonstrated that using more finely subdivided taxonomic groupings produced a more precise view of underlying diversity in the rocks. Furthermore, although there will always be a certain amount of error in taxonomic compilations, Sepkoski (1993) and Adrain & Westrop (2000) both demonstrated that such error was random and thus did not pose a serious problem to this approach.

The fossil record is of course notoriously incomplete, so to compensate for this a technique called range interpolation has been employed. Range interpolation removes some problems of a patchy fossil record by assuming that a taxon is present in each time interval between its first and last occurrence, whether or not it has actually been found in those time intervals. Because the fossil record is dominated by organisms with mineralized skeletons, the history of those taxa with hard parts is taken as a proxy for all marine diversity. The exact ratio of mineralized to unmineralized taxa is unimportant so long as it has remained broadly similar throughout the Phanerozoic. By assuming that sampling is more or less uniform through time, the relative numbers of taxa described from each time interval (or that cross boundaries between intervals) can be used as a measure of how diversity has changed.

This taxon-counting approach is simple to employ and seemingly robust to certain potential problems, but makes the following three critical assumptions: (1) all time intervals are equally well sampled; (2) preservation potential is uniform over time; (3) taxonomists partition taxa in a uniform manner. Each unfortunately is beset with problems.

Sampling of the rock record

There are two aspects of sampling that need to be considered: geographical bias and variation in sampling intensity.

Geographical bias. Jackson & Johnson (2001) and Johnson (2003) argued that any diversity curve constructed simply from cataloguing the numbers of fossils already described was doomed to failure because well-studied parts of the world that contribute most to taxonomic compilations were not necessarily representative of global diversity. Specifically, they found that the Neogene record of the tropics was woefully undersampled compared with temperate regions, a view later reinforced by Valentine et al. (2006).

European and North American data certainly contribute disproportionately to taxonomic compilations, simply because fossil collecting has been intensely pursued in those regions for much longer (Kidwell & Holland 2002). However, extreme unbalance in sampling between, say, Indo-Pacific faunas and those of temperate North America is no problem if this bias applies equally to all time intervals through the Phanerozoic.

Unfortunately, continental plates have migrated out of the tropics over time (Allison & Briggs 1993; Walker et al. 2002; Fig. 2). Because diversity is highest in the tropics, a long-term trend of decreasing diversity could be created artificially simply because the well-studied parts of the world have shifted over time from equatorial to temperate latitude through plate migration. Indeed, some palaeontologists are starting to factor out this bias from their analyses (e.g. Bush & Bambach 2004; Crampton et al. 2006b). On the other hand, the smaller-scale rises and falls in diversity from stage to stage that have been taken as the signatures of mass extinction and radiation cannot be explained by such slow changes in the positions of continental blocks (Smith 2001).
Sampling efficiency. If some time intervals are less well sampled than others for whatever reason, then simple taxon counting fails as a reliable proxy for diversity. A rise in numbers of fossil taxa collected and recorded might simply be due to better preservation, more outcrop availability and/or more collection effort. Miller & Foote (1996) and Alroy et al. (2001) correctly identified this as a major concern, and have been at the forefront of developing methods that can overcome such sampling problems.

The rock record as an unbiased sample
Various lines of evidence (gap analysis: Foote & Sepkoski 1999; Crampton et al. 2006a; Fara & Benton 2000: collector curves: Benton 1998; Paul 1998: phylogenetic trees: Benton et al. 2000) suggest that paleontologists have done a passably good job in exploring surface outcrops for their fossil content, at least in better-studied parts of the world. Of course, there is a strong geographical bias to these databases as discussed above; yet for any one region, rocks at outcrop have probably been approximately equally well sampled over the Phanerozoic. But are those rocks themselves an unbiased subset of past sedimentary deposits? It is this aspect that Smith (2001), Peters & Foote (2001, 2002), Smith et al. (2001), Crampton et al. (2003, 2006a) and Peters (2005, 2006a) have all expressed concern about. What these studies highlight is that the amount of sedimentary rock preserved has fluctuated significantly over time, as has the nature of the sedimentary environments captured.

Rock record bias has been calculated in two ways. Smith (2001) estimated post-Palaeozoic rock outcrop area for western Europe by counting the number of geological maps with outcrops assigned to each of the 42 stages (an equal grid sampling method) (Fig. 3). Peters & Foote (2001) used the number of named formations from a lexicon of geological names as a proxy for both rock extent and habitat heterogeneity (Fig. 4). Both found that the rock record showed a strong and worryingly close match to sampled diversity. More detailed study of the New Zealand rock record later found that measuring rock outcrop area was a more accurate estimator of number of fossil localities sampled (Crampton et al. 2003), whereas the number of formations might be a better estimator of habitat heterogeneity (Crampton et al. 2006a).

Computer simulation studies (Holland 1995, 2000; Holland & Patzkowsky 1999), backed up by detailed field studies (Patzkowsky & Holland 1996; Gale et al. 2000; Smith et al. 2001; Crampton et al. 2006a), have provided a broad understanding of how sequence stratigraphic packaging of sediments, driven by lower-order cycles of marine onlap–offlap, can bias perceptions of marine diversity. Transgressive phases to early highstand leave a good record as sediment accommodation space increases with rising sea level. In highstand phases accommodation space is low, progradation rates are high and much of the shelf tends to be covered in coarse-grained sediments, especially in siliciclastic systems, creating less conducive conditions for fossil preservation. Late highstand and regressive phases leave the poorest record, as the last sediments to be deposited are the first to be exposed and eroded from perched positions. Productivity of the carbonate factory is governed by the aerial extent of shallow illuminated sea floor and also varies through a sequence cycle, affecting preservation potential. Crampton et al. (2006a) observed the best preservation of molluscan faunas in mid-cycle (around time of maximum flooding) at the top of transgressive systems tract intervals, and poorest preservation towards the end of highstand system tract intervals (Fig. 5). Lower-order sequence cycles certainly play an important role in determining diversity patterns on a regional scale, but probably do no more than cause faunas to migrate across the shelf, passing in and out of the windows of preservation afforded to palaeontologists.

Highest-order cycles, driven by tectonic processes (primarily sea-floor spreading, mantle plume related phenomena, and continental assembly and disassembly: Dewey & Pitman 1998), are more complicated to interpret (see below). This is because major onlap–offlap cycles across continental blocks concurrently change the surface area of shallow marine habitats and the surface area of sedimentary deposits that eventually are preserved in the rock record. This creates distinct cycles of marine sediment abundance through the Phanerozoic that clearly follow higher-order sequence stratigraphic cycles (Figs 3–5).

Changing proportions of marine sedimentary rock preserved at outcrop mean that paleontologists have been dealt a biased deck of cards before even a single fossil is collected. Thus, even if diversity had remained constant over the Phanerozoic, the fact that successive time intervals preserve greater or lesser amounts of sedimentary rock at outcrop can create variation in the numbers of fossils recorded through time.

Change in the proportion of onshore to offshore sediments preserved in the record is probably as important as changes in the volume of rock preserved (Smith et al. 2001; Crampton et al. 2003, 2006a). The diversity of macrofossils preserved in the rock is a function of both the original biological diversity gradient and changing preservational potential across the shelf, controlled by sedimentation rates (e.g. Kidwell & Baumiller 1990). The ratio of onshore to offshore sediments is also affected by the angle of the continental shelf and thus varies through transgressive–regressive cycles and according to tectonic setting. The diversity of the New Zealand molluscan fauna drops in late highstand intervals (Crampton et al. 2006a), probably because of dilution by high sediment delivery in this siliciclastic setting (Cooper et al. 2006). A similar marked drop in diversity is recorded over the Cenomanian–Turonian boundary, when high-diversity faunas in oligotrophic onshore facies became displaced by lower-
diversity faunas, as sea levels rose and stratified oceanic conditions became more widespread across the continental shelf (Gale et al. 2000; Smith et al. 2001). Thus perceived diversity and the environment of origin of major clades (Jablonski & Bottjer 1991) are both potentially affected by changes in the proportion of onshore–offshore marine habitats captured by the rock record in any given time interval.

Changes in the extent of continental flooding affect both the absolute amount of sediment deposited over the cratons and the geographical extent of the habitats that those sediments capture, and so could affect biological diversity just as much as sampling, as is discussed more fully below.

**Fossil preservation**

The fossil record is palpably incomplete and, even within groups such as bivalve molluscs that might be expected to have good preservation potential, maybe a quarter of all taxa never enter the fossil record (Cooper et al. 2000; Smith et al. 2001). Thus perceived diversity and the environment of origin of major clades (Jablonski & Bottjer 1991) are both potentially affected by changes in the proportion of onshore–offshore marine habitats captured by the rock record in any given time interval.

Changes in the extent of continental flooding affect both the absolute amount of sediment deposited over the cratons and the geographical extent of the habitats that those sediments capture, and so could affect biological diversity just as much as sampling, as is discussed more fully below.

![Fig. 3. Global marine genus diversity v. rock at outcrop record for western Europe (France, England and Wales) since the late Triassic.](image-url)

(a) Log diversity of genera whose taxonomic duration is 45 Ma or less (data from Rhode & Muller 2005). (b) Marine sedimentary rock at outcrop, based on number of 1:63 360 (UK) and 1:80 000 (France) geological maps, from Smith (2001). (c) Detrended plots of genus diversity and rock at outcrop diversity (from Smith & McGowan 2005). (d) Second-order sequences in western Europe, from Hardenbol et al. (1998).
et al. (2006) could find evidence for major large-scale bias over
time in molluscan preservation. Furthermore, the relative com-
monness of a mollusc in the fossil record bears no correlation to
the robustness of its shell (as measured by size, thickness,
reinforcement and mineralogy) (Behrensmeyer et al. 2005).

Therefore the preservational bias created by shell composition
and durability on Phanerozoic diversity patterns is apparently
weak and largely random, even though it may be highly signifi-
cant in bed-to-bed comparisons and on a regional scale. Kidwell
(2005) speculated that this was possibly because each time
interval manages to capture an adequate range of taphonomic
settings.

Taxonomic practice

Taxonomic data are messy. Taxa of the same rank cannot strictly
be considered equal or even equivalent, as different taxonomists
may have applied very different recognition criteria when erect-
ing them. Although earlier studies (e.g. Adrain & Westrop 2000)
found no evidence that taxonomic errors and arbitrary taxonomic
splitting of lineages was systematically biased, Ausich & Peters
(2005) provided some evidence to the contrary. Using Ordovi-
cian–early Silurian crinoids as an example, they showed that
there was a preferential bias against extending taxonomic names
across the major Epoch boundaries. This creates an underestima-
tion of early Silurian biodiversity and an overestimation of rates
of extinction in the late Ordovician.

Studies that look at rates of occurrence or standing diversity
through time are relatively immune to taxonomic problems, as
only the presence of hidden synonymy (morphologically indis-
tinguishable taxa being given different names) might inflate
estimates. However, serious problems arise when last occurrences
of taxa are equated with extinction. Last occurrences of taxa are
not necessarily extinctions, but could represent arbitrary cuts
(pseudoextinctions) of a continuous lineage imposed by taxono-
mists engaged in chopping up long branches (Smith & Patterson
1988). At genus level maybe a third of last occurrences represent
such pseudoextinctions (Patterson & Smith 1989).

The trend recently has been to treat databases compiled at
finer taxonomic resolution (e.g. genus level) as more reliable
than those at a high taxonomic resolution (e.g. family or order).
This largely stems from the study by Robeck et al. (2000), which
showed, using computer simulation, that the finer the taxonomic
category used in a study the more precise a view of underlying
diversity resulted. However, their computer simulations only
explored the accuracy of methods assuming that the preserved

Fig. 4. Global marine genus diversity plotted against various rock record and
sampling bias estimates for the Phanerozoic. (a) Global marine genus-level
diversity taken from Sepkoski (2002) with single occurrences and poorly dated
occurrences removed. Geological time scale revised and smooth trend curve (dashed)
added by Rhode & Muller (2005). (b) As (a) but data detrended. (c) Estimated genus-
level sampling probability per stage based on analysis of first and last occurrences of
genera, from Foote (2003). (d) Number of marine formations in the USGS lexicon,
from Peters & Foote (2001). (e) Number of marine sedimentary rock sections in North
America based on the COSUNA charts (continuous line), from Peters (2005), and
number of through-ranging gaps in deposition (dashed line), from Peters
(2006a). (f) Major sequence stratigraphic cycles of North America, based on Sloss
(1963).
fossil record was an unbiased representation of true underlying biodiversity. Reduced to absurdity, species counts give the best estimate of how many species can be collected at outcrop, but species data are also the most affected by biases in the rock record. As the focus of study moves down the taxonomic category from order to family to genus to species, the average taxon duration declines, with the result that a greater proportion of taxa come to have durations equal to or shorter than the sizeable gaps in the rock record separating sediment bundles (Peters 2006a, b). Therefore, deriving diversity curves from higher taxa with longer average duration should help to smooth out some of the rock bias that affects genus-level databases.

Finally, higher taxa are being used as a proxy for numbers of species that existed in each time interval. Naturally, one would want the average number of morphospecies assigned to a higher taxon such as a genus or family to be constant through time. Unfortunately, this is not the case. Plotting numbers of described species and families through time (see Fig. 7) shows that significantly more species are assigned to a Neogene family than to a Palaeozoic family (Flessa & Jablonski 1985). Why this should be so is still unclear, but, if we want to use higher taxa as a measure of species diversity, then we also need to correct for this bias.

**New approaches to overcome biases of the fossil record**

In the face of overwhelming evidence that the fossil record is seriously biased by a wide range of factors, two approaches are currently being developed to try to improve diversity estimates:

- **Sample standardization**
  Both Jackson & Johnson (2001) and Alroy et al. (2001) recognized that sampling unevenness could be seriously biasing the Phanerozoic diversity curve, and that data standardization was needed. However, they did not agree on what that standardization procedure should be. Jackson & Johnson (2001) and Johnson (2003) argued that palaeobiologists should be applying standard ecological approaches that use abundance data to determine comparable measures of sampling completeness through rarefaction. Furthermore, they believed that new data needed to be gathered to remove biases in global coverage. From the number of specimens collected and the number of species those specimens represent one can easily determine how many species we would expect to find in each sample if a standard number of specimens were collected. This provides a clear and unambiguous way of comparing alpha diversity where sampling has been very unequal. The problem is that insufficient data of this form are yet available to allow a global biodiversity curve to be constructed.

  Alroy et al. (2001) opted for a different form of standardization, one that made best use of the available published literature. They used two methods, both of which employed randomized subsampling of taxonomic lists. The first involved rarefaction of individual taxonomic occurrences in the database, as previously employed by Miller & Foote (1996). The second method drew entire taxonomic lists and counted numbers of taxa added. For
each time interval a list of fossiliferous localities is compiled, along with a list of the taxa that each locality has yielded. Drawing localities at random from this pool until a predetermined number of species is reached and repeating many times provides a relative measure of taxonomic diversity.

Both sample size standardization and subsampling of taxonomic lists are methods for removing uneven sampling and thereby allowing meaningful comparison of average taxonomic diversity between time intervals. However, subsampling of taxonomic lists makes a number of unrealistic assumptions that are not met by real data, and can be strongly influenced by the spatial organization of the diversity they encompass (Bush et al. 2004). More critically, all methods of standardization simply remove the biases created by local incompleteness within known collections, they do not tackle the larger problem of general incompleteness, where faunas are not preserved in the first place (Foote 2001, 2003).

**Modelling large-scale preservation bias**

Instead of trying to standardize for sampling differences through time, an alternative approach is to estimate how preservation potential has varied and use this to adjust the raw diversity curve. This has been successfully applied to extinction studies where there are good data on sampling and species diversity, and a variety of methods can be employed to establish the relative completeness of the fossil record for each time interval (see Crampton et al. 2006a, and references therein). The raw taxon count is used to construct an initial diversity plot, which is then subsequently modified through applying correction factors based on the quality of the fossil record. However, this approach has yet to be applied to the global Phanerozoic marine diversity curve.

There are numerous papers that have tried to assess the quality of the fossil record (e.g. Foote & Sepkoski 1999; Benton et al. 2000), but these look at the record as a whole rather than stage by stage. Currently, four methods can provide the sort of high-resolution temporal data on the quality of the fossil record that are needed: gap analysis, ghost range analysis, modelling of origination and extinction patterns, and rock record analysis.

**Gap analysis.** Paul (1998) summarized how gap analysis can be used to determine the relative quality of the fossil record. Put simply, between the first and last occurrence of a taxon there will be time intervals when it has been captured (i.e. preserved, collected and identified) and time intervals where it has yet to be recorded (gap). The ratio of gap to known record gives a measure of the quality of the fossil record specific for that time interval. Periods with a high gap to observed ratio have comparatively poorer fossil records than periods with proportionally fewer gaps. A plot of the ratio of gap to observed distributions through the Phanerozoic provides an estimate of how successful we have been at capturing the taxa that are known to range through time intervals. This was the approach used by Paul (1998) to plot the quality of the Palaeozoic fossil record as exemplified by echinoderms, and was one of the techniques used by Crampton et al. (2006a) in assessing the New Zealand fossil record of Cenozoic molluscs (Fig. 5c). It has also been used to explore the quality of the Jurassic to Eocene continental fossil record of vertebrates (Fara & Benton 2000; Fara 2002). However, this method cannot account of any missing range beyond the first and last recorded appearance of a taxon, and also takes no account of pseudoextinction: name changes imposed by taxonomists onto a single lineage typically coinciding with a major preservation gap. To take account of these two problems we need a different approach.

**Ghost lineage analysis and molecular clocks.** The construction of phylogenetic trees encompassing large numbers of taxa provides an alternative method for estimating the relative quality of the fossil record over time. Phylogenetic relationships amongst taxa are constructed without reference to their stratigraphic occurrence, being based solely on the distribution of morphological characteristics. Yet when the resulting cladogram is optimized and calibrated against the observed stratigraphical distribution of fossils, the resultant tree identifies periods of time when lineages must have existed even though they have not yet been found (Norrell 1992). Counting both the inferred and observed ranges in such calibrated trees compensates for missing parts of the fossil record much better than range interpolation, as it allows for some range extension beyond first and last occurrence datums. Treating plesiomorphic taxa that are stratigraphically older than their sister group as potential ancestors avoids spurious range extensions (Smith 1994; Lane et al. 2005).

Phylogenetic trees constructed from molecular data or both morphological and molecular data can provide an additional estimate of origination times that is largely independent of the stratigraphic occurrence of fossils. Where morphological and molecular trees agree, multiple local (minimal age) fossil calibration points are used and methods that allow molecular rates of evolution to vary across the tree are employed, accurate and precise molecular estimates of divergence can be obtained (Near et al. 2005; Smith et al. 2006; Yang & Rannala 2006).

Molecular clocks are limited in use to groups that have diverse extant representatives, but can be used as probes to test and compare the quality of their fossil record at Epoch level (e.g. Teeling et al. 2005; Smith 2007).

So far, phylogenetic trees and molecular phylogenies have been used only in small-scale studies and/or across a limited number of taxa (mostly terrestrial vertebrates). The construction of a Phanerozoic curve mapping the quality of the marine fossil record at stage level from phylogenetic and molecular data has yet to be attempted. However, given the rate at which high-quality cladograms and molecular phylogenies have appeared over the last decade it will not be long before such an attempt can be made. The advantage of this approach is that it should be more effective at picking up sampling and rock record bias than gap analysis (Lane et al. 2005).

**Rate modelling of originations and extinctions.** Foote (2001, 2003, 2005) has been developing a sophisticated model-based approach that uses data on first and last occurrences of taxa to simultaneously estimate the most likely rates of origination, extinction and preservation for each time interval. His approach is to extrapolate the expected pattern of preserved first and last appearances corresponding to a given evolutionary model, starting from an initial specified time series of rates of origination, extinction and preservation. Optimization is used to find the set of rates that best predict the observed pattern of first and last occurrences.

Using this approach, Foote (2003) developed a curve for the rate of preservation (i.e. quality) of the fossil record through the Phanerozoic (Fig. 4c). He was also able to demonstrate that origination and extinction rates had varied over time, although the precise timing and duration of peaks in extinction or origination was sensitive to model assumptions. That is to say, we know that origination and extinction rates have varied in a way that is not directly predictable from sampling bias. However,
as noted by Foote, his approach still assumes uniformity of sampling with respect to environments and geography, and takes no account of taxonomic problems with the database such as pseudoextinction.

**Rock record.** A plot of how the rock record changes through the Phanerozoic can be used as a proxy for the quality of the fossil record. Counts of rock outcrop area or number of geological formations provide estimates of the number of fossiliferous localities and/or the environmental heterogeneity captured by the rock record at any one time interval. Smith (2001), Peters & Foote (2001, 2002) and Crampton et al. (2003) have all provided such estimates (Figs 3–5). An alternative approach is to measure the distribution of sedimentary hiatuses in the geological record (Peters 2006a). Reassuringly, the preservation rate plot of Foote (2003) and the rock record plot of Peters & Foote (2002), although derived from very different assumptions, show excellent agreement (Fig. 4c and d), which gives confidence to their approximate correctness.

As our knowledge of how the quality of the fossil record varies over time advances, more sophisticated methods will surely be used to correct the Phanerozoic marine diversity curve. The studies of Foote (2003), Bush & Bambach (2004) and Crampton et al. (2006a, b) hint at what might be achieved with high-quality datasets.

**Current controversies concerning the marine Phanerozoic diversity curve**

So far I have discussed why traditional methods of estimating Phanerozoic marine diversity, although precise, are no longer considered accurate in the face of various biases. I have also outlined some of the alternative approaches that workers are turning to, to overcome these biases and derive more accurate diversity estimates. To end I want to highlight two aspects of the marine Phanerozoic diversity curve that are currently a matter of considerable debate: has marine diversity increased significantly over the last 100 Ma and what is driving cycles of diversity?

**Has taxonomic diversity risen steeply since the Mesozoic?**

One of the most striking features of the Phanerozoic marine diversity curves of Sepkoski and Benton is the dramatic rise in diversity that starts in the late Cretaceous and continues to the present day. This trend is apparent in all compilations that have used traditional methods of taxon counting (e.g. Fig. 1) and at all taxonomic levels, although it is at its most pronounced for genus-level data (Lane & Benton 2003). Furthermore, such a rise intuitively makes sense. Many important groups appear to have diversified greatly since the start of the Mesozoic, with origination greatly outstripping any extinction (actinopterygian fishes, echinoids and starfishes, decapod crustaceans, corals, cheilostome bryozoans, neogastropods). Indeed, it is striking how the diversity curve tracks the rock record curve throughout most of the Phanerozoic, but starts to diverge from the Late Cretaceous onwards very markedly (Fig. 6). No other time interval shows anything comparable. On the other hand, some analyses, after taking sampling into account, have suggested that diversity was either static or declining through the Cenozoic (Alroy et al. 2001; Peters & Foote 2001; Crampton et al. 2006b).

There are several potential artefacts that could be creating this Cenozoic rise in diversity. Probably the most significant of these is what Raup (1972, 1976) dubbed ‘the pull of the Recent’. Traditional analyses, such as those of Sepkoski and Benton, use...
range interpolation to compensate for gaps in the fossil record. However, if there is one level (the Recent) where the biota is completely known (more or less) this will skew diversity counts. The closer in time we approach the Recent, the more taxa will have their ranges extended forward from their last fossil record to the Recent. Extinct taxa cannot benefit from this additional correction for missing record so the net result is an increase in the ratio of observed to inferred taxa towards the present day.

One way to ameliorate the problem of the ‘pull of the Recent’ is to omit all Recent records so that taxon counts extend only to their latest fossil occurrence. A preliminary study by Sepkoski (1997) using this approach suggested that there might indeed be a small effect, but that the rise in overall diversity during the Cenozoic was not an artefact. Peters & Foote (2001) were sceptical of this result and re-examined Sepkoski’s genus-level data, removing all Recent taxa not recorded from the Plio-Pleistocene. They estimated that about 50% of extant genera with a fossil record were missing from rocks of the last 5 Ma. If true, this would be a major distorting bias, as it would greatly inflate estimates of sampled diversity in these youngest time intervals.

Modelling by Peters & Foote (2001) suggested that the rise in diversity since the late Miocene was probably overestimated four- to eight-fold.

However, Jablonski et al. (2003) took a detailed look at Sepkoski’s molluscan generic database and found that only a small number of extant bivalves with a fossil record had no Plio-Pleistocene fossil record. The reason for this marked discrepancy was because Sepkoski’s data were not standardized or updated but simply took taxonomic records as originally published. Jablonski et al. corrected specific taxonomic assignments and found that only 9% of extant taxa with a fossil record were not present in the Pliocene. Additional records from sources not consulted by Sepkoski reduced this figure to 5.4%. Thus the ‘pull of the Recent’ may not be such a problem if care is taken over taxonomic assignments.

Two additional factors might also act to artificially elevate marine diversity towards the Recent.

(1) There may still be an overall trend towards a better fossil record in younger rocks, simply because fossils are better preserved and easier to collect from younger, less lithified sediments. This seems to be the case in the most detailed analysis of sampling bias yet undertaken, that by Crampton et al. (2006a). They found that the probability of sampling fossil molluscs increased through the Cenozoic by between 20 and 50%, most markedly in the Pliocene and Pleistocene (Fig. 5). They ascribed this to the greater ease of recovery of fossils from younger strata as well as to increasing surface outcrop area preserved.

(2) Cenozoic turnover rates are low compared with the Mesozoic and Palaeozoic and time intervals comparatively shorter, so a greater proportion of genera extend across several time intervals (nearly 80% survive through each Cenozoic epoch; Foote 2000). Consequently, Peters & Foote (2001) estimated that this could create a two-fold increase in genus diversity for the Cenozoic, and would have greatest effect on the Plio-Pleistocene, where durations are the shortest.

Countering these two biases are four biases that act in the opposite direction, leading to underestimation of global marine diversity towards the Recent.

(1) All measures of rock bias in the Phanerozoic (Peters & Foote 2001; Smith 2001; Crampton et al. 2003; Peters 2005) have demonstrated a marked reduction in outcrop area and/or sampled habitat heterogeneity through the Cenozoic (Figs 3–5). Thus even if sampled diversity levels rose slowly (e.g. Alroy et al. 2001) or remained more or less uniform over time (Crampton et al. 2006b), this trend is maintained in the face of decreasing numbers of fossil localities and/or habitat heterogeneity. Decreased rock outcrop must surely ameliorate, to a large extent, any increase in the quality of preservation and ease of collection that also occurs towards the present.

(2) The average number of species assigned to a family has more or less doubled since the Jurassic (Fig. 7; Flessa & Jablonski 1985), so that counts of higher taxa will progressively underestimate species diversity towards the Recent. This could underestimate Neogene species diversity compared with mid-Jurassic species diversity by as much as a factor of two.

(3) There has been a shift of the well-studied parts of the world away from the tropical belt during the Cenozoic (Fig. 2). Estimates of diversity in which the tropical fauna is progressively undersampled would be expected to result in lower sampled diversity towards the present (Jackson & Johnson 2001; Bush & Bambach 2004).

(4) There may be a large-scale bias against aragonitic shell preservation towards the Recent as a result of increased depth of bioturbation (Bush & Bambach 2004), causing a one- to two-fold reduction in diversity (but see discussion above).

The preliminary analyses by Alroy et al. (2001) of their database of biotic inventories using various methods of standardization found comparable diversity levels in the Palaeozoic and Mesozoic–Palaeogene, and evidence of only a modest rise in diversity between the end of the Middle Jurassic and the end of the Palaeogene. They concluded that there was ‘only a weak Mesozoic–Cenozoic radiation’, much smaller than had been predicted from raw taxon counts. However, Bush et al. (2004) showed that one of the techniques being used was highly unreliable, and that the other methods actually produced results that were comparable with the Sepkoski curve (Fig. 8).

Diversity has three components, alpha diversity (diversity within communities), beta diversity (diversity added by different communities in a region), and gamma diversity (diversity added by different geographical regions) (Whittaker 1972). Bush & Bambach (2004) compared alpha diversity in the Palaeozoic (Late Ordovician to Late Devonian), with communities in the Late Miocene to Pleistocene. After correcting for various possible biases (changing palaeolatitude, aragonitic dissolution, and inequality in the proportional representation of onshore–offshore communities) they concluded that alpha diversity had increased by a factor of about 3–4 since the late Devonian, broadly comparable with the results of Sepkoski et al. (1981). Kowalewski et al. (2006) did the same for Jurassic and Neogene communities and also found that alpha diversity had increased, although not by as much as expected. These observations are important because, for diversity to remain static or decline in the face of rising alpha diversity, beta or gamma diversity must have declined. A decline in beta diversity could come about through progressively restricting the set of communities available for palaeontologists to sample at outcrop; for example, through progressive loss of deposits representing mid- to outer shelf on land as sea levels fell through the Cenozoic.

In summary, there are good reasons for believing that the practice of through ranging between first and last occurrence, and the increase in the quality of the fossil record towards the present could be responsible for overestimating the rise in sampled marine diversity during the Cenozoic. However, there are also factors that bias the curve in the opposite direction (i.e. underestimation of the rise in diversity), and the most powerful biasing factor towards overestimation, the ‘pull of the Recent’, has been demonstrated to be less of a problem than at first
suspected. Habitat sampling is a key issue here because changing sea levels change the range of habitats and communities preserved on land for palaeontologists to sample; this effect could be artificially driving beta diversity and cannot easily be corrected for by undertaking more fieldwork. Because there are so many biases that need to be taken into account, it is not yet clear whether the rise in marine diversity towards the Recent is genuine or artificial.

Cycles of extinction or cycles of non-preservation?
The idea that taxonomic diversity might rise and fall episodically goes back to the catastrophists of the early 19th century. However, belief in a more regular pattern to extinction and origination is much more recent. Ager (1977) hinted at a 60 Ma cycle but thought it probably fortuitous, Thompson (1977) showed a c. 30 Ma cycle in vertebrate diversity, and Raup & Sepkoski (1984, 1986; see also Rampino & Haggerty 1995) advanced evidence of a 26 Ma cycle of extinction in marine invertebrates. The latest resurrection of this is that of Rhode & Muller (2005), who used Fourier spectral analysis to argue that the genus-level Phanerozoic marine diversity curve has a strong 62 Ma cyclicity (Fig. 4b). However, like all previous claims for cyclicity, Rhode & Muller’s 62 Ma periodicity turns out to be an artefact of data manipulation and the analytical method used (Omerbashich 2006).

The possibility that marine diversity has risen and fallen repeatedly on a regular basis over the Phanerozoic has, nevertheless, intrigued palaeontologists. Although not strictly periodic, the regularity of this pattern suggests that marine diversity, rather than being the result of a complex, often haphazard set of biological and physical chance interactions, might be driven by a single recurring extrinsic factor. Furthermore, the long time scale of these cycles of diversity implies a major tectonic or extraterrestrial driver of change. Rhode & Muller (2005), like others before them, sought to identify the driver by seeking to match the rises and falls in diversity against various geological and astronomical signatures, including $\delta^{18}$O and $\delta^{13}$C (as a proxy for global climate change and biomass), time of emplacement of large igneous provinces, sea-level curves, ages of impact crater, and number of rock formations. Two of these showed weak correlation with marine diversity (emplacement of large igneous provinces and the Exxon sea-level curve (Haq et al. 1987)) but the correlations were neither statistically robust nor convincing. Subsequently, Smith & McGowan (2005) demonstrated a highly significant correlation between rock outcrop area and genus diversity for the post-Palaeozoic (Fig. 3c). They showed that more than half the variance in genus-level diversity is explained by the amount of rock at outcrop. A similar close match is seen in the paper by Peters (2005), where a more refined database of stratigraphic sections was provided (Fig. 6). Basically, highest diversity is observed at times of widespread sediment deposition, whereas lowest diversity coincides with system bases when marine sediment at outcrop is at its lowest. Critically, peaks in the number of last occurrences of taxa mostly coincide with the transition from time intervals with a good marine rock record to those with a poor rock record (Smith 2001; Peters & Foote 2002, figs 3–5).

It is now clear that sampled marine diversity, as measured by taxon counts at genus, family and presumably species levels, fluctuates regularly through time. This recurrent signal is unlikely to have arisen simply by chance (e.g. Sepkoski 1989; Benton 1999), so some forcing mechanism seems to be involved. Furthermore, the high statistical correlation between amount of rock at outcrop and marine diversity rules out any of the postulated extraterrestrial drivers (periodic passage of the Solar System through space, oscillation of the Sun through the galactic plane, solar activity cycles, companion star effects or planet X; see Rhode & Muller (2005) for a succinct summary), as there is no plausible mechanism to link any of these extraterrestrial drivers to rock preservation at outcrop. There are, on the other hand, several potential geotectonic processes that work on the time scale required and that affect rock record through sea-level change (Dewey & Pitman 1998; Miller et al. 2005).

Changes in rock record, as measured by map outcrop area (Smith 2001; Smith & McGowan 2005), number of named formations (Peters & Foote 2001, 2002) or onset of stratigraphic gaps (Peters 2005, 2006a, b) correspond to second- and third-order sequence stratigraphic cycles (Figs 3–5). Consequently, they are best explained as a direct response to major cycles of craton flooding and draining. Sloss (1963) first recognized that major regional unconformities on the shelf (surfaces of erosion and non-deposition) divide the stratigraphic record into sequences and provide the key to understanding eustatic change.
The sedimentary record ultimately preserved at outcrop during one of these cycles will depend upon the interplay of total surface area inundated by the sea, and length of time and severity of erosion following retreat of the sea. The amount of marine rock preserved should be greatest around the maximum flooding surface, when the continents are most extensively inundated, and will be at its lowest at around system base. However, the flooding record of continents is not a direct measure of eustatic change because variations in subsidence and sediment supply also influence shoreline location (Dewey & Pitman 1998), and tectonically active regions may differ somewhat from stable cratons in their response (Crampton et al. 2006a).

There are three major drivers that control global sea-level curves on time scales of tens of millions of years (Dewey & Pitman 1998; Miller et al. 2005): (1) emplacement of large igneous provinces in the form of oceanic plateaux, which produces fast rises in sea level (at rates in the region of 60 m Ma\(^{-1}\)) and slow falls as a result of thermal subsidence (10 m Ma\(^{-1}\)); (2) changes in ocean crust production primarily affected by variations in sea-floor spreading rates or ocean ridge lengths (100–300 m amplitudes at rates of 10 m Ma\(^{-1}\)); (3) continental collision and crustal shortening (amplitude 50 m at rates of up to 1 m Ma\(^{-1}\)). All three drivers are ultimately related to mantle plume activity; increased plume activity results in elevated sea-floor spreading rates and the emplacement of major volcanic provinces, all of which reduce ocean basin volume and force seas to rise and flood over the continents. Times of reduced plume activity result in ocean basin deepening, and general regression of seas from the continental blocks. In addition, variations in sedimentation rates can cause slow changes (60 m amplitude changes at rates of 10 m Ma\(^{-1}\)). Although glacial cycles can trigger massive changes in sea level they act over much shorter time scales than those of interest here (amplitudes of 200 m at rates of 20 m ka\(^{-1}\)).

The well-known c. 300 Ma Wilson cycle of continental accretion and fragmentation (Fischer & Arthur 1977), presumably linked to superplume development, is responsible for the two first-order sequence stratigraphic cycles of the Phanerozoic (Dewey & Pitman 1998, fig. 2) and is clearly expressed in the rock record by a major shift in the proportion of terrestrial to marine sediments (unpublished data) and in continental flooding (Fig. 2b). Thermal uplift, prior to rifting and break-up of a supercontinental block, causes sea level to drop and terrestrial sediment to be deposited across those areas of the continent that can today be sampled by geologists at outcrop. Sedimentation in shallow marine habitats still takes place, but over more limited areas on outer continental shelves or even upper slopes, areas not currently accessible to field palaeontologists. With plate rifting and dismemberment, continental trailing edges develop, which, as they sink, provide accommodation space for marine sediments to build through time. Maximum flooding is expected to coincide with maximum continental dispersal. Finally, towards the end of the cycle, as plates converge and coalesce there is net destruction of marine sediments through subduction, shortening and flexion of continental blocks at collision margins, all resulting in sea-level fall and enhanced erosion of the most recently deposited sediments.

Smaller-scale cycles of continental flooding and draining occurring over a time scale of tens of millions of years are superimposed over this primary cycle and consequently have different signatures depending upon where each falls in the first-order cycle. When first- and second-order flooding surfaces coincide there is a breakdown of shelf-edge fronts and stratified, oligotrophic oceanic conditions spread widely over the shelf, as in the early Late Cretaceous (Gale et al. 2000). Conversely, the degree to which terrestrial sedimentation dominates at lowstands of lower-order cycles will be less as first-order cycles approach their maximum flooding surface than towards their start or end. The driver of these secondary cycles is less clear and possibly not a single factor. As the rate of ridge production has remained more or less uniform over the last 180 Ma (Rowley 2002), this cannot be responsible for driving the sea-level cycles. However, large igneous province emplacement possibly shows strong periodicity over parts of the Phanerozoic at about the correct wavelength (Prokoph et al. 2004). Furthermore, there is prima facie evidence for a link between the emplacement of major oceanic basaltic provinces and the onset of rapid sea-level rise, as, for example, in the Cenomanian (Kerr 2005). Evidence for cycles of sediment erosion comes from stronitum, carbon and oxygen isotopes and shows a 30 Ma cyclicity (Prokoph & Veizer 1999). In summary, a combination of plate assembly and disassembly on a time scale of c. 300 Ma, overlain by a second-order c. 30–100 Ma cycle of continental flooding and regression driven by changing ocean basin volume can best explain the cycles in rock record that have been observed. But can these also explain cycles in diversity?

What causes the link between sampled diversity and rock record?

The connections between the primarily tectonic processes that alter sea level and marine diversity are complex (Fig. 9). Flooding of cratons increases the surface area of shallow marine shelf that organisms can inhabit, while also increasing the surface area over which marine sediments are deposited and ultimately become incorporated into the rock record. Regression reduces habitable shelf area and reduces the volume and surface area of previously deposited marine sediments from the rock record by erosion. Consequently, therefore, there are two competing and non-exclusive hypotheses that would explain the link between rock outcrop area and marine diversity: a biological and a sampling hypothesis (Peters & Foote 2001, 2002; Smith 2001; Peters 2005, 2006a).

Biological hypothesis. Diversity changes through speciation and extinction in response to increases and decreases in the surface area of shallow continental shelves. Shallow-water habitats (within storm wave base) are nutrient rich because of vertical mixing and/or lie within the ecologically important photic zone. They therefore support large and heterogeneous marine communities. As sea levels rise and seas start to spill across the continental shelves marine life diversifies to occupy the ever-increasing area of shallow-water habitats. Biodiversity later decreases as continental shelf area covered by sea decreases through competition and extinction. Biodiversity also decreases at times of maximum flooding because that is when the waters covering large parts of the shelf become stratified as the sea floor becomes too deep to allow adequate nutrient recycling through storm mixing, and open ocean, oligotrophic conditions spread across much of the continents.

Sampling hypothesis. Diversity changes only slowly in marine settings, as communities migrate across the shelf to keep pace with slowly changing sea levels, but the amount of rock that is preserved at times of continental flooding is much greater than at times of regression. The diversity of taxa sampled in any time interval is simply controlled by the amount of rock at outcrop that remains and the proportions of shallow- to deep-water
Phanerozoic and at the level of individual stages. Furthermore, he showed that the proportion of genera having their last occurrence in a stage was not correlated with the duration of time encompassed by the overlying gap. Because the sampling hypothesis predicts that the proportion of genera having their last occurrence should increase as the duration of the overlying gap increases, and genus duration should be independent of sequence duration, Peters concluded that genus diversity was intimately tied to habitat continuity within local depositional basins; that is to say, that sea-level change drove both rock record and biological diversity independently. He also showed that rates of genus origination and extinction were positively correlated with the areal extent of hiatuses.

Peters (2005, 2006a) makes a strong case for why the match between marine biodiversity and rock record might be more biological than sampling artefact in origin. However, there are reasons for questioning his conclusions. First, Peters showed that the average duration of a stratigraphic gap is around 100 Ma whereas the average duration of a genus and a sequence is about 25 Ma. Therefore few if any genera are expected to span gaps in individual sections, making the duration of the hiatus academic to local survivorship (as demonstrated by the lack of correlation). Second, Peters ignored the problem of pseudoextinction, which is estimated to affect about one-third of genera in well-studied groups (Smith & Patterson 1988) and probably more in stratigraphically important groups such as ammonites and graptolites.

The disappearance of a genus along with marine sedimentation in a region is no guarantee that it has not survived somewhere else, only to be given a new name by taxonomists when its descendents reappear at a later time. Indeed, taxonomists are very likely to have used gaps in the rock (and fossil) record as convenient and objective points to subdivide otherwise continuously evolving lineages. Marine conditions continued uninterrupted throughout the Phanerozoic whether or not our biased sample of rock deposits captures this. Therefore the correspondence between average section duration and average genus duration could equally well be explained as a reflection of taxonomic practice if pseudoextinctions are clustered around peak occurrences of gaps in the stratigraphic record. That pseudoextinction might have a sizeable effect is suggested from gap analysis at higher taxonomic level (Paul 1998), where it is clear that time intervals when there are peaks of last occurrence are followed by time intervals when there are elevated proportions of missing fossil record.

Today the world’s continental marine shelves represent the most productive marine ecosystems but cover only about $28 \times 10^6$ km$^2$ or 0.1% of the Earth’s surface (Carlton et al. 1999). Therefore one might expect major changes in sea level that alter marine shelf area to be a driver of extinction. We can test this assumption by turning to the Pleistocene fossil record. Pleistocene glacial cycles had a profound effect on sea level. For example, the last glacial maximum at around 20 ka ago caused sea levels to drop rapidly ($1 \text{ cm} \text{a}^{-1}$) to C. $120 \pm 5 \text{ m}$ below that of today (Miller et al. 2005). This shifted sea level close to the continental shelf-edge, and did so an order of magnitude faster than any tectonically driven sea-level fall. The net result was a dramatic reduction in the area of shallow water marine habitats, with those lying within the photic zone, the most productive areas of all, shrinking to about 20% of their former extent (Schaaf 1996). Yet Valentine et al. (2006) have shown that bivalve genera were remarkably immune to Pleistocene sea-level changes. Although there are major shifts in geographical range (Valentine & Jablonski 1991), extinction at genus level was less than 6%. Regional extinction was maybe three to four times

habitats sampled. With less rock preserved, a smaller range of habitats will be sampled and fewer rare taxa discovered.

Of course, the truth probably lies somewhere between, with both biological and sampling factors playing their part. But is it possible to tell which has the upper hand in any one cycle? There seems little doubt that the drop in diversity associated with the continental accretion and formation of Pangaea at the end of the Permian was real. It resulted in a fundamental reorganization of marine community structure and the concomitant loss of major clades (Erwin 2006; Wagner et al. 2006) on a scale seen at no other time interval except possibly at the end of the Cretaceous (Sepkoski 1981, 1997; Peters 2005). However, there is much less certainty about whether marine diversity during other time intervals really fluctuated significantly, despite all the claims made for mass extinctions (e.g. Jablonski 2005).

To distinguish between the biological and sampling hypotheses, Peters (2005, 2006a) compared the duration of genera with the duration and timing of unconformity-bounded sedimentary sequences, based on over 4000 sections at 541 North American localities. He found that there was a strong and statistically significant match between genus duration in the fossil record and sequence duration in sections, both averaged over the entire sequence.

Fig. 9. Schematic flow chart showing the two ways in which tectonic and climatic drivers can alter sampled fossil diversity through sea-level change.
higher in high-latitude, isolated island regions, where the effects of glaciation were strong and latitudinal migration impossible (e.g. New Zealand: Crampton et al. 2006b). It is hard to think of a more extreme situation than having sea level drop to the continental shelf-edge so rapidly, and yet diversity at genus level changed only modestly. This and several other lines of evidence suggest that regression or transgression by itself cannot generate high levels of extinction in marine populations (see Gale et al. 2000; Smith et al. 2001). Other factors, such as perturbation of primary productivity in the oceans, must be critical.

The concept that mass extinctions profoundly affected the history of life on Earth through the Phanerozoic dominated the last decade and a half of palaeontological research in the 20th century. Yet much of the evidence for such events (and their supposed cyclicity) came from taxon sampling that was neither standardized nor took account of variation in the rock record.

There are, nevertheless, two time intervals when extinction is greater than predicted even after preservational and rock record biases (but not pseudoextinction or habitat biases) are taken into account. These occur at the Permo-Triassic and Cretaceous—Cenozoic boundaries and stand as evidence that life has been affected by extraordinary events. Other postulated times of mass extinction are much less convincing, as they occur at times when the rock record at outcrop was also becoming markedly poorer and when pseudoextinction is likely to be most rife. This means that their precise duration, timing and even existence remain ambiguous (Foote 2003, 2005).

**Summary and conclusions**

The fossil record provides our only realistic means of estimating how diversity has changed through time. Currently, however, our understanding of Phanerozoic marine diversity patterns is in a state of flux. We know that past attempts to measure biodiversity through time are probably inaccurate because of several major systematic biases, although we are still unsure as to how profound an effect these will turn out to have had in distorting our view. The primary biases that blight current efforts to estimate Phanerozoic marine diversity are: (1) uneven sampling effort of rocks at outcrop both geographically and over time; (2) a cyclical systematic bias in the rock record from which we have to make our fossil collections, driven ultimately by episodes of tectonic activity; (3) taxonomic inconsistency, especially associated with interpreting the meaning of last occurrence data. More sophisticated methods of estimating diversity patterns are now being developed that can take sampling and rock record bias into account and are already challenging some long-held views. These approaches need to be combined with a deeper appreciation of the nature of taxonomic data before we arrive at a definitive view of how marine diversity has changed over time.

It is possible that the Phanerozoic marine diversity curve will turn out to be very different from the one prevalent just a few years ago. The relentless rise in marine diversity through the late Cretaceous and Cenozoic may still be there, although once sampling artefact has been properly accounted for it will probably be seen to be flattening out over the last 60 Ma. More importantly, many of the marked stage-to-stage changes in diversity present in unstandardized data, and that provided the ‘bread and butter’ for ‘mass extinction’ hypotheses during the 1980s and 1990s, may be no more than artefact. Only the end-Permian and end-Cretaceous events currently seem secure. Even if there are other periods of elevated extinction that cannot be explained solely by sampling bias (Foote 2003), there is growing evidence that changes in diversity are principally the result of slow-acting tectonic events that control the temporal continuity of the geological record (Peters 2005), not catastrophes. And so we seem to be returning more or less full circle to the views of Newell (1952), who first produced cogent evidence in support of cyclical sea-level change as the primary driver of the marine fossil record. The critical question facing us now is whether it does this biologically, through altering the ecological landscape of shallow marine seas, or indirectly as a consequence of the volume of sedimentary deposits that enter the rock record.

I am grateful to K. Johnson, A. McGowan, N. MacLeod, P. Taylor, R. Fortey, M. Benton and D. McIlroy for discussion and helpful reviews of an earlier version of this paper. This work was supported by a grant from the Leverhulme Trust.
1 the fossil record of echinoids in the Triassic and early Jurassic using sampling data, phylogenetic analysis and molecular clocks. Paleobiology, 00, 000–000.


Received 30 November 2006; revised typescript accepted 23 February 2007.

Scientific editing by Duncan McIlroy

1. Spelling Norrell not as in ref. list. Which is the correct version?
2. Carlton et al. date not as in ref. list. Which is the correct year?
3. Alroy et al. - please give the first three authors before et al.
4. Cherns & Wright - please check the page range.
5. Dewey & Pitman - please give editor's.
7. Smith 2007 - please update if possible.