Fig. 3 The size ranges of North American rodents — an early group, and two later groups — to show the shift from small sizes to a broad range of body sizes including many large ones. The index of size is the length of the first lower molar, which varies directly with overall body size. (After Stanley 1973.)

rapidly — the Late Eocene *Basilosaurus* was 21 m long, after 15—25 million years of evolution.

Stanley (1979) noted that, in contrast, 'large' molluscs took much longer to evolve. The first large free-swimming clam was *Megalomoidea* which appeared after nearly 100 million years of radiation. The first large epifaunal bivalves, the inoceramid rudists of the Jurassic and Cretaceous, took nearly 400 million years to appear.

Amongst land plants, large size arose at the end of the Devonian, and especially in the Carboniferous, with the first tree-like clubmosses (*Lepidodendron*, 45 m high) and horsetails (*Calamites*, 16 m high). This had taken 50—60 million years of land plant evolution. Really giant gymnosperms (*Sequoia* and other redwoods) are known from the Jurassic, as much as 250 million years after the origin of land plants, and 150 million years after the origin of gymnosperms.

Advantages and disadvantages of large size

Numerous advantages of large size have been postulated (Stanley 1973): improved ability to capture prey or escape from predators, greater reproductive success, increased intelligence (large bodies have large brains), better stamina, expanded size range of possible food items, decreased annual mortality, extended individual longevity, and increased heat retention per unit volume. Protection from predation would seem to be a great advantage. Adult elephants and rhinoceroses have no regular threat from carnivores today. However, thick-skinned mammals of the Oligocene to Pleistocene of the Northern Hemisphere and South America were subject to attacks by specially adapted sabre-toothed cats — the Machairodontidae in North America, Europe, Africa, and Asia, and the Borhyenidae in South America. The sauropod dinosaurs are assumed to have been immune from attack since the largest predatory dinosaurs could only have tackled very young sauropods, or dying adults.

A disadvantage of large size may be greater proneness to extinction. This is not simply an attribute of large size, but rather an expression of specialization. Large animals are often more restricted in their niches, in their scope for adaptation, than smaller relatives. Their need for large amounts of food, or for particular environmental conditions, may make them more likely to suffer when habitats change. Also, the fact that large animals tend to have small population sizes, and hence small gene pools, makes their hold on life seem more precarious. The death of a few more individuals than normal may precipitate species extinction.

Bakker (1977) showed that terrestrial tetrapods surviving mass extinctions in the Late Palaeozoic and Mesozoic tended to be of small body size. Thus, the large dicynodonts and dinocephalians of the Late Permian died out, leaving smaller dicynodonts and cynodonts to cross the system boundary. A similar explanation has also been given for selectivity in the Cretaceous—Tertiary event on land (Section 2.13.7). In more general terms, Stanley (1979) suggested that species longevity varies with the reciprocal of body size: small species tend to survive longer than large species. This is supported by evidence from the Pliocene and Pleistocene mammalian fossil record. The only modern species that can be tracked back before 3 Ma are small mammals. All the large ones arose after that, and this is probably not an artifact of a poor fossil record since such forms are more readily fossilized than small ones.

Within any clade, lineages of large organisms may be expected to display shorter taxon durations, lower rates of speciation, and higher rates of extinction (Stanley 1979), and hence greater volatility in the face of environmental stress. These ideas have yet to be tested thoroughly. They are of added interest since they could be seen as charac-

teristics that are subject to species selection (since these are not organism-level features). They could also potentially be interpreted as examples of the 'effect hypothesis' (Vrba 1983; see also Section 2.6). This hypothesis suggests that species-level characteristics, such as species duration or broad ecological adaptation, may be incidental effects of individual characters, such as dietary or habitat preferences. Natural selection, acting on organisms, might select for large body size, which in turn might produce higher extinction rates within a lineage. These higher rates could be interpreted as an incidental *effect* of natural selection, rather than as a result of species-level selection. These ideas are still highly controversial.

References

Alexander, R.McN. 1985. Mechanics of posture and gait of some large dinosaurs. *Zoological Journal of the Linnean Society* **83**, 1—25.

Bakker, R.T. 1977. Tetrapod mass extinctions — a model of the regulation of speciation rates and immigration by cycles of topographic diversity. In: A. Hallam (ed.) *Patterns of evolution as illustrated by the fossil record*, pp. 439—468. Elsevier, Amsterdam.

Fraser, A.I. 1962. Wind tunnel studies of the forces acting on the crowns of small trees. *Reports on Forest Research* **1962**, 178—183.

Hokkanen, J.E.I. 1986. The size of the largest land animal. *Journal of Theoretical Biology* **118**, 491—499.

Stanley, S. M. 1973. An explanation for Cope's Rule. *Evolution* **27**, 1—26.

Stanley, S.M. 1979. *Macroevolution: pattern and process*. W.H. Freeman, San Francisco.

Vrba, E.S. 1983. Macroevolutionary trends: new perspectives on the roles of adaptation and incidental effect. *Science* **221**, 387—389.

2.11 Rates of Evolution — Living Fossils

D. C. FISHER

Introduction

The study of rates of evolution encompasses a wide variety of approaches to characterization of the amount of evolutionary change within particular groups of organisms, over specified time intervals. The high level of interest that palaeontologists and evolutionary biologists have shown in this subject is not surprising, since rates are a common focus in the analysis of any process. The importance of rates, however, is often only marginally attributable to intrinsic interest in 'how rapidly' or 'how slowly' a process operates. Rather, information on rates tends to be used as a means of investigating the underlying dynamics of the process in question, or sometimes as input for analysing the dynamics of a related process. Much of the work on rates of evolution has thus been directed toward a better understanding of the dynamics of evolutionary change. Studies have been designed with the intent of comparing rates of evolution in a variety of ways — within and between particular taxonomic groups, ecological settings, and lineage geometries (e.g.,

ancestor—descendant sequences that include lineage splitting versus ones that do not). While interesting generalizations are emerging, a greater appreciation is also being gained of the difficulties of quantifying rates of evolution.

'Living fossils' is a term frequently used to denote extant representatives of groups of organisms that have survived with relatively little change over a long span of geological time. Such groups are implicitly recognized as having displayed unusually low rates of evolution. In both professional and popular literature, living fossils collectively appear to have attracted more attention than have groups displaying unusually high rates of evolution. This may be partly because, in keeping with the inherent paradox of the term 'living fossil', evolutionary history is expected to involve conspicuous change, and it is surprising when it does not. In addition, evolutionary rate statements are commonly (though not exclusively) framed in terms of putative ancestor—descendant pairs, and it is easier to recognize these when the total amount of change has been small than when it has been large. Instances of

living fossils are thus more likely to be accepted on prima facie grounds than are instances of higher evolutionary rates. In any event, living fossils have frequently provided a focus for discussions of evolutionary rate and have helped to clarify some of the factors that may be involved in promoting or inhibiting evolutionary change.

Three ranges of values for evolutionary rates

G.G. Simpson was one of the early major contributors to the quantitative study of evolutionary rates. He proposed that rates be classified by their absolute value as 'low', 'medium', or 'high'. Although this might be considered trivial, Simpson (1944, 1953) argued that frequency distributions of evolutionary rates for sufficiently inclusive sets of taxa typically contain three discrete modes, allowing low, medium and high categories to be recognized on non-arbitrary grounds. This empirical claim suggests some degree of disjunctness in the operation of the processes and/or constraints that interact to produce evolutionary change. Simpson coined the term 'bradytely' to refer to the phenomenon of supra-specific taxa that have shown consistently low rates of evolution. Bradytely thus encompasses the same general concept implied by 'living fossil', but without the arbitrary stipulation that a representative of the group be alive today. Simpson also suggested 'horotely' to refer to taxa comprising the middle mode in the spectrum of observed evolutionary rates and 'tachytely' to refer to supraspecific taxa showing consistently high rates of evolution. Although Simpson's (1953) demonstration of the multimodality of evolutionary rates has subsequently been shown to be flawed (Gingerich 1983; Stanley 1985), the terms denoting these rate categories (especially bradytely and tachytely) have had considerable heuristic value. They are now commonly used to refer to ranges of rate values regardless of whether multimodality has been demonstrated independently. For instance, in a study applying the terms in this latter fashion, Raup and Marshall (1980) showed that rates within several orders of mammals were significantly higher (e.g. Cetacea and Rodentia) or lower (e.g. Perissodactyla and Carnivora) than the mean for all mammalian orders. However, whether evolutionary rate distributions (at a given rank, within some more inclusive group of organisms) tend to show some 'typical' form and, if so, whether that form is multimodal, unimodal but non-normal, or unimodal and normal, are presently open questions.

Qualitative categories of evolutionary rates

Evolutionary rates may also be categorized by the aspect of evolutionary change that is measured. Three commonly discussed categories are genetic, morphological, and taxonomic rates. However, various subdivisions of each of these are also significant. For instance, genetic rates include rates of DNA nucleotide substitution and rates of gene rearrangement, among others. These two kinds of rates refer to different processes of genetic change, acting at different levels in the hierarchy of genetic structure. Each offers its own perspective on the general phenomenon of evolutionary change, and it is conceivable that each will show a different frequency distribution, even over the same large group of taxa. In the same way, morphological rates are sometimes subdivided into 'size' rates and 'shape' rates, since these two factors are commonly treated as different, though not unrelated, aspects of morphology. Finally, taxonomic rates include various approaches to measurement of the longevities and rates of origination and extinction of taxa. Terminology for categories of taxonomic rates varies somewhat among authors, and each category may be further subdivided according to the taxonomic rank treated. In each case, the meaning of such rates depends critically on the underlying taxonomic philosophy. The type of taxonomic rate that will be focused on here is the rate of origination of new taxa of specified rank, since this corresponds most closely to a 'rate of evolution' (i.e. without introducing aspects of extinction rate).

In this categorization of evolutionary rates, genetic and morphological rates refer to changes in the genotype and phenotype, respectively. An alternative convention is to distinguish between molecular and morphological rates of evolution. This retains all aspects of genotypic change within molecular evolution, but adds to it components of protein evolution that would ordinarily be considered changes in the phenotype, albeit at a molecular level.

Although molecular data are usually available only for living organisms, increasing effort is being focused on extraction of some molecular data from appropriately preserved fossil material (Section 2.1). Still, except for the success of such efforts, molecular rates can only be measured directly over relatively short timespans. Alternatively, they may be computed from the cumulative divergence of contemporaneous taxa. In this case, some parsimony assumption is used to partition change between or among the separate lineages involved. Although

this approach may seem to remove molecular rates from the domain of palaeontology, we must still relate measured divergence to the time interval over which it has developed — the time since the most recent common ancestor of these taxa. Tectonic or palaeogeographical data suffice for this in certain instances, but palaeontological data provide the most commonly applicable constraints on the time of splitting of lineages. For this reason, and because of their common focus on analysis of the pattern and process of evolution, palaeontology and studies of molecular evolution are closely related (Section 2.1).

Measurements of morphological rates may also be based on comparisons among contemporaneous taxa for which the divergence history is relatively well known. However, when morphological features can be sampled in a succession of stratigraphic intervals, we have the option of calculating rates 'directly' from the fossil record. Since any source of morphological disparity between samples will contribute to perceived evolutionary rate, it is important to be aware of, and if possible control for, non-evolutionary components of variation (e.g. differential ontogenetic representation, differential taphonomic biases, or range shifts in clinally varying populations). If it can be argued that consecutive samples represent a series of ancestors and their descendants within a species-level lineage — an ideal situation that approximates 'tracking' morphology through time — the resulting rate is referred to as a 'phyletic' rate. However, if the phylogenetic context of consecutive samples is more complex or unresolved than this, the rate is better referred to as a 'phylogenetic' rate (Raup & Stanley 1978). Phylogenetic rates imply a disclaimer recognizing that increments of change may have been measured between samples that do not bear a direct ancestor—descendant relation to one another. Depending on the history of morphological change and the pattern of phylogenetic relationships linking consecutive samples, phylogenetic rates may be either greater or less than corresponding phyletic rates (i.e. the phyletic rates that might be measured if an arguably ancestor—descendant sequence were available). Both of these types of rate represent transformation within a 'lineage' (broadly construed, possibly at a supraspecific level), but they differ in the degree of resolution with which the lineage can be traced.

Taxonomic origination rates are likewise designed to quantify change through time, but they differ fundamentally from the rates discussed thus far. To the extent that new taxa are erected to recognize some increment of morphological change within lineages, origination rates incorporate a transformational component comparable to that assessed by molecular and morphological rates. However, origination rates also include a component representing the cladogenetic (or lineage splitting) aspect of evolutionary change. The relative contributions of these two components — lineage transformation and lineage splitting — are difficult to quantify and rarely reported. They vary from group to group depending both on taxonomic practice and on the actual evolutionary history of the group under study.

Units of measurement for evolutionary rates

Genetic or molecular rates are sometimes quantified in terms of the number of events involving a particular type of change, per time interval. Comparisons of molecular rates may be normalized for the number of entities 'at risk' for change (e.g. number of nucleotide substitutions per site, per million years), but this is not practical in all instances (e.g. computing the number of potential gene rearrangements). Molecular rates based on distance measures (e.g. DNA—DNA hybridization, immunological distance) are given in units appropriate to the distance measure utilized.

Morphological rates may be expressed as change in the value of some morphological variable (any appropriate units of measurement), per time interval. However, variables of different dimensionality (e.g. lengths versus areas) must be divided by an appropriate factor before they can be properly compared. Moreover, we are usually interested in proportional rather than absolute changes in morphology. Given the scaling relationships of most morphological variables (and their variances), a convenient solution is to measure morphological rates in terms of differences in the logarithm of the value of the variable of interest. A difference of a factor of e (base of natural logarithms, 2.718) per million years was defined by Haldane (1949) as a morphological rate of 1 darwin (d).

Rates of taxonomic origination may be measured as the number of new taxa (within a given higher taxon) per time interval. This is often expressed as a percentage increase, normalized for the length of the time interval. Rate of origination may also be calculated from the rate of change in total diversity at a given taxonomic level and the rate of extinction at that level. In interpreting origination rates, it is important to consider such possible complications as differential effects of taphonomic and mono-

graphic biases, and differential application of taxonomic practice within and between groups being compared (Raup & Marshall 1980). From an evolutionary standpoint, however, a more fundamental issue with rates of taxonomic origination is that they lump together information on lineage transformation and lineage splitting. Given the current unevenness of our detailed phylogenetic knowledge of most groups, this may be an unavoidable compromise, and indeed, it offers some benefits of convenience and succinctness in the representation of evolutionary history. However, it is to be hoped that more phylogenetically discriminating approaches to studying diversification will be developed in the future.

The effect of measurement interval on evolutionary rates

Measured rates are commonly treated as independent of the interval length over which they are measured. For processes occurring at approximately constant rates, this characterization is acceptable. However, for any variable-rate process, the measured rate is an average and may be influenced strongly by rate fluctuations during the measurement interval. Depending on the temporal structure of rate fluctuations and the range of intervals being considered, measured rates will be more or less susceptible to biasing effects from interval length.

Some molecular rates appear to behave in 'stochastically constant' fashion, at least over certain time-spans (commonly of the order of tens of millions of years). The relative constancy of these rates (with both rate and constancy varying from one molecular system to another) has led to the proposal of the 'molecular clock' hypothesis (see also Section 2.1). According to this hypothesis, molecular difference, once calibrated to reflect rate of change, can be used as a measure of time since lineage divergence (Fitch 1976). However, even for molecular clocks that are relatively 'well behaved' over a particular time interval within a given group, there is growing evidence that observable change has either accelerated or decelerated at other times during the history of that group (Goodman *et al.* 1982; Gingerich 1986). For divergence times that span periods of significant rate change, systematic biases can be anticipated.

The factors thought to influence morphological rates (see below) are known to fluctuate on a variety of time-scales. Because neither the highest nor the lowest rates are likely to be maintained over protracted periods of time, the largest range of variation should be observed in comparing rates measured over the shortest time intervals. For the same reason, there should be a tendency toward intermediate values, which are due to averaging of rate fluctuations, when measuring over longer intervals. Since morphological rates are typically expressed in terms of net change in the value of some morphological variable, changes in the *direction* of morphological change, as well as in the rate of change *per se*, contribute to the moderation of rates measured over longer time intervals. This interaction is partly responsible for the decline in maximum observed morphological rates with increasing measurement interval (Fig. 1). However, as Gingerich (1983) pointed out, the lower, and to some extent the upper bounds of the distribution of observed rates in Fig. 1B are also influenced by factors unrelated to evolutionary process. The lower bound corresponds to a practical limit of measurement precision, beyond which earlier and later forms would not usually be recognized as different, yielding a rate of zero. The upper bound, on the other hand, represents an effective limit beyond which pairs of earlier and later forms differ so strongly that their relationship, and hence their appropriateness for a rate calculation, is likely to be questioned. The result is a tendency for longer measurement intervals to yield lower rates. Because of these biasing factors, comparison of rates measured over very different time intervals is a non-trivial problem. Many comparative studies of evolutionary rates have not adequately dealt with this issue.

Taxonomic rates are also affected by measurement interval, but not in all the ways noted above. As with morphological rates, rates of origination calculated over longer intervals are likely to be damped by averaging a range of shorter-term values. However, rates of origination are not moderated by changes in the 'direction' of evolution; 'new taxa' are new taxa, even if they show reversals in certain attributes. In addition, with rates of origination, low values do not suffer an interval-related bias based on measurement precision, nor do high values necessarily engender suspicion of lack of relationship.

The effect of stratigraphic completeness on evolutionary rates

Stratigraphic completeness (see also Section 3.12) could in principle affect the precision of palaeontologically documented divergence times, but in practice, phylogenetic uncertainties and disconti-
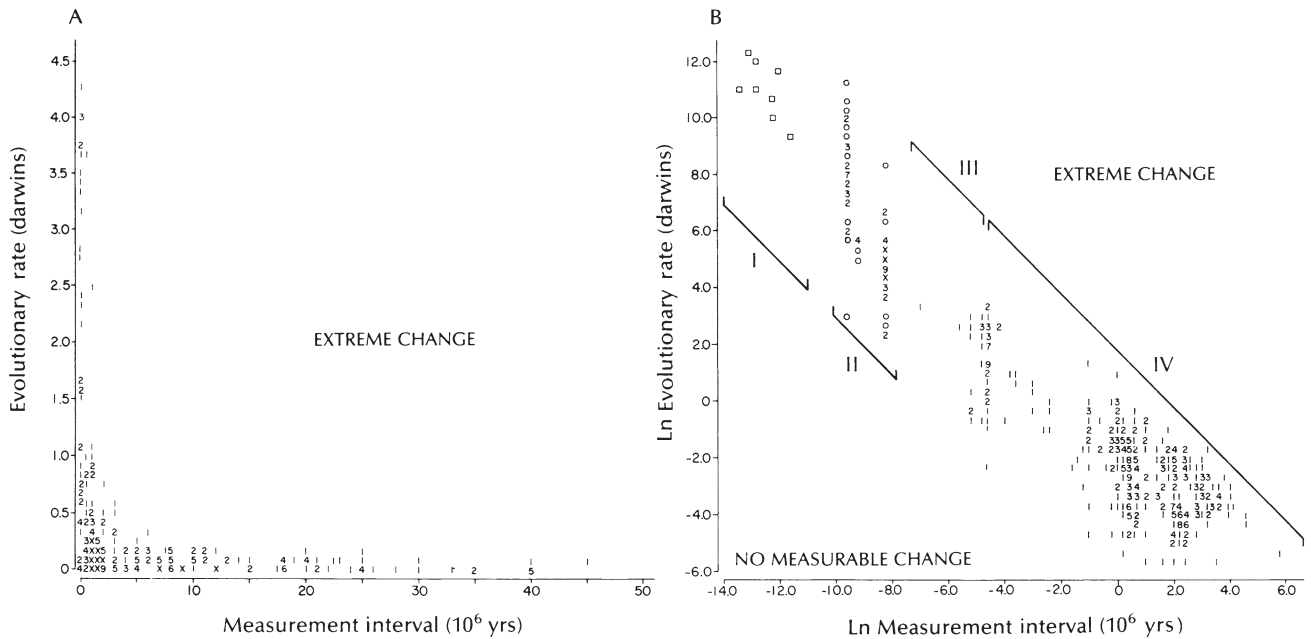
Fig. 1 Inverse relationship between morphological rates of evolution and the time intervals over which they are measured (after Gingerich 1983), illustrating some of the biasing effects discussed in the text. A, Subset of rates shown in B, plotted on linear axes. B, Logarithmic transformation of rate distribution; time intervals range from 1.5 to 350 million years. Rates plotted in domain I (open squares) represent laboratory selection experiments; domain II (open circles, and digits for multiple cases; X > 9) represents historical colonization events; domain III (digits) represents post-Pleistocene events; and domain IV (digits) is drawn from the pre-Holocene record of invertebrates and vertebrates.

nuities in the preserved record of taxa (even within intervals that have a sedimentary record) are more important sources of error in these estimates. Stratigraphic completeness increases in importance, however, when morphological or taxonomic rates read 'directly' from the fossil record are considered. In a relatively incomplete section, the actual age difference between two samples may be either much greater or much smaller than their estimated age difference based on linear interpolation from dated levels. This translates into substantial imprecision in rate measurements. In order to minimize this problem, Dingus & Sadler (1982) suggested that rates only be measured at levels of resolution for which stratigraphic sections can be considered complete (i.e. for which each included interval of given magnitude is likely to be represented by sediment). Following this recommendation, relatively incomplete sections limit us to longer time intervals for rate measurement and thus, through the biasing effect of interval length, tend to yield lower rates than might be seen in more complete sections. Relatively incomplete sections also tend to reduce measured rates of origination.

Factors affecting actual rates of evolution

Having explored some of the factors that tend to distort perceptions of evolutionary rates, the sources of real variation in such rates will now be discussed. Among the more conspicuous of these are controls of the rate of transformation within established, species-level lineages. These include: mutation rate; generation time; degree of resource specialization; and the nature, amount, and distribution of variability within populations. Population size may also be important but is probably overshadowed by population structure — the pattern and scale of subdivision of populations and the degree of reproductive interaction between those subdivisions. Other factors are at least partly extrinsic to the species in question: rate of environmental change; ecological factors such as the level of interspecific competition; and, in general, the intensity of selection (assuming selection and fitness are defined so that intensity of selection is not trivially equivalent to rate of evolution).

Another group of controls overlaps somewhat with the first but may be distinguished as operating at a different level in the genealogical hierarchy. It

consists of factors that determine the rate of initiation of new species-level lineages. Speciation rate assumes particular importance in a punctuated view of evolution, but its role in influencing evolutionary rate is not dependent on the predominance of a punctuated mode of evolutionary change. Intrinsic controls on speciation rate include such factors as dispersal ability (also relevant as a determinant of population structure) and degree of resource specialization. There are also extrinsic controls, such as rate or incidence of habitat fragmentation by geomorphic or tectonic processes.

Living fossils — alternative definitions

Living fossils figure in discussions of evolutionary rates as a conspicuous and yet potentially tractable case in which the relationship between a large-scale evolutionary pattern and its underlying causes may be explored (Eldredge & Stanley 1984; Schopf 1984). As noted above, the central concept in the definition of living fossils is survival over long periods of time with minimal morphological change. Auxiliary criteria have been appended by various authors and do indeed apply to certain cases traditionally recognized as living fossils. However, they are much less applicable to others. For instance, a relict geographical distribution and greatly diminished present (relative to past) diversity characterize *Sphenodon* (a rhynchocephalian) and *Nautilus* (a nautiloid cephalopod), but not *Limulus* and related genera (horseshoe crabs). Likewise, *Latimeria* (a coelacanth) and *Neopilina* (a monoplacophoran) represent clades once thought to be extinct, but *Lepisosteus* (a gar) and *Lingula* (an inarticulate brachiopod) have long been known from both fossil and Recent biotas. Living fossils are sometimes referred to as 'species' that have persisted for inordinately long periods of time, but few if any instances are actually founded on well documented species-level identity. The most generally useful definition therefore focuses on supraspecific taxa that have shown unusual morphological conservatism.

One of the most commonly cited living fossil groups is the Xiphosurida, or horseshoe crabs. Fig. 2 provides some sense of the morphological conservatism that can be seen within this group, comparing the extant species *Limulus polyphemus* with the Triassic *Limulus vicensis*. While the generic identity of these two species may be questioned (Fisher *in* Eldredge & Stanley 1984), their overall anatomical similarity is evident. Other species with-
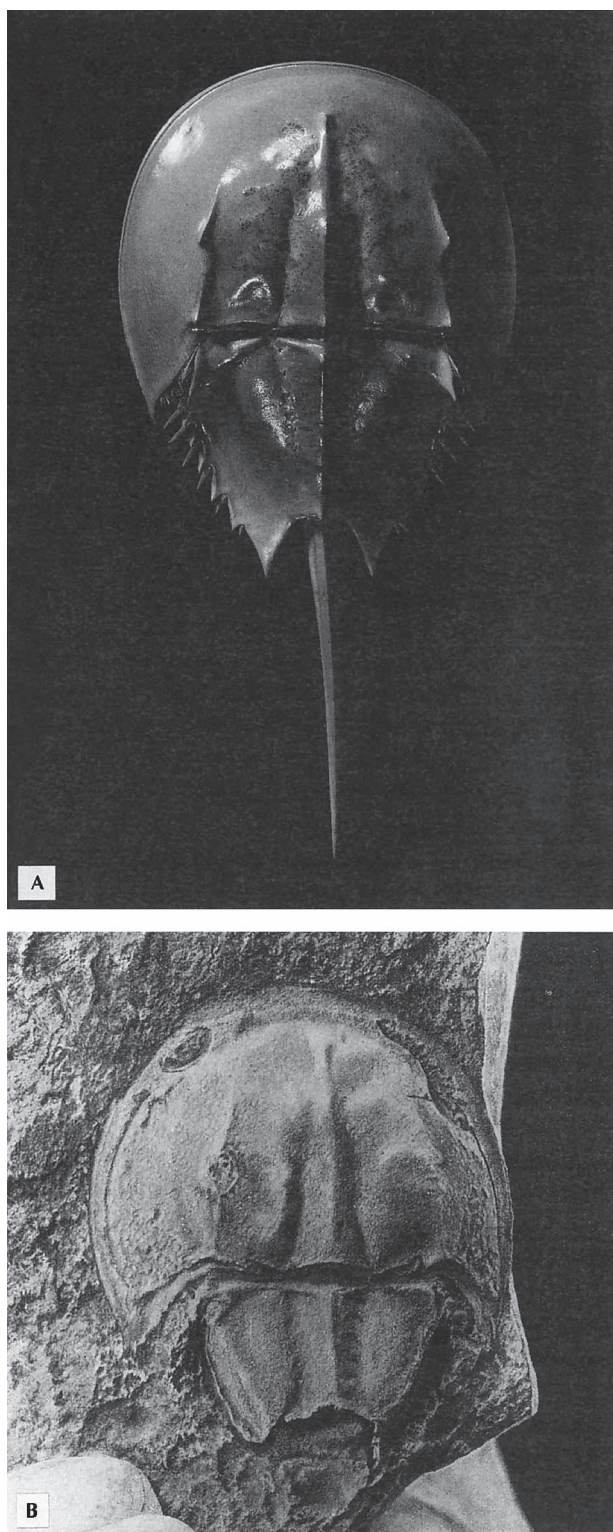

Fig. 2 Horseshoe crabs, a commonly cited living fossil group. A, Dorsal aspect of a juvenile *Limulus polyphemus*, Recent, distributed along much of the eastern coast of North America; *c.* one half actual size. B, Dorsal aspect of a specimen of *Limulus vicensis*, Triassic, France; *c.* actual size. The tail spine is not preserved on this specimen, but it was presumably present originally. (From Bleicher 1897.)

in the group show greater morphological divergence, but the reputation for bradytely has focused on comparisons such as that given here.

Bradytely — alternative explanations

The problem posed by living fossils is to explain the general phenomenon of bradytely. Simpson's (1944, 1953) interpretation was that the low rates of long-term morphological evolution shown by bradytelic lineages are a consequence of unusually low rates of intraspecific phyletic transformation (Fig. 3A). This appears to be a testable proposition, but it has thus far received little direct, empirical evaluation (perhaps because few bradytelic groups have a sufficiently continuous fossil record). However, some of the factors that have been suggested as responsible for low rates of phyletic transformation (e.g. unusually low levels of morphological or genetic variability) have been assessed within bradytelic groups and found not to differ significantly from values typical of nonbradytelic taxa (e.g. Selander et al. 1970). Other factors that could in principle be responsible (e.g. extreme habitat stability, or strongly canalized development) are difficult to test. Some factors do seem to hold for a wide range of bradytelic groups and have been thought to contribute directly to low rates of intraspecific change (e.g. ecological generalization and broad physiological tolerance; Simpson 1953). Nevertheless, consideration of alternative explanations is clearly warranted.

Another approach to interpreting bradytely steps up a level in the hierarchy of evolutionary processes — from intraspecific interactions to the circumstances surrounding speciation events (cladogenesis). It depends, furthermore, on the proposition (associated with the concept of punctuated equilibrium) that most morphological change is accomplished during and driven by cladogenesis, and that the subsequent history of species tends to be dominated by morphological stasis. Under this characterization of evolution, a low rate of intraspecific transformation would be the norm and would not be seen as a sufficient cause of bradytely. However, bradytely might be due to unusually low rates of speciation within bradytelic lineages (Fig. 3B); according to this interpretation, low speciation rate would allow few opportunities for morphological change and would thus restrict a lineage to a relatively low rate of change averaged over the long term (Eldredge 1979). As long as speciation is understood as a process that is not itself dependent
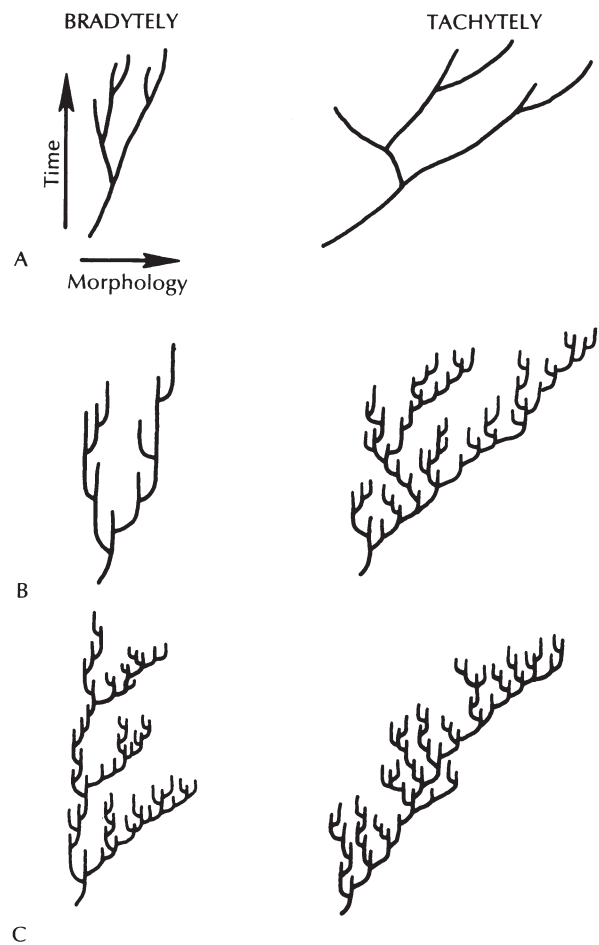

Fig. 3 Schematic representation of three explanations of controls on long-term rate of morphological evolution. (After Fisher in Eldredge & Stanley 1984.) A, The contrast between bradytely and tachytely may be due to differences in the rate of intraspecific morphological transformation. B, The same contrast may be due to differences in rate of speciation. C, Bradytely and tachytely may also reflect higher-order patterns of differential survival and cladogenesis.

on morphological change, this interpretation represents a novel perspective on bradytely. Yet there is still a question as to why certain taxa, retrospectively recognized as bradytelic, show such a low rate of speciation. One possible answer has been suggested by the observation that a number of bradytelic taxa also show tendencies toward eurytopy — i.e. they have, at least in many respects, relatively broad, generalized ecological requirements. In this, they contrast with stenotopic taxa, which have relatively narrow, specialized requirements. It has been suggested that, relative to stenotopic taxa, eurytopic taxa are less subject to directional selection, often have broader geographical

ranges, and tend to have populations that are less susceptible to range disruption and consequent reproductive isolation (Jackson 1974; Eldredge 1979; Vrba *in* Eldredge & Stanley 1984). This may result in a lower rate of speciation and a lower likelihood of morphological divergence during speciation. The case studies of bradytely in Eldredge & Stanley's (1984) compendium offer qualified support for the association of bradytely, low speciation rate, and eurytopy, but rigorous evaluation of this pattern is difficult because of the lack of quantitative indices of morphological conservatism or eurytopy. In addition, measurements of speciation rate are subject to significant sampling problems, such that even an evaluation of the relationship between speciation rate and subjective assessments of bradytely and eurytopy would be complicated.

A third interpretation of bradytely is that it arises at an even higher level in the genealogical hierarchy, as a result of differential survival of relatively primitive and relatively derived lineages within a clade (Fig. 3C). Treated simply as a phylogenetic pattern, bradytely may or may not have any single lower-level cause, but whether it does or not, it could be independent of any systematic difference in intraspecific rates of transformation or rates of speciation (Fisher *in* Eldredge & Stanley 1984).

While any of these three explanations of bradytely might operate in isolation from the others, they are not mutually incompatible. Nor can the possibility be ruled out that different instances of bradytely are traceable to different mixes of factors operating at a variety of levels. Although there are thus no simple answers, the investigation of bradytely has led to an expanded appreciation of the possible controls of long-term evolutionary rates.

References

Bleicher, M. 1897. Sur la découverte d'une nouvelle espèce de limule dans les marnes irisées de Lorraine. *Bulletin des Séances de la Societé de la Societé des Sciences de Nancy* **14**, 116–126.

Dingus, L. & Sadler, P. M. 1982. The effects of stratigraphic completeness on estimates of evolutionary rates. *Systematic Zoology* **31**, 400–412.

Eldredge, N. 1979. Alternative approaches to evolutionary theory. *Bulletin of the Carnegie Museum of Natural History* **13**, 7–19.

Eldredge, N. & Stanley, S.M. (eds) 1984. *Living fossils.* Springer-Verlag, New York.

Fitch, W.M. 1976. Molecular evolutionary clocks. In: F.J. Ayala (ed.) *Molecular evolution*. Sinauer, Sunderland, Mass.

Gingerich, P.D. 1983. Rates of evolution: effects of time and temporal scaling. *Science* **222**, 159–161.

Gingerich, P.D. 1986. Temporal scaling of molecular evolution in primates and other mammals. *Molecular Biology and Evolution* **3**, 205–221.

Goodman, M., Weiss, M.L. & Czelusniak, J. 1982. Molecular evolution above the species level: branching pattern, rates, and mechanisms. *Systematic Zoology* **31**, 376–399.

Haldane, J.B.S. 1949. Suggestions as to quantitative measurement of rates of evolution. *Evolution* **3**, 51–56.

Jackson, J.B.C. 1974. Biogeographic consequences of eurytopy and stenotopy among marine bivalves and their evolutionary significance. *American Naturalist* **108**, 541–560.

Raup, D.M. & Marshall, L.G. 1980. Variation between groups in evolutionary rates: a statistical test of significance. *Paleobiology* **6**, 9–23.

Raup, D.M. & Stanley, S.M. 1978. *Principles of paleontology*, 2nd edn. Freeman, San Francisco.

Schopf, T.J.M. 1984. Rates of evolution and the notion of living fossils. *Annual Review of Earth and Planetary Sciences* **12**, 245–292.

Selander, R.K., Yang, S.Y., Lewontin, R.C. & Johnson, W.E. 1970. Genetic variation in the horseshoe crab (*Limulus polyphemus*), a phylogenetic "relic". *Evolution* **24**, 402–414.

Simpson, G.G. 1944. *Tempo and mode in evolution*. Columbia University Press, New York.

Simpson, G.G. 1953. *The major features of evolution*. Columbia University Press, New York.

Stanley, S.M. 1985. Rates of evolution. *Paleobiology* **11**, 13–26.

2.12 Mass Extinction: Processes

2.12.1 Earth-bound Causes

A. HALLAM

Introduction

The idea that mass extinctions could be caused by strictly Earth-bound phenomena is an old one, dating back to the so-called heroic age of geology in the early part of the nineteenth century. Following the pioneering extinctions research of his compatriot G. Cuvier, the French geologist Elie de Beaumont proposed that catastrophic, virtually instantaneous upheavals of mountain ranges at infrequent intervals through geological history caused drastic environmental changes leading to the destruction of a high proportion of the Earth's biota. The correlation between episodes of diastrophism and times of major organic turnover was also noted by the American geologist T.C. Chamberlin at the beginning of this century, and by European geologists such as E. Suess and J.F. Umbgrove (Hallam 1981a). Modern research on tectonic activity suggests, however, that it is too localized geographically and insufficiently 'catastrophic' in time to account satisfactorily for mass extinction events. Attention must be confined to phenomena global in scale that can give rise to drastic changes in the physical environment. The only plausible contenders are changes in sea level and climate, and episodes of increased volcanicity.

Sea-level

The American palaeontologist Newell (1967) was the first person to make an explicit correlation between mass extinction episodes among Phanerozoic marine invertebrates and eustatic falls in sea-level, attributing the extinctions to increased environmental stress consequent upon substantial reduction of habitat area of shallow epicontinental seas. He distinguished six such episodes: end-Cambrian, end-Ordovician, Late Devonian, end-Permian, end-Triassic and end-Cretaceous. The first two are especially well marked by trilobite extinctions and the last three by ammonite extinctions. Extensive communities of reef-dwelling organisms

were destroyed in the Late Devonian (Section 2.13.3) and end-Triassic (Section 2.13.5) episodes and the calcareous plankton (foraminifera and coccolithophorids) drastically reduced at the end of the Cretaceous (Section 2.13.6). The biggest event of all was at the end of the Permian (Section 2.13.4), when many important Palaeozoic groups went completely extinct, including fusulinid foraminifera, camerate and inadunate crinoids, trepostome and cryptostome bryozoans, rugose corals, and productid brachiopods.

All but the first of these extinction episodes have subsequently been accepted by palaeontologists as the most significant extinction events in Phanerozoic history (Raup & Jablonski 1986). The correlation between major sea-level falls and Newell's mass extinction events is indeed striking (Fig. 1; Jablonski 1986). On a smaller scale, there is an equally striking correlation between the extinction of environmentally sensitive groups such as ammonoids and other episodes of widespread regression, probably correlating with sea-level fall, in both the Palaeozoic and Mesozoic (e.g. Hallam 1987a). Following ecological research on island biogeography, it is clear that smaller habitat areas can accommodate fewer taxa, so reduction in area must lead to lower diversity as the extinction rate increases. Whether the extinction is due to reduced habitat diversity, increased competition, crowding effects, or whatever, the basic empirical relationship appears to be well established.

Critics have pointed out that inferred episodes of significant marine regression do not always correlate with notable mass extinctions of marine organisms. This is most obviously true for eustatic falls of sea-level in the Quaternary and Middle Oligocene, the latter being probably the largest in the Tertiary (Haq *et al.* 1987). At least two explanations can be put forward, both of which take into account the phenomenon of biological adaptation. Quaternary regressions were followed by equally rapid transgressions after geologically short time intervals, limiting the effect of reduced habitat area and permitting a sufficient number of organisms to survive and expand their populations during the succeeding transgressions. Quaternary faunas are likely to have been relatively eurytopic, or environmentally tolerant, because they represent survivors of environmentally stressful Late Cenozoic times. The same
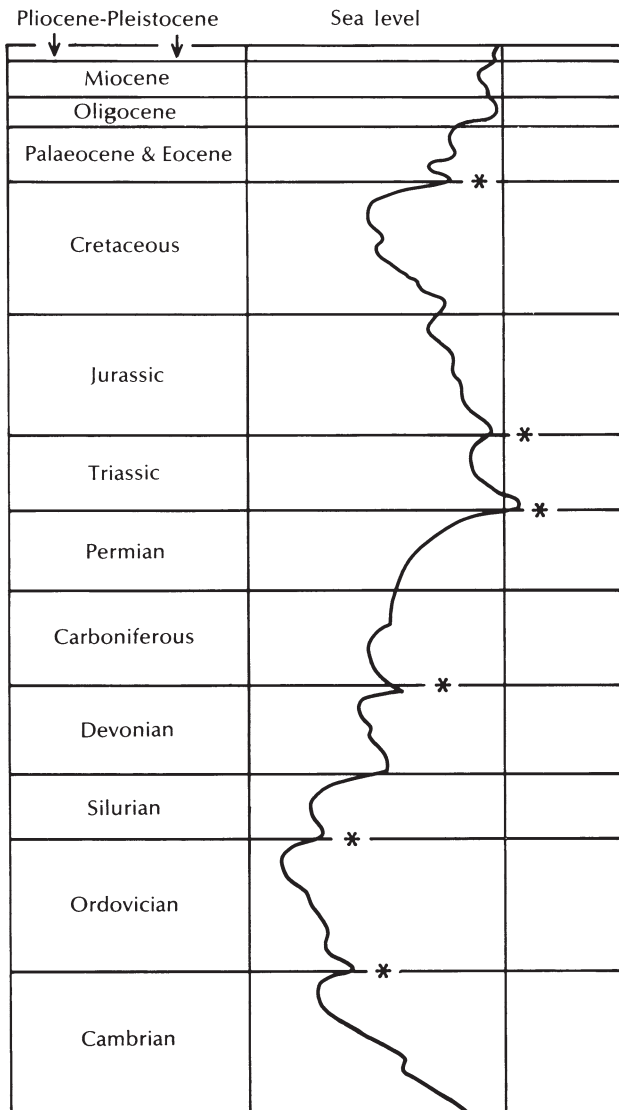
Pliocene-Pleistocene   Sea level

Fig. 1 Phanerozoic sea-level curve with Newell's (1967) six extinction events shown as asterisks. High sea-level to left. (After Hallam 1984.)

consideration may apply also to the Middle Oligocene regression, which followed closely on a significant increase in marine extinction rates across the Eocene−Oligocene boundary. It is likely that for long periods of Phanerozoic time most organisms became so well adapted to conditions of relative environmental stability, including equable climate, that even modest changes of sea-level could have had a striking effect on so-called 'perched faunas' in extensive and extremely shallow epicontinental seas. Such palaeogeographic phenomena cannot be closely matched at the present day, which marks an unusually regressive episode in Earth's history.

The latest Exxon sea-level curve from the Triassic

to the present (Haq *et al.* 1987) does not show unusually large falls at the times of the two greatest marine extinction episodes during this interval, the end of the Triassic and the end of the Cretaceous. The Exxon curve is based largely, however, on seismic stratigraphy, and should not be treated as more than a tentative model to be subjected to testing by other evidence. There is indeed considerable evidence of a major end-Triassic regression (Hallam 1981b), and some strong indications that the extent of the end-Cretaceous regression has been underestimated by Haq *et al.* (Hallam 1987b).

For several events, namely the end-Permian, end-Triassic and end-Cretaceous, mass extinctions in the marine realm appear to correlate closely with mass extinction of some terrestrial vertebrates, notably those large in size, which, because of their relatively low population numbers and reproductive rates, would be more vulnerable to environmental disturbance than smaller organisms (see also Section 2.10). Obviously such extinctions cannot be accounted for by reduction in land area, and a more likely explanation is bound up with the increased continental seasonal temperature contrasts induced by regression of epicontinental seas.

While much attention has been paid to regression as a promoter of extinctions it should be noted that there is a strong association between inferred sea-level rises that follow directly after falls and the spread of anoxic water in epicontinental seas, as recorded for instance by widespread laminated black shales. Habitable areas can be as severely reduced by this means as by regression, with a mass extinction event ensuing. For many extinction events, both major and minor, a clear correlation exists with extensive deposits of black shales. Among the major events the best examples are the basal Silurian and basal Famennian (Devonian), effectively equivalent to the end-Ordovician (Section 2.13.2) and end-Frasnian (Section 2.13.3) extinction events. Among minor events the clearest examples are the Cenomanian−Turonian boundary and Early Toarcian (Hallam 1987a). The spread of anoxic bottom waters may possibly also be implicated as a contributory factor in the end-Permian (Section 2.13.4) and end-Triassic (Section 2.13.5) events.

For much of Phanerozoic history the ocean might have been poorly stratified, in marked contrast to the present-day situation (Wilde & Berry 1984). In consequence the deeper ocean would be more or less anoxic and could not have served as a refuge for

shallow-water organisms at times of regression, or if they were outcompeted by other organisms. It is more than possible that the great bulk of the modern deep-sea fauna, which contains representatives of most phyla, is no older than Tertiary. Since the Late Eocene there has evidently been a system of strong currents induced by Antarctic glaciation, which have served to aerate bottom water in the deep ocean (Hallam 1981a). Lack of a deep water anoxic zone could help to explain why there is no significant extinction recorded for the major Middle Oligocene regression.

The cause of sea-level changes is bound up either with the melting and freezing of polar icecaps or with tectonics, such as the uplift and subsidence of ocean ridges and the splitting or collision of continents. The end-Ordovician event might well have had a glacioeustatic cause, associated with growth and disappearance of the Saharan ice sheet, but for the other major extinction events the most likely cause is tectonoeustatic. This poses a problem, because the rates of sea-level rise and fall produced by plate tectonics are approximately three orders of magnitude lower than for glacioeustasy, thereby allowing more time for organisms to adjust to a changed environment and hence avoid extinction. Unfortunately there are as yet insufficient data from the stratigraphic record, on amount and rate of sea-level change, to resolve this problem satisfactorily.

There remains another possibility, that rapid regressions and transgressions on a regional rather than a global scale could be produced as a result either of changes in the pattern of lateral stresses in the crust (Cloetingh *et al.* 1985) or by the rise of mantle plumes to cause epeirogenic uplift, with related volcanism associated with subsidence (Loper & McCartney 1986). The fact that such changes would not strictly come under the category of eustatic is irrelevant as far as the organisms are concerned, provided that the changes in question are both geographically extensive and rapid, thereby leading to drastic changes in the environment.

Climate

Changes of sea-level could have, as a by-product, some climatic consequences, but climate could of course fluctuate with time independent of eustasy. Stanley (1984, 1987) has been the strongest advocate of the view that temperature changes in the marine realm have been the dominant causal factor in Phanerozoic mass extinctions. This interpretation involves a gross extrapolation from his detailed studies of Plio-Pleistocene molluscan extinctions off the Atlantic and Gulf coasts of the U.S.A. Whereas there is a high rate of species extinctions in this region, there is negligible evidence of contemporary extinctions around the Pacific margins, or the Mediterranean. Stanley maintained that, because the extinctions are regional not global in extent, eustatic changes cannot be invoked. Instead he argued for a more pronounced lowering of temperature on the American east coast than elsewhere, as a result of palaeogeographical factors.

Extending back through time, the next major marine extinction event for which temperature decline can plausibly be invoked is across the Eocene−Oligocene boundary. This 'event' is decidedly not sudden in geological terms and is marked more by a pronounced increase in extinction rate rather than a drastic change over a narrow time interval. There is good independent evidence from oxygen isotopes of a fall in both surface and bottom water temperatures, but no indication from the curve of Haq *et al.* (1987) of sea-level changes significantly larger than at other times in the Tertiary. For pre-Tertiary times, however, the evidence implicating temperature as a causal factor is weak to non-existent, forcing Stanley to resort to some special pleading (though it could be argued that the end-Ordovician event (Section 2.13.2) had an ultimate climatic causation, if the glacioeustatic interpretation is accepted). For example, the largest extinction event of all, at the end of the Permian (Section 2.13.4), took place during a period of climatic amelioration, marked by the Middle Permian disappearance of the Gondwana ice sheet. It is conceivable, of course, that the end-Permian event was induced by an episode of temperature rise, but no plausible case has been made for this.

One of the points that Stanley cited in favour of his temperature control hypothesis is that the most extinction-vulnerable organisms, such as reef dwellers, were tropical in distribution throughout Phanerozoic history. While this may be true, it does not necessarily establish temperature as the key control, because tropical organisms tend to be generally stenotopic, as they are relatively sensitive to a variety of environmental factors.

A really extensive overturn of deep anoxic water at the beginning of episodes of climatic change has been suggested as a possible contributing factor to mass extinction events in the oceans (Wilde & Berry 1984). As discussed above, the rise and spread onto continental shelves of anoxic water is often associated with marine transgressions, so that it may be

unnecessary to invoke climatic change as well.

As regards changes in air temperature, the only satisfactory record comes from Late Cretaceous to Recent terrestrial plants. No striking extinction event has been recorded among these organisms for the Cenozoic, but at the end of the Cretaceous there were significant extinctions in the North Temperate Realm of western North America and Eastern Asia. Whereas the palaeobotanical consensus has related such extinctions to gradual temperature decline through the Late Cretaceous, the most recent research in the North American Western Interior suggests a temperature rise in the Maastrichtian and no significant change across the Cretaceous—Tertiary boundary (Wolfe & Upchurch 1987). Further back in time the evidence from terrestrial plants is more obscure, and has so far not been adequate to establish a convincing picture of climatic change.

Volcanism

The end-Cretaceous extinction event is the one that has received by far the most attention (see also Sections 2.13.6, 2.13.7). Notwithstanding the claims made for extra-terrestrial impact, there is strong evidence for marine regression at this time, suggesting that this phenomenon is involved in the extinctions. Sea-level change cannot account, however, for the drastic extinctions at the Cretaceous—Tertiary boundary of calcareous plankton, nor for such physico-chemical evidence as an anomalous enrichment on a global scale of iridium, and the presence locally of quartz grains with shock-metamorphic laminae, in Cretaceous—Tertiary boundary layers (see also Section 2.12.2). Evidence of this sort has been claimed as conclusive for bolide impact, but in fact a case of at least equal plausibility can be made for terrestrial volcanism on a massive scale (Hallam 1987b). It is known that aerosols enormously enriched in iridium compared with crustal rocks can be expelled from the mantle during flood basalt eruptions. Eruptions of this kind on a sufficient scale over several 100 000 years could produce the observed global enrichment of the element. The Deccan Traps of India, erupted during the magnetic zone that embraces the Cretaceous—Tertiary boundary, are the most obvious candidate. There is good evidence of contemporary explosive volcanism in other parts of the world, and reasonable grounds for believing that such volcanism can generate the pressures required to produce shock-metamorphic laminae in mineral grains.

Massive volcanism over an extended period would have deleterious environmental consequences. It is known that flood basalt fissure eruptions that produce individual lava flows with volumes greater than 100 km³ at very high mass eruption rates are capable of injecting large quantities of sulphate aerosols into the lower stratosphere, with potentially devastating atmospheric consequences. Such volatile emissions on a large enough scale would lead to the production of immense amounts of acid rain, reduction in alkalinity and pH of the surface ocean, global atmospheric cooling, and ozone layer depletion. Atmospheric cooling would be reinforced by ash expelled into the atmosphere by contemporary explosive volcanicity.

Thus for the end-Cretaceous extinctions a compound scenario seems to be required, involving both sea-level fall and volcanicity on an exceptionally intense scale, with associated climatic changes (there is as yet, however, no evidence to support the notion that volcanicity was a direct causal factor for other mass extinction events). Loper and McCartney (1986) noted that increased end-Cretaceous volcanism correlates with a significant change in the geomagnetic field, with a long Cretaceous reversal-free period coming to an abrupt end in the Maastrichtian. They proposed a model involving periodic instability of the thermal boundary layer at the base of the mantle. This layer accepts heat from the core and transmits it upward by way of mantle plumes. As it thickens by thermal diffusion it becomes dynamically unstable and hot material erupts from it. Heat is extracted from the core at a greater rate, increasing the energy supply and hence the magnetic reversal frequency of the dynamo in the fluid outer core. Hot material rises through mantle plumes to the surface to give rise to volcanic activity. Both non-explosive and explosive volcanism can be produced, depending on the condition of the lithosphere, which varies regionally. Increased mantle plume activity has the potential for causing uplift of extensive sectors of continents and hence regression of epicontinental seas. Present-day hotspots are associated with regional topographic bulges, so it is reasonable to infer that most epeirogenic uplifts reflect hot, low density regions in the asthenosphere, derived from plume convection. Epeirogenic subsidence on the continents and marine transgression might be expected to follow episodes of substantial volcanic eruptions.

Fischer (1984) put forward a general hypothesis that relates changes of sea-level, climate, and volcan-

icity to produce two supercycles during Phanerozoic time. Times of high rates of ocean floor spreading and oceanic volcanicity correlate with buoyant ocean ridges and consequently high sea-level stands. Less carbon dioxide is removed from the atmosphere by terrestrial weathering because of reduced continental area, and the volcanicity brings more of the gas to the Earth's surface. Thus the carbon dioxide content of the atmosphere is high, and because of the greenhouse effect the climate is equable, with no polar ice caps. The converse tectonic situation gives rise to low sea-level stands, low atmospheric carbon dioxide, and stronger climatic differentiation between the tropics and the poles. The rates of change involved in such processes appear, however, to be too low to account for mass extinction events. The most promising line of approach in generating terrestrial models is probably a closer investigation of the relationship between sea-level change, continental uplift, volcanism, and mantle plume activity, as has been proposed for events across the Cretaceous—Tertiary boundary (Sections 2.13.6, 2.13.7). The end-Permian extinction episode (Section 2.13.4) is an especially promising candidate for this type of investigation.

References

Cloetingh, S., McQueen, H. & Lambeck, K. 1985. On a tectonic mechanism for regional sea level variation. *Earth and Planetary Science Letters* **75**, 157−166.

Fischer, A.G. 1984. The two Phanerozoic supercycles. In: W.A. Berggren & J.A. van Couvering (eds) *Catastrophes and Earth history*, pp. 129−150. Princeton University Press, Princeton.

Hallam, A. 1981a. *Facies interpretation and the stratigraphic record*. W.H. Freeman, Oxford.

Hallam, A. 1981b. The end-Triassic bivalve extinction event. *Palaeogeography, Palaeoclimatology, Palaeoecology* **35**, 1−44.

Hallam, A. 1984. Pre-Quaternary sea-level changes. *Annual Review of Earth and Planetary Sciences* **12**, 205−243.

Hallam, A. 1987a. Radiations and extinction in relation to environmental change in the marine Lower Jurassic of northwest Europe. *Paleobiology* **13**, 152−168.

Hallam, A. 1987b. End-Cretaceous mass extinction event: argument for terrestrial causation. *Science* **238**, 1237−1242.

Haq, B.U., Hardenbol, J. & Vail, P.R. 1987. Chronology of fluctuating sea levels since the Triassic. *Science* **235**, 1158−1167.

Jablonski, D. 1986. Causes and consequences of mass extinctions. In: D.K. Elliott (ed.) *Dynamics of extinction*, pp. 183−229. Wiley, New York.

Loper, D.E. & McCartney, K. 1986. Mantle plumes and the periodicity of magnetic field reversals. *Geophysical Research Letters* **13**, 1525−1528.

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

Raup, D.M. & Jablonski, D. (eds) 1986. *Patterns and processes in the history of life*. Report of Dahlem Workshop, 1985. Springer-Verlag, Berlin, Heidelberg.

Stanley, S.M. 1984. Marine mass extinction: a dominant role for temperature. In: M.H. Nitecki (ed.) *Extinctions*, pp. 69−117. University of Chicago Press, Chicago.

Stanley, S.M. 1987. *Extinction*. Scientific American Books, New York.

Wilde, P. & Berry, W.B.N. 1984. Destabilisation of the oceanic density structure and its significance to marine extinction events. *Palaeogeography, Palaeoclimatology, Palaeoecology* **48**, 143−162.

Wolfe, J.A. & Upchurch, G.R. 1987. North American non-marine climates and vegetation during the late Cretaceous. *Palaeogeography, Palaeoclimatology, Palaeoecology* **61**, 33−78.

2.12.2 Extra-terrestrial Causes

D. JABLONSKI

Introduction

Extra-terrestrial causes have long been invoked for mass extinctions, but only in the past decade has the general scientific community taken the idea seriously. Geochemical, sedimentary, and other signals in the stratigraphic record are sufficient to suggest that it is impossible to ignore extra-terrestrial impacts as potential explanations for the biotic crises that punctuate the fossil record. The case is not fully proven for any single mass extinction, although it is strongest for the end-Cretaceous event (W. Alvarez 1986; L.W. Alvarez 1987; see Hallam 1987 and Officer *et al.* 1987 for different views; see also Sections 2.12.1, 2.13.6, 2.13.7). In any event, the initial discovery of iridium and other geochemical anomalies at the Cretaceous—Tertiary boundary has sparked an immense amount of interdisciplinary research on the problem of mass extinctions and potential extra-terrestrial forcing agents.

Potential mechanisms

Proposed extra-terrestrial causes for mass extinctions have included variation in solar heat output, massive solar flares, sudden influx of cosmic rays owing to a nearby supernova or the Solar System's crossing of the Galactic plane, and collisions with comets, asteroids, or other extra-terrestrial objects

(collectively termed bolides). Until recently such factors were at best subject to only the weakest verification based on approximate correlations in timing, and at worst simply reflections of desperation in the face of seemingly inexplicable biotic upheavals. New lines of evidence for possible bolide impacts at one, and perhaps as many as five, extinction events have shifted these speculations into the realm of testability.

Earth-crossing asteroids (asteroids whose orbits cross that of the Earth or could cross as a result of long-range gravitational perturbations) are sufficiently common that significant bolide impacts must have occurred in the geological past. The Earth should suffer impacts by *c*. six 1 km asteroids per million years, and by *c*. two asteroids of 10 km or more per 100 million years, i.e. about a dozen large impacts since the beginning of the Phanerozoic (Shoemaker 1984). Effects of 1 km objects are uncertain but, as discussed below, most workers believe that impact by a 10 km bolide would have severe, global consequences.

The average collision rate for comets is almost certainly lower than that for asteroids. Cometary impact rates could occasionally be raised, however, by perturbing the Oort cloud of comets that surrounds the Solar System far beyond the outermost planets (inner edge about 10^4 Astronomical Units (AU) from the Sun, where 1 AU is the distance from the Sun to the Earth). Passage through the higher stellar densities in the spiral arms of the Galaxy might raise collision rates by about 10% (Shoemaker 1984). This low-frequency modulation of cometary impacts would be punctuated approximately once per 100 million years by short-lived bursts (1−3 million years) triggered by close passage of individual stars (Hut *et al*. 1987).

Evidence for periodic extinctions, still hotly debated, suggests (but does not prove) a more regular and frequent perturbation of the Oort cloud. Hypothesized mechanisms include: oscillations around the Galactic plane, where encounters with stars and molecular clouds would be most probable; a tenth planet in a highly eccentric orbit beyond Pluto (at *c*. 100 AU); and a dim solar companion star, christened Nemesis in advance of discovery (at distances variously estimated in the order of 10^4−10^5 AU). Debates on the astronomical plausibility of these mechanisms, with Nemesis maintaining a slight edge, are reviewed by Shoemaker & Wolfe (1986) and Hut *et al*. (1987) (see also Section 2.12.3).

The magnitude and geographical scale of an impact's effects depend on bolide size and velocity but thresholds have not been determined. An asteroid 10 km in diameter was estimated for the end-Cretaceous event on the basis of global iridium levels, and although potential effects are still poorly understood they would probably have been severe. W. Alvarez (1986), L.W. Alvarez (1987) and Prinn & Fegley (1987) emphasize the following possibilities:
1 *Darkness* caused by the global cloud of fine dust particles generated by the impact. For 2−11 months, this darkness may have been sufficiently profound to halt photosynthesis, thereby causing the collapse of marine and terrestrial food chains.
2 *Cold* would accompany the darkness, with temperatures dropping below freezing in continental interiors. Maritime climates would be less severely perturbed, owing to the thermal inertia of oceanic waters.
3 *Greenhouse effects* and global warming could follow the cold-temperature excursion if the bolide(s) struck in the ocean. After dust grains coagulated and settled from the atmosphere, the remaining burden of water vapour could trap infrared energy reflected from the Earth and raise global temperatures by as much as 10°C. The duration of this greenhouse episode is uncertain, with estimates ranging from months or years to much longer spans than the immediate cold, dark aftermath — perhaps as long as 1000 years (Prinn & Fegley 1987).
4 *Nitric acid rain* might result from shock heating of the Earth's atmosphere during impact (see Prinn & Fegley 1987, whose calculations are followed here). Energy from atmospheric entry and, especially, the supersonic plume ejected upon impact would produce very large amounts of nitric oxides. These compounds would undergo a series of reactions and ultimately rain out as nitric and nitrous acid. On land this would severely damage foliage (and, presumably, animals) both directly and through mobilization of trace metals. In the ocean, within a decade or less, the acid rain could lower the pH of the mixed layer (especially the upper 30 m) to 7.5−7.8, sufficient to dissolve calcite and thus severely stress calcareous organisms. Further, injection of so much strong acid into the atmosphere would elicit a significant exhalation of oceanic CO_2, which, combined with the accumulation of CO_2 in the atmosphere owing to depressed activity of marine phytoplankton, would yield greenhouse warming over thousands of years.

This impressive menu of impact-driven perturbations could be expected to cause mass extinctions of the observed magnitudes. Indeed, a

number of palaeontologists have argued that the hypothesized perturbations are too severe for the observed extinctions, even at the Cretaceous—Tertiary boundary (e.g. Hallam 1987). However, the impact-effect models are very poorly constrained and require extrapolation far beyond hard observational data; a new generation of more realistic and sophisticated models may provide an improved basis for critically comparing hypothesized causes with observed extinction patterns.

Biological evidence

The initial impetus for seeking extra-terrestrial impacts was of course the biological pattern of extinction in the fossil record, whether perceived as peaks in global extinction rates or as disappearances of taxa or biomass in local sections. Unfortunately, the biological consequences of impacts, massive volcanism, and other alternatives are not sufficiently understood or sufficiently unique to provide critical tests. Complex biological upheavals enacted on scales of months, years or decades, as postulated by impact scenarios, are extremely difficult, often impossible, to resolve in single stratigraphic sections, and challenge the limits of global correlation. Short-term events are superimposed on more protracted patterns in the expansion and contraction of taxa, owing to Earth-bound physical and biotic factors, so that the effect of a given boundary event on a particular taxon (particularly a waning one) is debatable. At present, the strongest constraints that palaeontological data can provide involve consistency between a given mechanism and the biological pattern observed in an imperfect fossil record.

Onset and aftermath. For extra-terrestrial impacts, biological responses include abrupt onset of extinction with an extremely brief crisis period, and a relatively short-lived reorganization and rebound during return to pre-impact conditions. In end-Cretaceous impact models, for example, most environmental perturbations would last only 1—10 years, an interval impossible to correlate among distant localities, and within which events are virtually unresolvable in the geological record. Geologically abrupt onset of mass extinction is a requirement but not a unique prediction of impact hypotheses: even such gradual processes as marine regression or transgression could in principle carry threshold effects that would produce sudden extinction pulses on stratigraphically-resolvable timescales.

Hypothesized greenhouse warming, and possibly other palaeoceanographic anomalies, would persist for some thousands of years beyond the impact itself. Some palaeontological (and geochemical) evidence supports a geologically brief — but ecologically protracted — recovery period, particularly in terrestrial plants (reviewed by Wolfe 1987) and marine plankton (reviewed by Zachos & Arthur 1986), although, again, these would not be unique to extra-terrestrial events.

Extinction patterns observed at critical boundaries cannot be taken at face value. Seemingly abrupt extinction can result from erosion or non-deposition of sediments during the critical time interval, so that biological events are compressed into single beds. At the same time, artificially gradational extinction patterns result when sampling deteriorates, or is simply uneven, in the interval approaching the boundary (a phenomenon termed backwards-smearing, or the Signor—Lipps effect — see Jablonski 1986a; Raup 1987).

Stepwise patterns of extinction, with pulses of extinction arrayed around a mass extinction boundary, have been claimed to reconcile the requirements of abrupt extinction with observations that seemed to suggest gradual loss of taxa. Such stepwise patterns — with up to 12 discrete extinction events claimed near the Cretaceous—Tertiary boundary — are also taken as the geologically rapid succession of extinction events expected during cometary bombardment. These stepwise patterns are distinct from prolonged patterns of decline such as suggested for Late Cretaceous ammonites, and are recorded near the Cenomanian—Turonian, Cretaceous—Tertiary, and Eocene—Oligocene boundaries (Hut *et al.* 1987). Unfortunately, such patterns cannot yet be taken at face value, because they can also be generated by sampling effects, local ecological changes, and/or minor breaks in sedimentation imposed on *either* abrupt or gradational extinction.

Lazarus taxa (which seem to suffer extinction but then reappear later in the stratigraphic record; Jablonski 1986a; Raup 1987) provide one means of partially controlling for unevenness in sampling and preservation: the proportion of Lazarus taxa, i.e. of observed last appearances that represent artificial extinction, permits a rough quantitative assessment of the reliability of extinction data within and around critical time intervals. Most stepwise extinction sequences contain some Lazarus taxa, suggesting that sampling effects are indeed a factor. More rigorous and comprehensive approaches are

required to place confidence limits on bed-by-bed extinction patterns.

Detailed studies of critical time intervals are urgently needed, but the plea for more centimetre by centimetre sampling near extinction events is somewhat misguided. At that scale, local ecological effects, the vagaries of sampling, and even bioturbation are likely to overwhelm the fine structure of global events. Careful sampling of relatively long geological sequences that encompass extinction events would be especially valuable, so that absences as well as presences could be recorded throughout, to provide some statistical control. Consistency of extinction patterns among widely separated localities also should be sought in a critical fashion; caution is necessary, particularly for apparent stepwise patterns, because different taxa — say, ammonites and benthic gastropods — have different sampling characteristics, even on broad geographical and temporal scales (see Jablonski 1986a on the biology of Lazarus taxa).

Selectivity has been claimed for most mass extinctions: large-bodied taxa, reef-dwellers or tropical organisms in general, and endemic taxa all appear to suffer preferential extinction (Jablonski 1986a, b). Critics (and some supporters!) of impact hypotheses have claimed that impact-driven extinction would be random rather than selective, so that any observed taxonomic or ecological selectivity would be contrary evidence. This claim seems inappropriate, however: taxa differ in their vulnerability to environmental change, so that any given perturbation, regardless of scale, should affect some groups more severely than others. Survivorship of widespread taxa, non-tropical taxa, small-bodied taxa, members of detrital food chains, freshwater taxa, deciduous plants, and plankton whose life cycles include resting cysts, has been claimed for the end-Cretaceous extinction (Jablonski 1986a, b; Hallam 1987). All are consistent with, but not exclusive to, impact hypotheses. Similarly, the possibility that mass extinctions are qualitatively different from background extinctions in their victims (e.g. see Jablonski 1986a, b) does not require impact events — any perturbation of sufficient magnitude could, for example, cross a threshold of extinction effects so that broad geographical range could determine survivorship but species richness was no longer important.

Periodicity. The apparent periodicity of post-Palaeozoic extinction events has sparked much re-search and speculation on extra-terrestrial forcing factors (see also Section 2.12.3). The periodicity itself, however, is not an adequate test for extra-terrestrial causes, although few alternatives have been advanced (Hallam 1987 reviewed a hypothesis of endogenous periodicity in mantle plumes; see also Section 2.12.1). Clearly, the critical role for palaeontological data in testing for extra-terrestrial causes of mass extinctions lies in the degree of correspondence between biological events and independent physico-chemical evidence for impacts or other extra-terrestrial forcing mechanisms. As discussed below, however, assembling such evidence is not as straightforward as was once hoped.

Physical evidence

Several physico-chemical phenomena have been proposed as independent evidence for extra-terrestrial impact. Although each has its critics, and some may not be strictly diagnostic, taken together the data make a strong case for the end-Cretaceous and Late Eocene extinctions, with weaker but suggestive evidence for several other post-Palaeozoic events (Raup 1987). The strongest Earth-bound alternative at this time appears to be volcanism (Hallam 1987; Officer *et al.* 1987; see also Section 2.12.1).

Geochemical. The anomalously high concentrations in Cretaceous−Tertiary boundary sediments of iridium, and other elements scarce in the Earth's crust but abundant in asteroids, launched the Alvarez hypothesis that an end-Cretaceous impact caused the mass extinction. Since 1979 this anomaly has been found at over 75 localities world-wide (Fig. 1) in deep-sea, shallow-marine, and continental palaeoenvironments, usually in a distinctive clay layer that coincides (within stratigraphic uncertainty limits) with the extinction event (W. Alvarez 1986; L.W. Alvarez 1987). Excursions in oxygen and carbon isotopes near the boundary also suggest a low-productivity episode that may have lasted 1.0 million years or more, accompanied by detectable but unexceptional temperature oscillations (Zachos *et al.* 1989). The direction of the stable isotopic fluctuations is appropriate to impact hypotheses, but the duration seems too long and the temperature changes too mild (but see above discussions on uncertainties in impact models and limits in stratigraphic resolution).

None of the other four major mass extinctions of the Phanerozoic has such strong geochemical
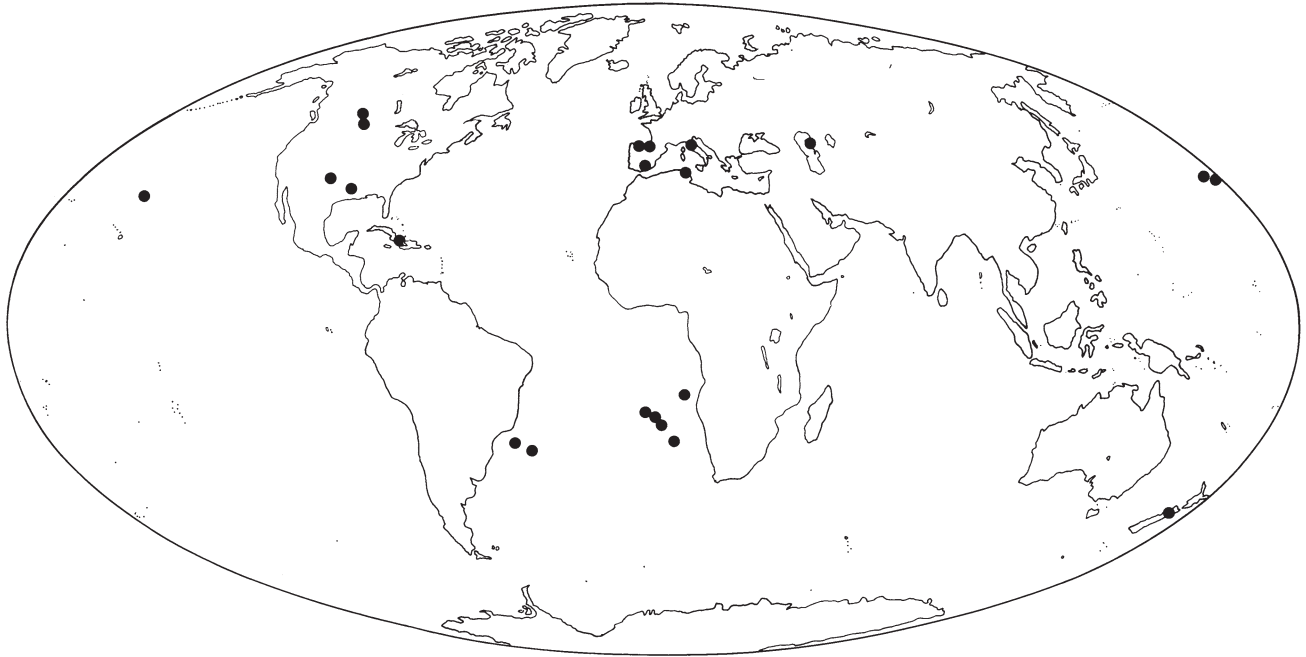
Fig. 1　Global distribution of iridium anomalies in Cretaceous–Tertiary sediments. (After L.W. Alvarez 1987.)

anomalies known from so many localities, although far less effort has been devoted to the search (Jablonski 1986a; Donovan 1987a; Raup 1987). Slight end-Ordovician iridium enrichments seem to be terrestrial in origin; the end-Triassic results are negative so far; the reported end-Permian anomaly, at the largest mass extinction of them all, has not been repeated by other laboratories, and the boundary clays seem volcanic in origin; the Late Devonian (Frasnian–Famennian) anomaly occurs in an unusual stromatolitic deposit and has not been replicated in other boundary sections.

Among lesser extinction events, iridium anomalies are geographically widespread near the Eocene–Oligocene extinction boundary, along with a series of microtektite horizons whose impact origin is virtually uncontested (Hut *et al*. 1987). An iridium anomaly was recently discovered (L.W. Alvarez 1987) for the small Middle Miocene extinction that forms the most recent peak in periodicity analyses, although the global extent of the iridium is as yet unknown. The Cenomanian–Turonian boundary has excess iridium, but other impact signatures are lacking and a terrestrial origin may be involved. An anomaly at the Middle–Upper Jurassic boundary — where no extinction event occurs but is predicted by periodicity models — occurs (like the Frasnian–Famennian example) in stromatolitic sediments, raising the spectre of biological or dia-

genetic concentration. Age uncertainty of an iridium anomaly in a 2–3 mm iron-rich crust at an unconformity in the Southern Alps overlaps with another weak or 'missing' (i.e. predicted by periodicity models) extinction peak in the Bajocian (Rocchia *et al*. 1986). An iridium anomaly, with other cosmic debris, is recorded from Late Pliocene sediments in the Southern Ocean, coinciding in time but not in space with a regional extinction event in the North Atlantic. The situation is further complicated by an anomaly near the base of the Cambrian, at a level lacking mass extinction and well after the beginning of the Cambrian radiation of skeletonized organisms (Donovan 1987b).

The degree to which all of these iridium anomalies denote impacts is still debated (Hallam 1987; Officer *et al*. 1987; Section 2.12.1). Iridium enrichments may extend for metres around the Cretaceous–Tertiary boundary in some key sections; the significance of these new observations is unclear, with interpretations ranging from diagenetic mobilization from an impact-fallout layer to prolonged deposition from volcanic aerosols. An aerosol from the Hawaiian volcano Kilauea was highly enriched in iridium, apparently derived from the deep mantle; however, other elements in the aerosol do not mimic the extra-terrestrial abundances in end-Cretaceous boundary sequences (W. Alvarez 1986) so that, again, the significance of these data is uncer-

tain. Boundary clay compositions do not always correspond to extra-terrestrial elemental abundances and isotope ratios, and yield conflicting evidence regarding the nature of the hypothesized bolide. It is not clear whether post-impact diagenetic overprint or multiple impacts by bolides of different compositions (expected in cometary bombardment?) can account for such inconsistencies. New analytical techniques (L.W. Alvarez 1987) will permit much more extensive stratigraphic coverage, both at extinction boundaries and at quiet times in between, and thus greatly improve understanding of the global iridium flux and potential nonextraterrestrial enrichment mechanisms.

Mineralogical. Potential independent evidence for impact comes from shock-metamorphosed quartz and other sedimentary particles. Like iridium, quartz grains with at least two and up to nine intersecting sets of shock lamellae have been found in Cretaceous—Tertiary boundary sequences throughout the world, in both marine and continental settings (Fig. 2) (Bohor *et al.* 1987a; Izett 1987). Such multiple lamellae are known only in particles from nuclear testing sites and impact craters. Shock-metamorphosed minerals do form near certain explosive volcanic eruptions (Hallam 1987), but the multiple lamellae and the world-wide distribution of the relatively large grains (0.1–0.2 mm in North Pacific and New Zealand sediments, up to 0.6 mm in North America) are difficult to reconcile with volcanic activity (W. Alvarez 1986; Bohor *et al.* 1987a). The search for shock-metamorphosed minerals at other extinction events has been negative so far, except for an intriguing preliminary report near the Triassic—Jurassic boundary in Austria (Badjukov *et al.* 1987).

Sedimentological. Microtektites (glassy droplets formed by bolide impacts) are almost undoubtedly present at three horizons near the Eocene—Oligocene boundary (Hut *et al.* 1987). A similar origin has been suggested for spherules of disordered potassium-feldspar (sanidine), glauconite, goethite, and magnetite found world-wide in Cretaceous—Tertiary sequences (W. Alvarez 1986), but recent evidence suggests an authigenic, non-impact origin for at least some spherules (Hallam 1987; Izett 1987). Microspherules of varying composition occur in Permo-Triassic boundary sediments in Sichuan, China (Gao *et al.* 1987); their significance is uncertain in the light of the seemingly volcanic origin of the boundary clays in China.
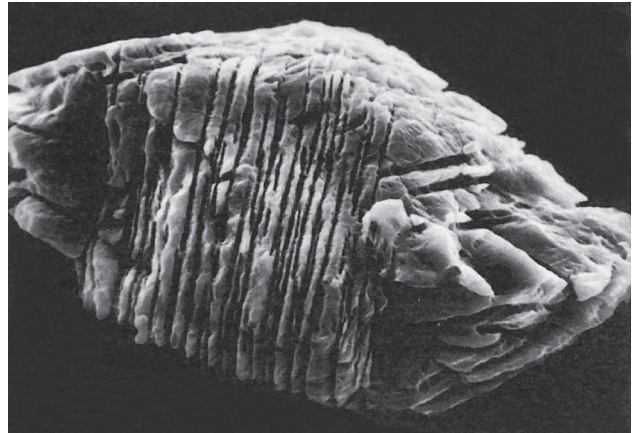

Fig. 2 Shocked quartz grain from Cretaceous—Tertiary boundary clay in a non-marine section at Brownie Butte, Garfield County, Montana. Scanning electron micrograph, width of field 0.14 mm. (Courtesy of B.F. Bohor.)

More work is needed in separating spherules of different origins before interpretations are possible (Bohor *et al.* 1987b).

Soot particles are abundant in Danish and New Zealand Cretaceous—Tertiary boundary clays (W. Alvarez 1986; L.W. Alvarez 1987). If these clays represent only one year of deposition, as postulated by most impact models, the carbon flux would have been 10^3–10^4 above background levels, suggesting extensive wildfires triggered by the heat of impact or propagated among the remains of forests killed by the hypothesized post-impact cold interval. However, the uniqueness of such soot occurrences is uncertain, and the high flux depends on the duration of clay layer deposition, which is still debated (Hallam 1987).

Cratering. Major impacts should leave craters at least an order of magnitude larger than the bolide itself. Age uncertainties are troublesome and the data are extremely sparse, but the association between extinction events over the past 250 million years and the 26 well dated craters of 5 km or more in diameter may be statistically significant (reviewed by Shoemaker and Wolfe 1986, who are sceptical). Simulations by Trefil & Raup (1987) suggest that this cratering record comprises about one-third periodic impacts (presumably comet showers) and two-thirds random collisions with asteroids. Shoemaker & Wolfe (1986) reach a similar conclusion by different means.

Questions emerge about the best-studied extinction event, however. The only well dated craters of appropriate size near the Cretaceous—Tertiary

boundary are in the U.S.S.R. (Shoemaker & Wolfe 1986), but the size and density of shocked quartz grains suggests an impact in North America (Bohor *et al*. 1987a; Izett 1987). Further, the shocked quartz suggests an impact in sedimentary rocks, i.e. a continental or shallow-water setting, whereas magnetite and other spherules suggest altered basalt, and thus an oceanic impact (although impact derivation of the spherules is now questioned, as noted above). These contradictions are perhaps resolvable with an end-Cretaceous comet shower and the consequent multiple impact, but the problem of impact site(s) remains (Hallam 1987). A volcanic interpretation is no more satisfactory in this regard.

Conclusion and prospects

Although no one indicator is definitive, at present the diverse physical and chemical evidence at the Cretaceous−Tertiary boundary is most readily explained by a bolide impact. Volcanism is the chief rival, but as W. Alvarez (1986) argued, evidently only quiet basaltic eruptions yield iridium aerosols and melt microspherules, whereas violent siliceous eruptions are needed to produce shocked minerals. Neither kind of eruption will produce all of the observed impact signatures, nor can either account for the world-wide distribution of shocked quartz, iridium and other geochemical anomalies. The periodic mantle plume hypothesis might yield both explosive and non-explosive volcanism on a global scale (Hallam 1987; Section 2.12.1), but this model awaits evaluation. The palaeontological data are generally consistent with, but provide little conclusive support for, impact-driven extinction mechanisms. As many authors have noted, marine regression at this and other extinction events obscures biological and physico-chemical signals and may even play a role in extinction (Section 2.12.1).

The most definitive evidence for or against extra-terrestrial factors in mass extinctions (apart from the discovery of the hypothesized solar companion, Nemesis) will come with an assessment of the strength of temporal association between Phanerozoic mass extinctions and physico-chemical signatures of bolide impacts. This work is under way, and it is impressive that the three or four most recent extinction peaks recognized in global data sets and/or local stratigraphic sections (Middle Miocene, Eocene−Oligocene, Cretaceous−Tertiary, and Cenomanian−Turonian) bear at least some impact indicators. The weak but significant clustering of crater ages at extinction events over the past 250 million years should prompt analyses around other boundaries, with ongoing refinement of hypotheses. Assessment of negative evidence remains a problem, however, so that impact hypotheses can be remarkably elastic and difficult to falsify: absence of craters, shocked quartz and even iridium anomalies are consistent with impact on now-subducted ocean, basaltic impact site, and cometary rather than meteorite impact, respectively. Additionally, not all major craters, microtektite horizons, or iridium anomalies coincide with extinction events. Better understanding of the potential effects of impacts, and of the distribution of potential impact signatures through the stratigraphic record, should lead to the framing of more refined hypotheses regarding the role of extra-terrestrial factors in the evolution of life on Earth.

References

Alvarez, L.W. 1987. Mass extinctions caused by large bolide impacts. *Physics Today* **40**, 24−33.

Alvarez, W., 1986. Toward a theory of impact crises. *Eos* **67**, 649, 653−655, 658.

Badjukov, D.D., Lobitzer, H. & Nazarov, M.A. 1987. Quartz grains with planar features in the Triassic−Jurassic boundary sediments from Northern Limestone Alps, Austria. *Lunar and Planetary Science* **18**, 38−39.

Bohor, B.F., Modreski, P.J. & Foord, E.E. 1987a. Shocked quartz in the Cretaceous−Tertiary boundary clays: evidence for a global distribution. *Science* **236**, 705−709 (see also 666−668).

Bohor, B.F., Triplehorn, D.M., Nichols, D.J. & Millard, H.T., Jr. 1987b. Dinosaurs, spherules, and the 'magic' layer: a new K−T boundary clay site in Wyoming. *Geology* **15**, 896−899.

Donovan, S.K. 1987a. Iridium anomalous no longer? *Nature* **326**, 331−332.

Donovan, S.K. 1987b. Confusion at the boundary. *Nature* **329**, 288.

Gao Zhengang, Xu Daoyi, Zhang Qinwen & Sun Yiyin 1987. Discovery and study of microspherules at the Permian−Triassic boundary of the Shangsi section, Guangyuan, Sichuan. *Geological Review* **33**, 203−211 (in Chinese with English abstract).

Hallam, A. 1987. End-Cretaceous mass extinction event: argument for terrestrial causation. *Science* **238**, 1237−1242.

Hut, P., Alvarez, W., Elder, W.P., Hansen, T., Kauffman, E.G., Keller, G., Shoemaker, E.M. & Weissman, P.R. 1987. Comet showers as a cause of mass extinctions. *Nature* **329**, 118−126.

Izett, G.A. 1987. Authigenic 'spherules' in K-T boundary sediments at Caravaca, Spain and Raton Basin, Colorado and New Mexico, may not be impact derived. *Bulletin of the Geological Society of America* **99**, 78−86.

Jablonski, D. 1986a. Causes and consequences of mass extinctions: a comparative approach. In: D.K. Elliott (ed.)