What can we learn about ecology and evolution from the fossil record?

Jeremy B.C. Jackson\textsuperscript{1,2} and Douglas H. Erwin\textsuperscript{3}

\textsuperscript{1}Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093-0244, USA
\textsuperscript{2}Smithsonian Tropical Research Institute, Box 2072 Balboa, Republic of Panama
\textsuperscript{3}Department of Paleobiology, MRC-121, National Museum of Natural History, Washington DC, 20013-7012, USA

The increased application of abundance data embedded within a more detailed and precise environmental context is enabling paleontologists to explore more rigorously the dynamics and underlying processes of ecological and evolutionary change in deep time. Several recent findings are of special theoretical interest. Community membership is commonly more stable and persistent than expected by chance, even in the face of the extreme environmental changes of the Ice Ages, and major evolutionary novelties commonly lie dormant for tens of millions of years before the ecological explosions of the clades that possess them. As we discuss here, questions such as these cannot be adequately addressed without the use of the fossil record.

Introduction

‘Our ability to reconstruct past communities and the history of life hinges on thinking like biologists doing field work in ancient environments.’ [1]

The fossil record is a uniquely rich source of information about the patterns of past biological diversity: the waxing and waning of various clades and their component species, of novel morphologies and of ecological communities unlike those alive today. From rigorously documented patterns, paleontologists can also infer much about basic ecological and evolutionary processes that cannot be revealed by the study of living organisms alone. Some of the best examples of the synthesis of paleobiological pattern and process include the influence of Quaternary [since 1.8 million years ago (Mya)] climate change on the distribution and composition of temperate North American plant communities [2], the causes and consequences of the end-Cretaceous mass extinction (65.5 Mya) [3] and the Paleocene–Eocene thermal maximum (55.8 Mya) [4], and the environmental setting of the spectacular radiation of animals during the Ediacaran–Early Cambrian (575–509 Mya) [5].

In each case, paleontologists have worked with other geologists to obtain extensive new collections of fossils tied to detailed and independent records of changes in climate, oceanography, tectonic events and other aspects of the physical–chemical environment through geochemical proxies. Such abiotic forcing drives many ecological and evolutionary processes, so careful attention to reconstructing environmental scenarios independent of the biological record is essential. The examples that we discuss here reveal how ecological and evolutionary processes can be reconstructed from fossil data. These new insights provide a deep-time perspective to longstanding and fundamental ecological and evolutionary questions for which biological data are inadequate, and have inspired paleontologists to address more rigorously, and with greater focus, the roles of intrinsic versus extrinsic drivers of evolutionary change, a question that Stephen Jay Gould described as one of the ‘eternal metaphors’ in understanding the history of life [6].

Here, we discuss three fundamental ecological and evolutionary questions for which breakthroughs based on the fossil record are occurring because of a more holistic paleontological perspective, the formulation of testable hypotheses about process, and new programs of field work and data collection specifically designed to test them. Testability and an emphasis on process have long been hallmarks of research in other areas of ecology and evolution and must be more widely adopted by paleobiologists to ensure continued progress toward our understanding of the history of life. Ecologists developed an approach to experimental design and testability that became more rigorous during the 1960s in response to what many believed was an overly descriptive approach and has set the standard for most empirical ecology ever since. However, with the advent of macroecology and the need to confront pressing problems of human impacts, ecologists are once again increasingly faced with deducing process from pattern without recourse to experimental manipulation, the situation that paleontologists have always faced. But paleontologists still have much to learn from ecologists about the nature of sampling and the data required to test alternative hypotheses about the processes driving biodiversity [7,10,54]. Until recently, most paleontological data collected was used to document new species, to understand the geographical, environmental and temporal extent of existing species, and to document and understand ecological relationships. New hypotheses commonly require new data and, as paleontologists more rigorously define hypotheses and better understand how to test them, they will have to better...
define the data required. This will often necessitate the acquisition of extensive new field collections.

Most of our examples in what follows here reflect our greater familiarity with the marine fossil record, although we have endeavored to also include terrestrial examples. Conspicuously absent from our discussion, however, is any analysis of the longstanding debate about the ups and downs of the diversity of life throughout the Phanerozoic (since 530 Mya). This is because the data so far used are so poor [7,8] that we still do not know whether diversity was similar between the Paleozoic (543–250 Mya) and Cenozoic (since 65.5 Mya) [9] or increased several fold [10]. We believe that this question can be answered but will require intensive fieldwork and systematic sampling rather than reliance upon previously published studies [7].

Are ecological communities stable through time and space?

Ecologists have long debated the degree to which ecological communities are either chance associations of species with overlapping environmental niches or more limited in membership owing to interactions among species and the details of their life histories [11–19]. In the first case, community composition should vary more or less continuously with changes in the environment and the vagaries of recruitment from adjacent areas, local extinction and speciation [11,18,19]. But if membership is more limited, communities should exhibit statistical differences in species composition among habitats that persist in space and time [12,13,16,17].

Community composition over spatial scales of a few meters or less typically fluctuates widely in response to local disturbances every few years [20]. By contrast, observations over scales of hundreds of meters to thousands of kilometers reveal a more predictable and persistent pattern of patches of different community composition (i.e. successional states) that result from the interactions of stochastic environmental fluctuations and differences among the life histories, resource use and vulnerability of species to disturbance [21–23]. However, the extreme brevity of most ecological observations begs the question of the persistence of community composition over longer timescales that encompass overlapping generations and greater environmental change [24]. For most organisms other than microbes, these questions can only be addressed using the fossil record, although new questions arise regarding the comparability of recent and fossil communities in terms of the preservation of species, the adequacy of sampling for rarer species and the amount of time encompassed by a single paleontological horizon [25,26]. Fortunately, numerous studies have shown that quantitatively well sampled assemblages of fossil species usually provide a reliable picture of past community composition, if only for organisms with preservable hard parts (Box 1).

Comparing taxonomic occurrences over time

Most studies of the persistence of fossil communities are based upon comparisons of taxonomic occurrences from similar paleoenvironments over time [24]. Some of the most detailed studies of terrestrial plant communities are for Early Carboniferous (c. 340 Mya) wetland plants preserved in coal deposits. Preservation is generally excellent, enabling detailed quantitative census of leaf litter and root systems as well as of spores, and temporal resolution is commonly on the order of 100–1000 years. Assemblages from the same environments exhibit striking similarity in rank-order abundance of species over periods of hundreds of thousands to several million years [27,28]. Persistence was terminated by comparatively brief episodes of major extinction. Similar patterns of apparent community stability over several million years punctuated by comparatively short episodes of taxonomic turnover, dubbed ‘coordinated stasis,’ were described for marine benthos from Paleozoic continental shelf environments [29]. However, these analyses used only presence–absence data and temporal resolution was coarse. Quantitative analyses of the same faunas using abundance data reveal more variable and statistically different taxonomic membership of local assemblages than were originally described, although several of the most common taxa do persist throughout the entire interval [30,31].

Comparing patterns of association

Most other studies of community persistence concern patterns of association among species during the environmental changes associated with the waxing and waning of the Arctic ice sheets of the Late Pleistocene and Holocene [since 126 thousand years ago (Kya)]. Assemblages of insects [32,33] and temperate [2,19] and tropical [34,35] plants varied greatly in species composition over periods of 1000 years or less and individual species behaved differently, and apparently independently, as their environments changed. This might have also been true for Late Pleistocene to recent molluscs along the Pacific coast of North America [36], but the data for recent species are for ranges rather than samples, so we cannot judge
whether the absence of a species reflects either real limits to distribution or sampling bias. In all these examples, climates very different from today (nonanalog) resulted in nonanalog communities with no apparent long-term associations of species.

Problems of interpretation

The problem with these studies is twofold. First, what does it mean for communities to be similar or different in taxonomic composition, and where do we draw the line [37]? With sufficient effort, almost any two assemblages can be shown to differ statistically from one another. The more interesting ecological question is whether differences in community membership reflect deterministic processes of community assembly and maintenance or just stochastic processes. Distinguishing between these requires testing whether similarities in community composition are significantly greater than the null model expectation of a random sampling of the available within-habitat species pool [18,37,38]. The second problem concerns the relationship between changes in community composition and the environment. It is hardly surprising that plant community composition changed radically when the massive North American ice sheets that blanketed the landscape retreated by thousands of kilometers in a few thousand years, or when the tropics were transformed from warm and wet to cool and dry. Rather, the crucial issue is the persistence of species associations during periods of comparative environmental stability and the thresholds of environmental change before community composition responds.

The only studies to address these questions systematically, of which we are aware, are for Pleistocene coral reef and North American small mammal communities. Pandolfi [39] compared the community composition of reef crest and reef slope coral communities using presence–absence data from transect samples of nine different reef terrace horizons at three sites ~15 km apart along the coast of the Huon peninsula in Papua New Guinea. The reef terraces are a product of tectonic uplift and glacial sea-level fluctuations over 95 Ka. A new reef formed after each episode of uplift or sea-level rise so that each new reef represents an independent natural experiment in the larval recruitment and assembly of a new coral reef community. Adequacy of sampling was assessed by collector’s curves. There were no significant differences in community composition or species richness among the nine terraces at any one site, but there were persistent differences between sites along the coast, reflecting differences in runoff from a nearby river. Moreover, community membership at each site was significantly more persistent than was expected from a random sample of the habitat-specific species pool, and communities were consistently dominated by <25% of the total 66–92 species encountered in the reef crest and reef slope environments, respectively. Comparable spatial and temporal persistence in Caribbean shallow-water coral community composition over the past 125 Ka has been observed elsewhere based on different methods [40–42] (Figure 1).

Exciting new work on Ice Age small mammal communities from North America [37] demonstrates similarity among sites in species composition far greater than expected by the null hypothesis of neutral theory [18] across 3500 km and nearly 1 My. What is truly remarkable about this and the coral reef studies is that community membership should be so persistent during the exceptionally large environmental fluctuations associated with the waxing and waning of ice sheets on a massive scale, a point we return to in regard to environmental drivers of evolutionary change.

Ecological dimensions of evolutionary novelty

Understanding major evolutionary transitions and the introduction of evolutionary novelty is an increasingly compelling research area in paleontology, particularly as the advances in comparative studies of development have provided new insights into the mechanistic basis of morphological novelties. Novelty has several distinct components: the developmental introduction of morphological novelties, fixation of these within a species, a subsequent increase in the number of taxa (species or more inclusive clades) containing the novelty and an increase in the number of individuals (or is a corresponding ecological impact. Although these components are frequently conflated, novelties are not necessarily successful in generating increases in either taxic diversity or ecological abundance. Similarly, not all technological inventions are economically successful, and economists and historians of technology have distinguished between inventions and their subsequent impact as economic innovations [43]. Examples from the spread of flowering plants and grasses illustrate that there is often a macroevolutionary lag between the first appearance of a group (the invention) and its subsequent rise to ecological significance (the innovation).

Example of the angiosperms

The origin and spread of flowering plants (angiosperms) illustrates the complexity of these issues. Today, flowering plants comprise >95% of total plant diversity and
Paleontologists long believed that they spread rapidly after their earliest appearance in the Early Cretaceous (c. 135 Mya) [44,45]. However, the discovery of a latest Cretaceous flora at Big Cedar Ridge in Wyoming [46] showed that the situation is more complex. Angiosperms comprised 61% of total species but, because the flora was exquisitely preserved beneath a volcanic ash, it represents a single community on a single day and thus also enables an ecological census. Despite tens of millions of years of prior diversification, angiosperms accounted for only 12% of the vegetational cover, and a single species of palm accounted for much of this abundance. In a broader study, Lupia and colleagues [47] used the fossil record of pollen and spores to track the evolutionary patterns of angiosperms, ferns and their allies, and gymnosperms through the Cretaceous of North America. Flowering plants increased in diversity and abundance through the Cretaceous, evidently displacing free-sporing plants but not conifers. Nevertheless, increases in angiosperm abundance appear to have lagged behind increases in diversity by ~10 million years.

Examples of the grasses and acropoid corals
The expansion of grasses across North America during the Cenozoic (since 65.5 Mya) exhibits a macroevolutionary lag similar to that of Cretaceous angiosperms. Although grasses had diversified taxonomically by the Oligocene (34 Mya) [48], they did not become ecologically significant until the Late Oligocene (c. 25 Mya) or Early Miocene (c. 20 Mya), probably as a consequence of climate change. Likewise, the ecological explosion of reef corals of the genus Acropora during the Pleistocene lagged several million years behind the origins of the clades and species that subsequently rose to dominance, an ecological transition that was probably triggered by the increased magnitude of sea-level fluctuations during the past 1 My [49]. All these examples reinforce the need to embed evolutionary studies in their environmental context, as well as the crucial role that quantitative abundance data can have in establishing the controls on evolutionary novelty. In the absence of such data, paleontologists can easily, but misleadingly, confuse time of origin or diversification with ecological impact [50].

Comparative developmental studies
Comparative developmental studies provide another example of macroevolutionary lags. It is now clear that much of the developmental toolkit required to build bilaterian animals (all animals except sponges and cnidarians) was present in members of the Phylum Cnidaria (sea anemones and allies) [51] well before the appearance and diversification of complex Bilateria. Although the formation of the developmental mechanisms needed to produce the complex body plans of arthropods, molluscs and chordates was necessary for the Cambrian radiation (530–510 Mya), it does not appear to have been sufficient for the radiation, and ecological dynamics were the probable driving force [5,52].

Evolution is often viewed as highly opportunistic, immediately taking advantage of new adaptations and possibilities. But the fossil record does not provide strong support for the common assumption that taxonomic diversification and ecological abundance are correlated, or for a close correlation between major developmental novelties and either taxonomic diversification or ecological impact. If this is generally true, then extaptations [53] have great evolutionary significance because they provide the raw material for later evolutionary advances, even though their initial appearance might have occurred for reasons unrelated to the subsequent success of the clade having them.

Environmental drivers of evolutionary change
There is a longstanding, unresolved debate about the relative importance of biotic interactions versus changes in the environment (Red Queen versus Court Jester hypotheses) as the principal drivers of mammalian evolution ([54] and extensive references therein) that hinges largely on the spatial and temporal scale and quantitative rigor of the data [54,55] and whether it was collected with specific hypotheses in mind. The only studies that we are aware of to take all these factors into account come down strongly in favor of climate change as the primary evolutionary force [54,55]. It is also hard to think of a major macroevolutionary event in the oceans or involving terrestrial floras that was not somehow linked to environmental change. Most paleontologists and paleoceanographers take this almost as a matter of faith and the circumstantial evidence seems overwhelming [5,56]. Nevertheless, with the exception of extreme cases, such as extinction associated with asteroid impact [3,5], it is often difficult to establish close stratigraphic correspondence rigorously between environmental change and the hypothesized biological response, and to establish independent evidence to infer cause and effect based on a clear biological mechanism, such as the differential response of taxa to global warming [3,5], a collapse in primary productivity [57–59], or widespread anoxia in the oceans [60].

We have seen that community composition can be highly dynamic or persistent over tens to hundreds of thousands of years in the face of the great climatic fluctuations of the past million years, with the best evidence for persistence coming from tropical coral reefs where environmental variability was much less than in temperate regions, and from North American small mammals, where it was not [19,37,39]. In none of these examples, however, were changes in community composition associated with a burst of extinction or origination of species [61,62]. Exceptions include the mass extinction of megafauna closely tied worldwide with the arrival of humans, possibly combined with climate change that notably had no comparable effects before our arrival [63]. Similarly, in the case of modern human disturbance of the oceans, major changes in community composition owing to overfishing, habitat destruction, warming and declines in water quality have altered community composition beyond recognition and wreaked havoc with ecosystem structure and function [64,65]. Yet despite the ecological extinction of many formerly superabundant species, few of the easiest species to monitor have yet to go extinct. This delay between the onset of environmental change and
extinction reflects the inherently nonlinear dynamics and thresholds in response to perturbations in ecological time [66–68], with obvious implications for the origins and extinctions of species.

Examples of climate change and ecosystem response
One of the best examples of a tight link between climate change and ecosystem response other than the past 20 Ky is the response of North American plant communities to rapid global warming at the Paleocene–Eocene boundary 55.8 Mya [4]. Temperatures rose by 5–10°C within 10–20 Ky, a rate comparable to warming after the last glacial maximum 12 Kya. Mass extinction was observed for benthic foraminifera in the oceans, but most marine and terrestrial animals and land plants underwent rapid, long-distance shifts in distribution rather than extinction, a pattern strikingly similar to that of northern hemisphere plant and animal communities over the past 20 Ky [2,19].

By contrast, the response of tropical marine communities to the oceanographic events associated with the rise of the Isthmus of Panama 3–5 Mya includes strong macroevolutionary as well as ecological responses [57–59,61]. However, there appears to have been a time lag of 1–2 My between the initial ecological responses to environmental change and the subsequent mass extinction of marine benthos, including reef corals and molluscs. The separation of the western Atlantic from the eastern Pacific caused large-scale collapse in coastal upwelling and primary productivity [57–59] and corresponding increases in carbonate deposition throughout the tropical western Atlantic (69, A. Odea, unpublished data). These environmental changes were spread out over several million years while the Isthmian seaway gradually closed, but the greatest changes occurred between 4 Mya and 3 Mya (A. Odea, unpublished data). Faunal analyses also show gradual shifts in community composition of total biota (A. Odea, unpublished data) as well as molluscs [61,70] and reef corals [71], including an increase in the extent and size of coral reefs [69].

Throughout this drawn-out period of environmental change, the trophic structure of some 800 genera of molluscs also changed. There was a significant decline in abundance of suspension-feeding bivalves and predatory gastropods compared with other molluscs, and reef dwellers became more common [58]. However, these shifts were invisible without abundance data, and diversity of molluscs and corals increased despite a gradual increase in extinction rates [61]. Then, within a few hundred thousand years, there was a mass extinction of 20% of molluscan genera and 40% of reef coral species [58,61] (Figure 2). Which taxa suffered most was predictable from the changes in environments during the preceding 2 Ma, strongly suggesting a threshold response in extinction comparable to the collapse of modern ecosystems owing to human disturbance, but at a grander macroevolutionary scale (A. Odea, unpublished data).

Conclusions and perspectives
Studies of the fossil record and the history of life have gone through several phases. Initially, paleontology was concerned with discovery of new fossils and using the most widely distributed to establish the relative sequence of events of the geological timescale by biostratigraphy. Discovery, systematics and biostratigraphy continued to be the central focus of many paleontologists into the middle part of the 20th century, and fossils also had an important role in reconstructing ancient environments. George Gaylord Simpson’s great synthesis of genetics, evolution and paleontology in Tempo and Mode in Evolution [72] formed the core of modern paleontology, with fossils as representatives of once-living organisms (previous generations of paleontologists had a regrettable habit of viewing fossils as inanimate blobs with limited spatial and temporal ranges).

With the advent of paleobiology during the 1960s, paleontologists began to address the evolutionary and ecological processes underlying the patterns documented by the fossil record. Much attention has been given over the past several decades to understanding the biases of preservation and the rock record, development of highly sophisticated quantitative tools for understanding diversity patterns, and the use of fossils in phylogenetic reconstruction. With this robust toolkit now available, exciting new challenges are at hand, which will require gathering primary data using consistent protocols and in conjunction with meaningful information on environmental context. The time has come to move to a more rigorous engagement with the ecological and evolutionary processes that define and structure the history of life.
Acknowledgements
We thank Ken Johnson, Aaron O’Dea, Travis Smith and Jon Todd for helpful discussions of many of the ideas in this article, and Tony Barnosky, James Crompton, Andrew Read and Andrew Smith for their helpful reviews.

References
12 Elton, C. (1927) Animal Ecology, Sidgwick and Jackson
20 Connell, J.H. et al. (1997) A 30-year study of coral abundance, recruitment, and disturbance at several scales in space and time. Ecol. Monogr. 67, 461–488
28 DiMichele, W.A. et al. (2002) Place vs. time and vegetation persistence: a comparison of four tropical paleoimires from the Illinois Basin at the height of the Pennsylvanian ice age. Int. J. Coal Geol. 50, 43–72
60 Knoll, A.H. et al. (1996) Comparative earth history and Late Permian mass extinction. Science 273, 452–457