The evolution of diversity in ancient ecosystems: a review

S. Conway Morris
Department of Earth Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EQ, UK

On a perfect planet, such as might be acceptable to a physicist, one might predict that from its origin the diversity of life would grow exponentially until the carrying capacity, however defined, was reached. The fossil record of the Earth, however, tells a very different story. One of the most striking aspects of this record is the apparent evolutionary longueur, marked by the Precambrian record of prokaryotes and primitive euukaryotes, although our estimates of microbial diversity may be seriously incomplete. Subsequently there were various dramatic increases in diversity, including the Cambrian 'explosion' and the radiation of Palaeozoic-style faunas in the Ordovician. The causes of these events are far from resolved. It has also long been appreciated that the history of diversity has been punctuated by important extinctions. The subtleties and nuances of extinction as well as the survival of particular clades have to date, however, received relatively little attention, and there is still a tendency towards blanket assertions rather than a dissection of these extraordinary events. In addition, some but perhaps not all mass extinctions are characterized by long lag-times of recovery, which may reflect the slowing waning of extrinsic forcing factors or alternatively the incoherence associated with biological reassembly of stable ecosystems. The intervening periods between the identified mass extinctions may be less stable and benign than popularly thought, and in particular the frequency of extraterrestrial impacts leads to predictions of recurrent disturbance on time-scales significantly shorter than the intervals separating the largest extinction events. Even at times of quietude it is far from clear whether biological communities enjoy stability and interlocked stasis or are dynamically reconstituted at regular intervals. Finally, can we yet rely on the present depictions of the rise and falls in the levels of ancient diversity? Existing data is almost entirely based on Linnean taxa, and the application of phylogenetic systematics to this problem is still in its infancy. Not only that, but even more intriguingly the pronounced divergence in estimates of origination times of groups as diverse as angiosperms, diatoms and mammals in terms of the fossil record as against molecular data point to the possibilities of protracted intervals of geological time with a cryptic diversity. If this is correct, and there are alternative explanations, then some of the mystery of adaptive radiations may be dispelled, in as much as the assembly of key features in the stem groups could be placed in a gradualistic framework of local adaptive response punctuated by intervals of opportunity.

Keywords: diversity; evolution; adaptive radiations; palaeontology; molecular clock; extinctions

1. INTRODUCTION

It is John Phillips (1860) who is generally credited with first plotting a history of diversity as revealed by the fossil record (figure 1, upper left). Although schematic, his diagram has several features of interest. It depicts the triple peaks of Palaeozoic, Mesozoic and Cenozoic diversity, divided by the caesuras of the end-Permian and end-Cretaceous events. Interestingly, the respective dips in the highs of Palaeozoic and Mesozoic diversity were attributed to taphonomic failure, a problem that still dogs any analysis of ancient diversity. Phillips (1860) also recognized the precipitous increase in Cenozoic diversity. Thus, within a year of the first edition of Darwin's *Origin of species* the basic information on the history of organic diversity was available, with Phillips (1860) not only drawing attention to taphonomic lacunae but also in the pages preceding his figure 4 making some preliminary attempts to quantify his analysis.

It took more than 100 years, however, for serious interest in this history to be rekindled. First were the pioneering formulations of Newell (1967), Valentine (1969, 1973) and Tappan (1969), but the subsequent enterprise has been strongly influenced by the massive compilations, in part unpublished, by Sepkoski (1992) (figure 1) and an important series of analytical papers (Sepkoski 1978, 1979, 1981, 1988, 1993; see, also, Sepkoski 1997). The analysis of diversity and the resurgence of interest in mass extinctions remain a principal focus for palaeobiology (see, for example, Signor 1990; Benton 1990; Benton 1990; Valentine 1990; Erwin 1996a; Foote 1996), and has reached a certain stage of maturity. In addition to analysis of marine invertebrate diversity, comparable compilations for land plants (Niklas *et al.* 1983) and non-marine tetrapods (Benton 1985) are also available. Earlier concerns that the patterns of diversity were largely artefacts, controlled by such factors as area or volume of fossiliferous sediment, diversity 'hot-spots' centred on monographic treatment and...
fossil Lagerstätten, and a bias in favour of the preservation of younger sediments (the ‘Pull of the Recent’) have been quelled if not dismissed (see Sepkoski et al. 1981). Even so, the shape of the curve of organic diversity remains contentious. Thus Benton (1995), employing the data from his edited compilation of Fossil record 2 (Benton 1993), identified in general an exponential increase in Phanerozoic diversity. This, however, was refuted by the critical reanalysis of Courtillot & Gaudemer (1996) who reaffirmed the basic pattern identified earlier by J. J. Sepkoski of diversity curves governed by logistical growth and interrupted on occasion by mass extinctions of varying severity.

There is a continuing shift in research emphasis in universities and happily to a lesser extent in natural history museums away from primary systematics and the monographic treatment of fossil groups. It is perhaps questionable, therefore, whether any updated databases produced in the next few years based on the precepts of Linnean taxonomy of generic or familial diversity will lead to a radical revision of current depictions of the histories of diversity, of which marine life remains the best known (figure 1). The only exception to this trend in data accumulation might appear to be the substantial growth in the study of fossil Lagerstätten, especially those displaying soft-part preservation. Paradoxically, in recent years, the research programme in this area has been largely on the taphonomy of Lagerstätten and the possible reasons for exceptional preservations. The documentation of primary systematics of many of these extraordinarily well-preserved biotas has lagged far behind.

The reader will realize that this review makes no pretence to completeness. So far as it has an aim, it is to touch on areas that perhaps have suffered from either relative neglect or are particularly topical. To forestall justifiable criticism I should say that among the many potential topics that will not be considered here are: the problem of selectivity in extinctions (see, for example, Bennett & Owens 1997), apart from some passing comments at the end of this paper any detailed discussion of the mathematical treatment of diversity (see, for example, Patterson & Fowler 1996; Solé & Bascompte 1996; Solé et al. 1996); the questions of stratigraphic completeness and the sampling of the fossil record (see, for example, Foote & Raup 1996; Marshall 1997); variable species longevity (see, for example, Levinton & Ginzburg 1984; Kammer et al. 1997); biogeographical changes engendered by continental splitting; the influence of seawater chemistry (Grotzinger 1990; Harper et al. 1997); the patterns of onshore–offshore diversity (see, for example, Sepkoski 1991); the merits (or otherwise) of Linnean taxa versus cladistic methodologies (see, for example, Sepkoski & Kendrick 1993; Foote 1996); the stability of taxonomic concepts (Hughes & Labandeira 1995); and the important role of refugia as ‘ecological bunkers’ in times of adversity and also ‘evolutionary museums’ of archaic diversity (see Oji 1996). Finally, no serious attempt is made to cover the influence of climate on diversity in terms of either originations (see Cronin 1985), extinctions (Clarke 1993; Coope 1987, 1994), or faunal replacements (Janis 1989; Jackson 1994). At the risk of an eclectic journey the reader is invited to continue.

2. ANCIENT DIVERSITY: POTENTIAL PITFALLS

In assessing ancient diversity there appear to be two fundamental problems, which need to be reviewed before I comment, with equal brevity, on four particular topics (§3–6). The first problem concerns the divergence in approach between those who seek regularities and, if not laws, at least a degree of predictability in the behaviour and fate of ancient ecosystems. Whereas the antithesis is by no means absolute, an alternative view prefers to emphasize the contingencies of history and often by implication the lack of discernible trends or patterns.

At first sight the attempt to seek regularities in the history of diversity appears to be frustrated by the obvious changes in the biosphere in the last 4 billion years (Ba). Most obviously this applies to the Archaean (3.8–2.5 Ba), when ecosystems may have been only occupied by prokaryotes (see Schopf 1983). The emergence of eukaryotes, which in terms of the fossil record apparently can be traced back to about 2.1 Ba (Han & Runnegar 1992; but, see Doolittle et al. 1996), presumably defined the onset of more complex ecosystems, but direct evidence for significant reorganization of microbial communities is at present effectively restricted to two areas. The first is the documentation in the changes in the diversity of the microbially generated structures known as stromatolites, and in particular a dramatic decline in their diversity in the later Proterozoic (Walter & Heys 1985). This decline is widely attributed to the rise of grazing metazoans (Garrett 1970; Avramik 1971; Walter & Heys 1985), but abiotic explanations such as changes in the carbonate saturation of sea water seem as reasonable (Grotzinger 1990). There is a further complication in as much as the relationships between macroscopic expression of the stromatolite and the complexity or otherwise of the microbial communities remain obscure. The second area where an understanding of Precambrian diversity has advanced

Figure 1. The diversity curve of marine metazoan families, and their distinction into the three great evolutionary faunas, identified by factor analysis, and referred to as Cambrian (I), Palaeozoic (II) and Mesozoic–Cenozoic or Modern (III). The residual area above the main curves corresponds to that part of the diversity that cannot be accommodated by the three main statistical factors. The figure of 1750 (right) is the estimated number of marine families in Recent oceans, many of which have a low fossilization potential. The smaller diversity curve inserted top-left is redrawn from Phillips (1860). (Main figure is redrawn from Sepkoski (1981, figure 5). Reprinted with permission from Paleobiology and the Palaeontological Society.)

Phil. Trans. R. Soc. Lond. B (1998)
considerably is the depiction of protistan diversity, especially in terms of acritarchs (Knoll 1994). Even with the appearance of the first animals, whose origins may be deeper in the Proterozoic than the fossil record at present indicates (Wray et al. 1996; see, also, Conway Morris 1997), the ecology of the planet seems to have operated at an effectively microbial level for at least 3 Ba.

In this regard the subsequent Cambrian ‘explosion’ (ca. 545–520 Ma) remains one of the fulcrum points in the history of diversity with the origin and radiation of macroscopic metazoans. It is, nevertheless, a matter of contention as to the extent to which the Cambrian radiation is an ‘explosion’ of fossils whose newly acquired skeletal hard-parts breached a taphonomic threshold as against a genuine revolution in diversity. This potential dilemma was explored by Runnegar (1982a), but the recent reviews of the fossil record in this interval (for examples, see Lipps & Signor 1992; Brasier 1992; Conway Morris 1998b) treat the Cambrian ‘explosion’ as an evolutionary event of the first magnitude. That the trace fossils show a parallel increase in diversity (see Jensen 1997) is a strong argument in favour of the latter view, although it would still be possible to argue that the Cambrian ‘explosion’ is more the result of a scaling-up in size, perhaps because of changes in the concentration of atmospheric oxygen, rather than a genuine series of innovations. The problem of cryptic diversity is returned to in a later section (§6), but here we need only note that at least so far as the immediately preceding Ediacaran faunas are concerned, evidence continues to grow for phyletic links with life in the Cambrian (for examples, see Gehling & Rigby 1996; Conway Morris 1993a; Fortey et al. 1996; Gehling & Rigby 1996; Waggoner 1996; Fedonkin & Wagggoner 1997).

The impact of the Cambrian ‘explosion’ in other ways is perhaps still less fully appreciated. For example, Logan et al. (1995) drew attention to a significant shift in the ratio of carbon isotopes (δ13C) in certain hydrocarbons extracted from sediments spanning the Proterozoic–Phanerozoic boundary. This they interpreted as a consequence of the evolution of planktonic grazers, possibly arthropods (see Butterfield 1994, 1997) adept at harvesting suspended organic matter which, after passage through the gut, was encapsulated into faecal pellets which then rapidly descended to the sea floor a kilometre or more below. Before the evolution of these grazers the descent of organic matter through the water column would have been markedly slower because much of it would have been as minute particles of marine ‘snow’. The transition from a microbial ecosystem to one with metazoan harvesters probably had a profound effect on ocean chemistry, especially in terms of oxygen demand mediated by bacterial utilization of the slowly sinking particulate ‘snow’ (Logan et al. 1995).

The literature is replete with many other examples of subsequent changes in Phanerozoic diversity and their ecological consequences. Of these the invasions of the land (for examples, see Labandeira & Beall 1990; Kenrick & Crane 1997) and air (Crepet 1979; Feduccia 1980; Labandeira & Sepkoski 1993) perhaps are the most familiar, but others e.g. increasing degrees of sediment bioturbation (for examples, see Droser 1991; Droser & Bottjer 1993), taphonomic feedback of substrate types (Sprinkle & Guensburg 1993; Kidwell & Brenchley 1996), the recruitment to and the reorganization of planktonic ecosystems (see Rigby 1997) are also highly significant. Ironically, those workers who emphasize the many historical contingencies in the history of diversity are often particularly vociferous in their denial of what to others appears to be strong evidence for increases in the complexity (Valentine et al. 1994; see, also, McShea 1996) and sophistication of both biotas and ecosystems. Such trends seem to deserve the name of progress (Rosenzweig & McCord 1991; see, also, Jackson & McKinney 1990), which need not be equated with orthogenesis.

The second general problem concerns the question of scales of analysis in terms of geological time. A good example is that given by McGhee (1996) in his discussion of the late Devonian (Frasnian–Emsianen) mass extinction. The type and style of extinction differs according to whether the analysis is done at the level of epoch, stage, substage or subzone (figure 2). Which is the ‘true’ pattern, or at least the one with greatest explanatory value? The crux of this problem seems to be the problem of deciding whether the pattern of biotic change viewed on short time-scales ('ecological time') may be extrapolated seamlessly to the long-term intervals of ‘evolutionary time’, or whether emergent properties occur that would remain invisible to observers, such as ourselves, if a fossil record happened to be unavailable. Whereas strong claims have been made for such emergence, especially in terms of species selection (but, see Budd & Johnson 1991; Lieberman 1993) and mass extinctions (but, see Jackson & McKinney 1990), it is not yet clear the extent to which such principles actually govern the history of diversity. The main purpose of this paper is to review four specific areas to further our understanding of the changes in diversity.

3. THE PROBLEM OF RAPID DIVERSIFICATION

A survey of the marine fossil record (Sepkoski 1981, 1997) reveals three notable episodes of significant diversification: the early Cambrian, the mid Ordovician and the Cenozoico (figure 1). At present, the first of these, the Cambrian ‘explosion’ has attracted considerable attention, in part because of the realization of the significance of Burgess Shale-type faunas (Conway Morris 1998a) and the interest in the advances in the molecular biology of bodyplan development and its relevance to the origin of metazoan bodyplans (Conway Morris 1994a; Valentine et al. 1996; Erwin et al. 1997). The abruptness, or otherwise, of the Cambrian ‘explosion’ is predicated on both the possibility of an extended Proterozoic history and the phyletic relevance, if any, of the Ediacaran faunas. If present attempts to identify plausible metazoans in the Ediacaran assemblages (for examples, see Conway Morris 1993a; Waggoner 1996; Fedonkin & Wagggoner 1997) win favour, then the interval for this radiation might be extended to at least 30 Ma. Alternatively, even if the evidence for at least some Ediacaran fossils being unequivocal metazoans is accepted, this could still represent a radiation largely independent of the subsequent Cambrian event (Conway Morris 1993b). Some support for this idea may come from the record of Ediacaran trace fossils. Although a considerable diversity of Ediacaran ichnotaxa have been described, and some with peculiar morphology interpreted as representing ‘bizarre’ behaviours (Crimes
Figure 2. An illustration of the problems of the analysis of a geological event, specifically the late Devonian (Frasnian-Famenian) mass extinction, according to the resolution of the time-scale and stratigraphic unit chosen. The upper figures refer to familial diversity, respectively at the stratigraphic level of epoch (a) and stage (b). The lower figures illustrate the extinction as resolved at finer stratigraphic levels. (c) Illustrates the diversity of brachiopods in New York State and (d) the southern Urals of Russia, whereas (e) depicts events as seen as the subzonal level with reference to two important sections in Germany. (Figures are redrawn from McGhee (1996, figures 3.1, 3.2, 3.4 and 3.5 respectively). Reprinted with permission from *Paleobiology* and the Palaeontological Society (a,b), Canadian Society of Petroleum Geologists (c,d), GAGP (Göttingen) and E. Schindler (e).
1994), a reinvestigation of the record suggests this diversity has been exaggerated so that the Ediacaran record of trace fossils in fact was very impoverished (S. Jensen, personal communication). The revolution in benthic behaviour, and by implication neurological complexity, in the succeeding Cambrian would be consistent with the key steps in this radiation being achieved very rapidly, perhaps in less than 10 Ma. These uncertainties have an obvious bearing on the difficulties of identifying plausible triggers to explain the initiation of the Cambrian ‘explosion’, although the role of atmospheric oxygen or tectonic reconfiguration (Kirschvink et al. 1997) remain perennial favourites.

The Cambrian diversifications are of pivotal importance both for the establishment of a wide range of metazoan designs, codified in orthodox taxonomy as phyla, and a corresponding occupation of marine ecologies. Nevertheless, as has been long recognized diversification at the lower levels of the taxonomic hierarchy, such as the family, seems to have been more restricted. The figures for the number of genera and families will certainly require upward revision as our knowledge of Cambrian faunas, especially those preserved in Burgess Shale-type settings (Chen et al. 1996) and also by phosphatization (Müller & Hinz-Schallreuter 1993), continues to improve. Nevertheless, in comparison with the preceding Cambrian, the subsequent radiations in the Ordovician (figure 1) remain far more dramatic in terms of increase of diversity, notably of cnidarians, brachiopods, cephalopods, echinoderms and ectoprocts. Not only that, but the main increase in diversity appears to have been achieved in only a few million years, at least for a number of important shelly groups that included the trilobites, brachiopods and certain molluscs (Miller & Foote 1996). Although new evidence supports the biological distinctiveness of these emerging faunas (see Sprinkle & Guensburg 1995) and the pattern of diversification can be modelled with some detail, as yet we still have no convincing explanation for what may have initiated this Ordovician radiation. Thereafter, the diversity, at least of well-skeletonized marine families, appears to have stayed at a plateau, possibly representing an equilibrium, for about 200 Ma and weathering two significant mass extinctions at higher levels and miss phylogenetic links at traditional systematics may exaggerate the magnitude of species loss to planet-wide trauma respectively— are connected by a continuum of increasing disaster. Another influential view that has received wide support is Gould’s (1985) identification of tiers of influence for evolutionary process that transcend those operating at the level of the species and presumably are open to neodarwinian explanations. In this formulation mass extinctions are said to play a key role, in resetting ecologies and shaping the future development of post-catastrophe diversity (see, also, Courtillot & Gaudemer 1996). But are matters as simple as this?

The one instance that has received general agreement is the end-Pennsian event (Erwin 1993, 1995), marking the replacement of the Palaeozoic faunas by those of the Mesozoic. Given the severity of this episode (see Ershet et al., 1993), with extreme estimates of species loss reaching 96% (Raup 1979), it is not surprising that after the protracted time for recovery (see below) the marine communities were markedly different. Yet there is still some danger of exaggeration. Thus, once important groups of Palaeozoic marine organisms such as the trilobites (Fortey & Owens 1990) and rugose corals (Scrutton 1988) were already small components of even early Permian ecosystems. Other groups, such as the ammonoids and sea urchins, certainly suffered badly but subsequently rebounded to even greater levels of success. The possible pattern of extinction in the gastropods is particularly intriguing (Erwin 1996a; Erwin & Pan 1996). This is because the post-catastrophe faunas are markedly depauperate, yet a simple reading of the data may be misleading. One possibility is that absences are only apparent, and reflect a taphonomic lacuna imposed by a general failure of aragonitic preservation. Alternatively, when the Triassic gastropods do reappear they have some striking similarities to the preceding Pennsian forms. Is this because they are so-called ‘Lazarus taxa’, or are they convergent with strongly homeomorphic shells? Erwin & Pan (1996, p. 229) conclude by noting ‘few phylogenetic studies have investigated clade relationships across the Permo-Triassic boundary, but those that have (for brachiopods, bryozoans and asteroids) all suggest that traditional systematics may exaggerate the magnitude of extinctions at higher levels and miss phylogenetic links at lower levels’. Even the celebrated transition between brachiopod-dominated Palaeozoic seas and the post-Pennsian abundance of bivalve molluscs, which had been cast into a fortuitous correlation of respective decline and rise by Gould & Calloway (1980), has been reanalysed by Sepkoski (1996; see, also, Rhodes & Thompson 1993). The Pennsian mass extinction certainly remains very significant for both groups, but Sepkoski (1996) convincingly reconstructs their histories into a new context of dynamic interaction and competition.

4. EXTINCTIONS

Parallel to the depiction of the history of diversity there has been extensive documentation of the evidence for past extinctions, notably the five mass extinctions. Of these the episode at the end of the Cretaceous (K/T) has attracted unparalleled attention. A convenient distinction between mass and background extinctions receives some support from an analysis of the factors favouring species survival in certain groups, such as the molluscs (Jablonski 1989), according to the severity of the event. On the other hand, Raup’s (1991, 1992) formulation of the so-called ‘kill curve’ suggests that its end-members— relatively benign with limited loss of species to planet-wide trauma respectively—are connected by a continuum of increasing disaster. Another influential view that has received wide support is Gould’s (1985) identification of tiers of influence for evolutionary process that transcend those operating at the level of the species and presumably are open to neodarwinian explanations. In this formulation mass extinctions are said to play a key role, in resetting ecologies and shaping the future development of post-catastrophe diversity (see, also, Courtillot & Gaudemer 1996). But are matters as simple as this?

The one instance that has received general agreement is the end-Pennsian event (Erwin 1993, 1995), marking the replacement of the Palaeozoic faunas by those of the Mesozoic. Given the severity of this episode (see Ershet et al., 1993), with extreme estimates of species loss reaching 96% (Raup 1979), it is not surprising that after the protracted time for recovery (see below) the marine communities were markedly different. Yet there is still some danger of exaggeration. Thus, once important groups of Palaeozoic marine organisms such as the trilobites (Fortey & Owens 1990) and rugose corals (Scrutton 1988) were already small components of even early Permian ecosystems. Other groups, such as the ammonoids and sea urchins, certainly suffered badly but subsequently rebounded to even greater levels of success. The possible pattern of extinction in the gastropods is particularly intriguing (Erwin 1996a; Erwin & Pan 1996). This is because the post-catastrophe faunas are markedly depauperate, yet a simple reading of the data may be misleading. One possibility is that absences are only apparent, and reflect a taphonomic lacuna imposed by a general failure of aragonitic preservation. Alternatively, when the Triassic gastropods do reappear they have some striking similarities to the preceding Pennsian forms. Is this because they are so-called ‘Lazarus taxa’, or are they convergent with strongly homeomorphic shells? Erwin & Pan (1996, p. 229) conclude by noting ‘few phylogenetic studies have investigated clade relationships across the Permo-Triassic boundary, but those that have (for brachiopods, bryozoans and asteroids) all suggest that traditional systematics may exaggerate the magnitude of extinctions at higher levels and miss phylogenetic links at lower levels’. Even the celebrated transition between brachiopod-dominated Palaeozoic seas and the post-Pennsian abundance of bivalve molluscs, which had been cast into a fortuitous correlation of respective decline and rise by Gould & Calloway (1980), has been reanalysed by Sepkoski (1996; see, also, Rhodes & Thompson 1993). The Permian mass extinction certainly remains very significant for both groups, but Sepkoski (1996) convincingly reconstructs their histories into a new context of dynamic interaction and competition.

In terms of three of the other mass extinctions, the evidence for wholesale remodelling of ecosystems is rather weaker. This is especially true of the late Ordovician event (see Droser et al., 1997), which is also enigmatic in as much as its proposed correlation with a short but severe glaciation, has no parallel with either of the other two episodes of Phanerozoic glaciation in the Permo-Carboniferous and Plio-Pleistocene, respectively. The Devonian mass extinction certainly marks some significant changes in marine ecosystems, especially among the reefs (Scrutton 1988; McGhee

Phil. Trans. R. Soc. Lond. B (1998)
recovery appears to have been relatively rapid and this feature may further underline the apparent peculiarity of events, whereas for at least the late Ordovician extinctions seems to be the case for the end-Permian and Cretaceous the extraordinarily protracted lag-times for recovery to (see, for example, Galbrun 1997).

Maastrichtian dinosaurs is widespread, in at least some Mexico, but also because of the oblique nature of the impact and the movement of the ejecta curtain north-

that this region suffered extensive devastation not only because of its proximity to the impact site at Chicxulub, but also because of the oblique nature of the impact and the movement of the ejecta curtain northwards (Schultz & D’Hondt 1996). Whereas evidence for Maastrichtian dinosaurs is widespread, in at least some areas there is a possibility of extinction before the impact (see, for example, Galbrun 1997).

There is another feature of the aftermath of mass extinctions that deserves further consideration, and that is the extraordinarily protracted lag-times for recovery to ecological stability and biotic diversity. This certainly seems to be the case for the end-Permian and Cretaceous events, whereas for at least the late Ordovician extinctions recovery appears to have been relatively rapid and this feature may further underline the apparent peculiarity of this event. The evidence for the post-Permian trauma has long been appreciated in terms of the conspicuously depauperate Lower Triassic faunas (see, for example, Schubert & Botijer 1995). More recently an even grimmer scenario has been presented by Botijer and co-workers (1995, 1996), who record a notable abundance of stromatolites that grew in what are interpreted as open marine settings. Whereas other such examples are known from elsewhere in the Phanerozoic, including the present day (Rasmussen et al. 1993; Reid et al. 1995), they remain exceptions. The relative abundance of the Triassic stromatolites recalls the state of affairs in the Precambrian when, as previously noted, the abundance of stromatolites is generally interpreted as a direct reflection of the absence or at least ineffectiveness of metazoan grazers. As also mentioned previously, the connection between metazoan activity and stromatolite diversity is not accepted by all workers, and Grotzinger (1990) has invoked an hypothesis that looks to amounts of carbon dioxide and degrees of carbonate saturation in the oceans as the factors controlling stromatolite growth. This too may provide a link with the Permian debacle, because Knoll et al. (1996; see, also, Grotzinger & Knoll 1995) have argued that rapid ocean turnover and release of carbon dioxide led to severe hypercapnia in animals and consequent physiological stress. In any event for the marine environment of the Triassic to be pushed towards a situation reminiscent of the Precambrian, underlines yet further the severity of the Permian extinctions and the scope of the ecological vacancies awaiting reoccupation later in the Mesozoic. On present evidence it seems more likely that the protracted lag-time in the Triassic was caused by extrinsic stress, rather than by an innate inability for communities to reaggregate and restore diversity.

The case for protracted ecological instability in the aftermath of the Cretaceous event is better documented, although overall the extent of the crisis seems to have fallen short of the equivalent early Triassic episode. Although the degree of stratigraphic resolution is still relatively crude, it is already clear that there was a time-scale of both disaster and more importantly recoveries, with some environments and biotas appearing to rebound much more promptly than others (table 1). This compilation is very preliminary, and apart from considerable uncertainty about the time intervals, especially at the shorter end of the scale, it also begs the question of what is meant by the term ‘recovery’. Nevertheless, such a dissection of response to the catastrophe is overdue, as is an extensive investigation of regional variations (Barrera & Keller 1994; MacLeod 1995; Keller et al. 1997; but, see Raup & Jablonski 1993).

The evidence for perturbations, before eventual recovery, comes from several sources. These include the slow re-emergence from unspecified refugia of the so-called ‘Lazarus’ taxa, such as certain generalized micro-morphic brachiopods (see, for example, Surlýk & Johansen 1984; see, also, Johansen 1989a,b). It is even more convincingly demonstrated in terms of the dramatic fluctuations in population size, morphological variability and species composition in the communities of planktonic foraminifera (Gerstel et al. 1986, 1987). Enhanced vital effects of autotrophic fractionation of carbon isotopes in the coccolithophorids may also be attributed to ecological

Phil. Trans. R. Soc. Lond. B (1998)
Table 1. Time-scales for destruction and recovery of biotas across the Cretaceous–Tertiary boundary
(All figures are approximate and reflect uncertainties of sedimentation rates, magnetostratigraphy and radiometric dating.)

<table>
<thead>
<tr>
<th>time</th>
<th>effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 s</td>
<td>annihilation around impact site (ca. 30,000 km²) of Chicxulub³</td>
</tr>
<tr>
<td>1 min</td>
<td>earthquakes, Richter scale ten³</td>
</tr>
<tr>
<td>10 min</td>
<td>spontaneous ignition of North American forests⁵</td>
</tr>
<tr>
<td>60 min</td>
<td>impact ejecta crosses North America⁴</td>
</tr>
<tr>
<td>1 h</td>
<td>tsunami swamp Tethyan coastal margins⁵</td>
</tr>
<tr>
<td>1 week</td>
<td>first extinctions</td>
</tr>
<tr>
<td>9 months</td>
<td>dust clouds begin to clear¹</td>
</tr>
<tr>
<td>10 years</td>
<td>very severe climatic disturbance (especially cooling) ends⁴</td>
</tr>
<tr>
<td>1000 years</td>
<td>continental vegetation begins to recover: end of ‘Fern Spike’¹¹</td>
</tr>
<tr>
<td>1500 years</td>
<td>initial recovery of some deeper-water benthic ecosystems⁶</td>
</tr>
<tr>
<td>7000 years</td>
<td>recovery of some deeper-water benthic ecosystems⁶</td>
</tr>
<tr>
<td>70,000 years</td>
<td>oceanic anoxia diminishes⁵</td>
</tr>
<tr>
<td>100,000 years</td>
<td>final extinction of dinosaurs¹⁰</td>
</tr>
<tr>
<td>300,000 years</td>
<td>final extinction of ammonites¹⁰</td>
</tr>
<tr>
<td>500,000 years</td>
<td>large fluctuations in oceanic ecosystems begin to moderate¹³</td>
</tr>
<tr>
<td>1,000,000 years</td>
<td>open oceanic ecosystems partly recovered⁶</td>
</tr>
<tr>
<td>2,000,000 years</td>
<td>marine mollusc faunas mostly recovered¹⁴</td>
</tr>
<tr>
<td>2,500,000 years</td>
<td>global ecosystems normal¹</td>
</tr>
</tbody>
</table>

Sources of information:

¹This assumes a crater diameter of ca. 180 km (Hildebrand et al. 1995).
²See Toon et al. (1997).
³See Alvarez et al. (1995).
⁴See Covey et al. (1994).
⁵See Pope et al. (1994).
⁶See Bucy & Bucy (1986); no precise estimate of the duration of the ‘Fern Spike’ appears to be available.
⁷See Coccioni & Galeotti (1994); these data refer to Spanish stratigraphic sections and elsewhere the benthic foraminifera seem to have been less effected (see, for example, Kaino 1992).
⁸See Zahn et al. (1991) for a possible instance, but Lofgren et al. (1990) for a rebuttal.
⁹See Rigby et al. (1987) for a possible instance, but Loqren et al. (1990) for a rebuttal.
¹¹See Barrera & Keller (1994) for data based on foraminifera.
¹²See Alcala-Herrera et al. (1992) for data based on calcareous nannoplankton.
¹³See D’Hondt et al. (1996), in addition many other workers on microfossils have quoted similar recovery time.
¹⁴See Hansen et al. (1993).
Figure 3. A comparison of the sedimentary record at Deep Sea Drilling Project Site 528 (South Atlantic) in terms of magnetic susceptibility (solid curve) and carbonate content (solid line) across the Cretaceous/Tertiary (K/T) boundary. The magnetostratigraphy and radiometric ages (Ma) are depicted on the left-hand side and the right-hand column is an enlargement of the K/T section; figures refer to metres below the sea floor with the K/T boundary at 407.31 m. The magnetic susceptibility is largely controlled by iron concentrations, which occur mostly in detrital clays. The carbonate is largely derived from calcareous microplankton, including coccolithophorids and foraminifera. The figure emphasizes the instability of the post-catastrophe oceanic system in comparison with the latest Cretaceous (from D’Hondt et al. 1996; modified with permission of the publisher, the Geological Society of America, Boulder, CO, USA. Copyright © 1996 Geological Society of America).
hard-hit (see, for example, Johnson & Hickey 1990), but made a rapid recovery (see Tschudy & Tschudy 1986). Discussion of pre- and post-impact floras, however, is not straightforward because in the early Tertiary there were also significant long-term climatic changes (see Wolfe 1990). Although the radiation of the mammals has received wide attention, only recently has it become a focus for quantitative investigations which point to a very rapid radiation (J. Alroy, personal communication).

As was first construed, the impact hypothesis for the K/T event emphasized global darkening resulting from the ejection of rock dust and, it was subsequently realized, soot into the upper atmosphere as the forcing mechanism for extinction by both the cessation of photosynthesis and climatic perturbations, especially cooling. This was no doubt highly deleterious, but present estimates suggest that severe darkening did not persist for more than a few months and that the postulated plummeting of surface temperatures was offset by the heat reservoir of the oceans so that climatic disturbances may have been severe rather than catastrophic (Covey et al. 1994). It is also possible that the trauma was enhanced by the production of acid rain, with a sulphur source either in the bolide itself or vaporized from the Chicxulub impact site that was rich in anhydrite (CaSO₄).

Paradoxically the main accumulation of data pertaining to the K/T impact has led to the comparative neglect of the biological effects of other impacts. In part this is probably because a rather extensive search for extraterrestrial signatures at other stratigraphic intervals recording mass extinctions has been for the most part negative (Orth et al. 1984, 1986; McGhee et al. 1986; Goodfellow et al. 1992; Erwin 1993) or even when detected (Hodych & Dunning 1992; Bice et al. 1992; Claeys & Casier 1994) still seems to be inadequate as a forcing mechanism. Nevertheless, recent calculations of the energy released from main impacts (figure 4) are sobering. In particular the calculations by Toon et al. (1997) deserve detailed consideration. Some effects, such as large earthquakes and local blast, are perhaps more germane to the fragile infrastructures of our civilization and may be less significant in terms of the biosphere as a whole. Others, however, are certainly not.
entirely lacking. Nor do the puzzles end here. A potentially gigantic crater has been identified in the Kalahari desert of South Africa, possibly up to 350 km diameter, and dated at about 145 Ma (Koeberl et al. 1997). The near coincidence with the Jurassic-Cretaceous boundary is intriguing, yet to date the evidence for severe extinctions across this interval is limited.

The proponents of coordinated stasis, a concept returned to in §5, also identify times of rapid turnover and it might be possible to equate some of these with extraterrestrial impacts. It is also likely that some so-called ‘event-horizons’, the stratigraphic potential of which has received close attention by stratigraphers, are a consequence of highly energetic collisions. Nevertheless, the fossil record of extinctions appears to show little evidence for catastrophic extirpation, although one wonders if the vagaries of range completeness (or lack thereof) and taxonomic splitting are masking some events. It is also the case that sedimentologists and stratigraphers are seldom attuned to the possible evidence for impact driven processes, not least because, so far as I am aware, there are no coherent models for sedimentological consequences of the spontaneous ignition of several million square kilometres of land, followed by its part inundation beneath a 50 m high tsunami.

While the influence of extraterrestrial impacts is surely more important than presently realized, certainly it may be premature to assume that mass extinctions per se, however caused, are the leading force in shaping the history of diversity (see, also, Jackson & McKinney 1990). Nor need they represent the only extrinsic factors that pummel the biosphere. The role of Milankovitch-driven climatic cycles in stressing the environment and promoting extinctions has received the forceful advocacy of Bennett (1997). Here too we have the possibility of recurrent stress and a mechanism that could, in principle, be equated to the turnover pulse hypothesis of Vrba (1995).

5. THE STABILITY OF COMMUNITIES

Ecologists have long emphasized the contrasting views of community structure as formulated by the individualistic and haphazard view of H. A. Gleason, versus the recurrence and ecological interlocking identified by F. E. Clements (and subsequently C. Elton) (see Walter & Patterson 1994; Jablonski & Sepkoski 1996). In terms of the fossil record of the most recent past, especially with respect to terrestrial floras, there has been a strong emphasis on the Gleasonian view whereby, driven by climatic shifts, plant species reaggregate in novel combinations so that the present-day distributions and associations may have

Figure 6. Estimates of the effects of impacts by iron asteroids and long-period (LP) comets in terms of production of deep-water tsunamis with a 10 m amplitude in a 4 km deep ocean (figure is redrawn from Toon et al. 1997, fig. 7); reprinted with permission from Reviews of geophysics and the American Geophysical Union.

Figure 7. The effect of ballistic ejecta re-entering the atmosphere in terms of the thermal radiative flux experienced on the Earth's surface. Debris infall is estimated to take 20 min, and that only half the energy generated by re-entry is radiated downward and half of that actually reaches the surface. The impactor is taken to be an icy object travelling at 50 km s$^{-1}$. The threshold for spontaneous ignition is bracketed by flash heating (upper line) and more protracted irradiation (lower line). Note that for impacts yielding an energy of $10^{8}$ megatons, fires will be ignited across much of the globe (figure is redrawn from Toon et al. 1997, fig. 15 (lower)); reprinted with permission from Reviews of geophysics and the American Geophysical Union.

Yet doubts remain (Conway Morris 1998). In part this is because of the difficulty in correlating Milankovitch-driven climate change with any known extinction, combined with the fact that most species' durations far exceed the observed cyclicitics. With such low species mortalities versus the frequencies of oscillation, this makes the hypothesis difficult to test. There is also the uncertainty of deciding whether the much-enhanced Milankovitch cyclicitics of the last couple of million years, and their accepted connection to wildly oscillating glacial and interglacials, would apply with equal force in a past where the influence of such oscillations was apparently more muted. Not only that, but as Bennett (1997) himself is at pains to point out, the possibly chaotic behaviour of the Solar System (for examples, see Laskar 1989; Sussman & Wisdom 1992) may mean that the orbital forcing of the deep past is such that the present cyclicitics provide a guide rather than a key to environmental stress and potential extinction.
no precise counterpart in floras even a few thousand years old (Overpeck et al. 1992; see, also, Bennett 1997). Similar conclusions have been reached by Coope (1979, 1987): in his extensive analyses of Quaternary beetle assemblages, which also point to an exceptionally dynamic and unpredictable history of migration and reassociation. These influential studies refer to relatively short periods of geological time, and co-occur with the very marked shifts in the environment imposed by the recurrent swings between glacial and interglacial regimes. A very much longer perspective was taken by Buzas & Culver (1994), who documented the changing species compositions of communities of shallow marine benthic foraminifera of the North American Atlantic Coastal Plain during much of the Cenozoic. In this case, embayments (known as the Salisbury and Albemarle embayments) were successively isolated and reconnected with the main species reservoir of the open shelf by regressions and transgressions respectively. Buzas & Culver (1994) demonstrated that each reimmigration established a community of benthic foraminifera with a markedly distinct taxonomic profile, and in this sense represented only a subset, apparently drawn at random, of the main species pool.

This turbulent view of diversity, as expressed in communities, is almost directly antithetical to a widespread identification of ancient communities as stable, recurrent entities with such properties as persistence, stasis, ecological similarity, and self-regulation. This notion of community stability that is only occasionally punctuated by turnover and extinction encompasses quite a wide range of views. Nevertheless, broadly consistent approaches have been presented both by some paleoecologists (see DiMichele & Phillips 1996), as well as by extensive series of investigations in shallow marine communities (for examples, see Aberhan 1993; Miller 1993; Brett et al. 1996; Schopf 1996; Tang & Bottjer 1996; see, also, DiMichele 1994), and also reef ecosystems (Pandolfi 1996). Not all palaeoecologists see such patterns, and in particular trenchant criticisms of the general concept of community stability as now enshrined in the concept of coordinated stasis has been offered by Alroy (1996). This analysis deserves particular attention, because it is based on a carefully prepared database of Tertiary North American mammals and also on account of the exactly formulated tests for the several hypotheses connected to the depiction of coordinated stasis. Significantly, Alroy (1996) failed to identify either sudden turnovers of diversity or a correlation between the rates of extinction and origination. A critique in a related vein is offered by Bambach & Bennington (1996). Their review, of whether communities can in any sense be said to evolve, is more sympathetic to the ideas of coordinated stasis, but also notes that even within each interval of supposed stability there is still considerable turnover. Such ripostes may only be a part-solution to solving this divergence of interpretation. One wonders whether the climatic extremes of the last few million years are especially disruptive, notably for terrestrial communities (see, also, DiMichele 1994). There is also the obvious criticism of the problems of time-averaging (see, for example, Bambach & Bennington 1996) and the difficulties of precise stratigraphic correlation in some older sequences. These factors might conspire to impose an apparent uniformity on community structure that originally was much more dynamic and labile.

### 6. CRYPTIC DIVERSITY

Despite the biases and imperfections of the fossil record the pattern of Phanerozoic diversity is generally accepted to be robust. One obvious pitfall, that of the relative proportion of soft-bodied to skeletalized taxa at various geological intervals, deserves specific mention because of the increasing level of documentation of fossil Lagerstatten such as the Burgess Shale-type faunas (Conway Morris 1998) or mid-Mesozoic Plattenkalk (Bernier & Gaillard 1994). Such an analysis, however, is not straightforward.

The first problem is that there is a spectrum of taphonomic resistances of animals, meaning there is no simple distinction between skeletal and soft-bodied. Another problem is the wide variety of preservations encountered in fossil Lagerstatten, including carbonaceous films, diagenetic carbonate nodules, pyritization, phosphatization and polymerized tree resin (amber). In the last case, for example, Henwood (1993) ingeniously compared the taxonomic profile of insects in amber faunas from the mid Tertiary of the Dominican Republic with those obtained in similar circumstances by entomologists using standard trapping techniques (e.g. light trap, fogging) to show that, notwithstanding the exquisite fossil preservation (see Poinar 1992), amber faunas contain comparable sampling biases. At present, all it seems possible to conclude is that fossil Lagerstatten are themselves incomplete repositories of diversity, may occur in ‘unusual’ environments (e.g. lagoons, deeper water), but still do not suggest that our existing ideas of ancient diversity are hopelessly warped by taphonomic factors.

This does not mean, however, that we can afford to be complacent because several potentially more severe problems exist. These concern the role of phylogenetic systematics, the apparent discrepancies between molecular and taxonomic data, and hitherto overlooked ‘hot-spots’ of diversity. The first point, that of the role of phylogenetic systematics and especially the cladistic methodology in the context of the study of diversity has not gone unnoticed. In brief, present-day diversity curves are for the most part plotted at the taxonomic level of family. Generic databases do exist, but are largely unpublished and seldom taxonomically up-to-date. The principal complaint of phylogenetic systematists is that many of the taxa are paraphyletic and even polyphyletic (see Smith 1994), and thereby give a skewed assessment of original diversity. The most powerful demonstration of this was by Patterson & Smith (1987) who reanalysed taxonomic data, specifically of fish and echinoderms of which they have expertise, respectively, that had been used to identify a periodicity in mass extinctions (Raup & Sepkoski 1986). Using criteria of phylogenetic systematics the periodicity of extinction, in at least these two groups, vanished. Since then the debate as to whether Linnean taxa, especially at the level of family, can provide sufficient proxy for diversity has continued in a rather half-hearted and inconclusive manner (Sepkoski & Kendrick 1993). Another area of phylogenetic systematics that may deserve more attention in this context is the concept of ghost taxa (see Norell 1992). These are taxa identified as logically necessary in the
construction of a cladogram, but unknown in actuality. Clearly where ghost taxa are abundant, again our perceptions of diversity may be skewed.

The other important area of cryptic diversity concerns the emerging discrepancies between established taxonomic data and information inferred from molecular biology. In brief, three categories of divergence stand out, of which the first two involve proposed times of origination. Consider angiosperms and mammals. The palaeontological record indicates the flowering plants originated in the early Cretaceous, whereas following the end-Cretaceous extinctions the latter underwent an important adaptive radiation in the Paleocene that resulted in appearance of most of the main mammal groups. The molecular data point to a very different story. In the case of the angiosperms, dates of divergence varying from early Jurassic to Carboniferous are inferred (for examples, see Wolfe et al. 1989; Brandl et al. 1992; Martin et al. 1993; Laroche & Bousquet 1995; Kolukisaoglu et al. 1995). The case of the mammals is somewhat less dramatic, but molecular data still suggest a significant divergence of mammal groups in the Mesozoic (Janke et al. 1994; Hedges et al. 1996; Springer et al. 1997), as well as discrepancies within the Cenozoic radiation (see, for example, Frye & Hedges 1995). Similar conclusions may also apply to the divergence times of some groups of birds (Harlby et al. 1997).

The standard responses to these divergences in the estimated times of origination are well known, and they may be correct. The principal objection, of course, is that the basis for the molecular estimates assumes a clock-like behaviour for substitution. If molecular clocks periodically run fast, then it may be easier to reconcile these data with the fossil record. It is sometimes supposed that such speeding-up of the molecular clock is coincident with adaptive radiations, although accepting the consequences of such an assumption appears to have been neglected. An alternative source of potential confusion, again widely acknowledged, is the lateral transfer of genetic material (Syvanen 1994). The third possibility is to decouple assumptions about molecular architecture and the taxonomic pigeon-holing that too often accompanies the recognition of major groups. One of the most valuable concepts of phylogenetic systematics is that of the stem-group, whereby the emergence of major groups is documented in the serial acquisition of characters. The implications of this approach for the angiosperms and mammals are interesting, albeit in somewhat different ways. In the former case there is a highly controversial record of pre-Cretaceous plants that have been interpreted as having angiospermous characteristics (see Crane et al. 1995). The case of the mammals is more complex, because broadly there are two schools of thought. One (Darling et al. 1996, 1997; Wade et al. 1996) has emphasized an apparently very ancient origin for the planktonic foraminifera, hundreds of millions of years before their appearance in the Lower Jurassic. The extreme branch lengths of the planktonic foraminifera are obscure, although Pawlowski et al. (1997) proposed that ultraviolet radiation in the upper water column may be responsible. This needs to be tested against other planktonic groups, such as the coccolithophorids, and within the planktonic foraminifera themselves according to the positions they adopt in the water column. Philippe et al. (1994) have drawn attention to a large discrepancy in the estimated origination of diatoms from molecular data versus the fossil record, the latter indicating a Jurassic appearance of limited diversity followed by a Cretaceous effloresence. They conclude that a cryptic history of ‘naked’ diatoms, possibly extending as far back as the Precambrian, preceded their acquisition of siliceous skeletons. Philippe et al. (1994) also review (and reject) the alternative explanation of molecular rate changes. This is, of course, the standard riposte of those
anxious to reconcile times of molecular divergence with the fossil record. This latter view raises some uncomfortable difficulties. First, if molecular clocks run fast during adaptive radiations it implies a linkage between phenotype and molecules that are generally thought to be ‘invisible’ to morphology. Alternatively, clock dates could be telescoped if the initial stages ran fast before it slowed along a hyperbolic function. The idea of a molecular clock ‘ageing’ seems to find little favour, as yet.

The case for an appearance of dinoflagellates long before the first fossil records relies on molecular data, but in the form of chemical fossils extracted from sediments that date back at least as far as 700 Ma (Summons 1992). The dinoflagellates are especially appropriate in this respect because of their possessing characteristic biomarkers, and the potential for inferring a cryptic history of this group has been explored by Moldovan et al. (1996).

Not all the data from molecular biology, however, point to cryptic intervals of evolution predating first appearances in the fossil record. In an extensive analysis of the molecular clock data of enzymes Doolittle et al. (1996) concluded that the times of divergence of many major groups of both prokaryotes and eukaryotes were substantially younger than indicated in the fossil record. This analysis has not been well received by palaeontologists, and in addition Martin (1996) has proposed that this analysis by Doolittle et al. (1996) is seriously confounded by the prevalence of lateral gene transfer. For the most part such transfer, whereas admitted for groups such as bacteria and plants, has been considered to be unimportant in animals. Such a view is coming under scrutiny with possible examples of hybridization in groups as disparate as cnidarians (Odorico & Miller 1997) and chordates (Spring 1997).

There is a somewhat different area of molecular biology which also hints at large gaps in our knowledge of the diversity of microbial communities, notably of prokaryotes. Research by Barns and co-workers (for example, see Barns et al. 1996a,b), by using genetic probes in microbial systems such as hot-springs, is revealing an astonishing array of otherwise unknown prokaryotes. Many show affinities to the group of Archaeabacteria known as Crenarchaeota, but these bacteria also expand the taxonomic range and suggest the divisions between the Kingdoms are less absolute than earlier thought. Although there is reason to believe that these prokaryotic groups (and especially the thermophiles) are phylogenetically very ancient, at first sight the prospects of recognition in the fossil record are not encouraging. Nevertheless, fossil hydrothermal systems are known (Walter 1996; Knoll & Walter 1996), and there is also the potential for diagnostic biomarkers to be recovered (Summons 1992; Summons et al. 1996).

7. CONCLUSIONS

The pattern of diversity as seen in the fossil record appears to be robust, but it is almost entirely based on a traditional Linnean taxonomy (see Foote 1996). The impact of phylogenetic systematics, with their insistence on the recognition of monophyly may lead to significant changes in the tabulation of diversity. Revisions may also stem from the recognition of cryptic diversities which may be inferred from the disconnecting gaps that appear to separate times of origination as seen in the fossil record, as opposed to inferences made on the basis of molecular biology. Reconciliation of these seemingly contradictory data may go some way to resolving the problems in understanding the speed and extent of adaptive radiations if it transpires that many key features are acquired before the ‘releasing event’ such as a mass extinction or climatic shift.

Nevertheless, whatever shifts in the pattern of diversity emerge, it is surely the case that the known curve (figure 1) is not simply constructed on the basis of artefacts. In particular, times of diversification are real. At present it is still very difficult to identify the key factors involved, and it would be unwise to identify an over-riding and unique principle. First, there is appeal to the many extrinsic factors. Here one might include: marine regressions; biogeographic changes of continental distribution and oceanic structure that are ultimately controlled by plate tectonics; changes in atmospheric composition and/or ocean chemistry; massive volcanism; and changes in substrate type as engendered by taphonomic feedback that includes skeletal production and microbial activity. Of those factors that are usually labelled as intrinsic there are also many examples. They tend to be rather specific, relating to particular groups, but they are many. A couple of examples, taken more or less at random, are: the development of the mammalian hypocone (Hunter & Jernvall 1994); and the ability of Ordovician echinoderms to colonize new substrates (Sprinkle & Guensburg 1995).

Is the history of diversity then merely the product of a concatenation of a vast number of independent variables, whose chance associations and timings result in what for us is a familiar world but one which in actuality is almost entirely contingent? Alternatively, is there a deeper mathematical structure to diversity? As it happens, this is a metaphysical question (Conway Morris 1998), and reveals the endless but potentially fruitful tension between the belief in a bedrock of physical realities and the operation of an historical process, with what to many of us comprises a clear directionality. The historical dimension has formed the bulk of the discussion here, but can we also identify a deeper architecture that might provide a theoretical underpinning to our enterprise? One possible avenue is the work by Sole et al. (1997). These workers identified self-similarity in different time series of ancient diversity, mostly at the taxonomic level of family but also genera. The fractal structure, it is suggested, could be consistent with self-organized criticality. The data source they used (Benton 1993) is open to serious, albeit inevitable, criticisms of incompleteness and bias (Conway Morris 1994b).

Thus, this analysis is perhaps more suggestive than compelling, especially as the observed power spectra consistently only approach the expected scaling exponent of $\beta=1$. Nevertheless, the conclusion of Sole et al. (1997, p. 766) that ‘the internal biotic organization (is) the basic component for the response of the biosphere to external perturbations’ is consistent with a widely held view of the world-biota being interactive, coherent and possessing its own dynamic that, while not immune to abiological factors, was and is not in a state of recurrent lability on account of a ceaseless pummelling by the external environment. In this sense, ecological factors are important determinants of diversity and during rare intervals, e.g.

*Phil. Trans. R. Soc. Lond. B* (1998)
early Triassic or Palaeocene, exert exceptional levels of stress on biological communities. Certainly on any familiar ecological time-scale such intervals of one or more million years are protracted, but otherwise long-term disruption that might be predicted from such factors as cometary impacts, Milankovitch forcing, or severe climatic fluctuations (e.g. El Niño) in point of fact appear to be muted.

Let us accept, then, the 4 Ba history of evolution follows a script that is largely based on biological interactions. In this context, the role of interspecific competition as a force for moulding diversity might seem pre- eminent, yet in actuality it has encountered intense scepticism (see, for example, Benton 1996), in part because of the problem of trying to define credible tests. In a three-month trial Schluter (1994) identified competition as the driving force in morphological divergence of stickleback fish, but many palaeontologists would question whether such results can be extrapolated into a geological time-scale, notwithstanding Schluter’s explicit link to adaptive radiations. On a much larger time-scale, Sepkoski’s (1996) reformulation of diversity plots points towards interactions between brachiopods and bivalve molluscs. This study suggests that competitive patterns. Sepkoski’s (1996) analysis also echoes the earlier conclusions of Raup (1981) who in charting a secular decline in trilobite diversity noted that it could be explained by competitive displacement. Nevertheless, one of the key objections to the notion of competition is the persistence of survivors representing the supposedly inferior group. How seriously should we take this point? When we consider the range of groups for which extirpation via an obvious mass extinction is unlikely, e.g. graptolites, thelodonts, creodonts, or have at best relic status as the trilobites did in the Permian or monoplacophorans do in Recent seas, then competitive displacement remains a reasonable possibility. This does not provide a complete explanation. A well-known problem concerns the encrusting cyclostome and chelostome bryozoans, in which notwithstanding the latter’s conspicuous success in overgrowth has still failed to dislodge the weaker cyclostomes in competition for substrate space (Lidgard et al. 1993). Nevertheless, as McKinney (1995) points out, subsequent to the rise of the chelostomes the cyclostomes are forced into a temporal refuge by restrictions in colony size and early sexual maturation. Thus, there are other ways of staying in the race. Not only that, but in the real world there are other groups, notably algae, sponges and tunicates, that are even more effective at overgrowth. Competition need not spell the doom of a group, but the net result is a more complex and possibly faster-paced world (McKinney 1993).

In this way the history of diversity can be seen to be progressive in as much as the diversity of species, the richness of communities and the complexity of ecologies in the Recent appear to be unrivalled in comparison with the geological past. To point out that bacteria are still with us really misses the point. Not only is their ‘expertise’ biochemical, but they are displaced into the refugia of hot springs, sediments, and animal guts. As we have seen, if the animal biosphere enters extreme crisis, as seems to have been the case in the early Triassic, then the microbial mats and stromatolites return. But if the history of diversity is now clearer, how we actually got to where we are is still obscure. With our present predicament, in one sense this hardly matters. The appearance of our species with unique capabilities means that, so to speak, all bets are now off. The fact we are a product of evolution and represent a tiny twig on the immense tree of life is not germane to our destruction of the richest ecosystem the world has ever seen.

I thank the organizers of this Discussion meeting, Anne Magurran and Bob May, for their invitation to participate. Sandra Last is warmly thanked for typing numerous versions of this paper, as is Sharon Capon for drafting the figures. My research is supported by the Natural Environment Research Council and St John’s College, Cambridge. Cambridge Earth Sciences Publication 555.

REFERENCES


Johnson, K. R. & Hickey, L. J. 1990 Megafloral change across the Cretaceous/Tertiary boundary in the northern Great

*Phil. Trans. R. Soc. Lond. B* (1998)


Oji, T. 1996 Is predation intensity reduced with increasing depth? Evidence from the west Atlantic stalked crinoid Endocorynus parvus (Gervais) and implications for the Mesozoic marine revolution. Paleobiology 22, 339–351.


Raymond, A. & Metz, C. 1995 Laurussian land-plant diversity during the Silurian and Devonian: mass extinction, sampling bias, or both? Paleobiology 21, 70–95.


---

*Phil. Trans. R. Soc. Lond. B* (1998)