INTRODUCTION

Given that 99.9% of species that have ever existed on the Earth are already extinct, the most effective approach to examining extinction over long time scales is to delve into the 3.7 billion year history of extinction preserved in the fossil record. The “discovery” of extinction, both as a philosophical concept and a biological reality, can be traced back to Baron Georges Cuvier and his studies of fossil mammoth and mastodon anatomy in the late 18th century (Rudwick, 1998). We have come a long way since then and much of this progress has taken place in the past 30 years. Two breakthroughs, both published in the early 1980s, launched a scientific movement that was recently dubbed “the extinction industry” (Bambach, 2006). Walter Alvarez and colleagues’ (1980) hypothesis that an extra-terrestrial event was responsible for the end-Cretaceous (K/T) mass extinction triggered a firestorm of controversy and sparked intense interest in catastrophic extinction in the fossil record. The identification and designation of the “Big 5” mass extinctions by Raup and Sepkoski (1982), based on Sepkoski’s compilation of first and last occurrence data of Phanerozoic marine families (e.g., Sepkoski, 1982), solidified this interest. In the past decade, the Paleobiology Database (www.paleodb.org) has built on Sepkoski’s (1982; 2002) and other early databasing efforts (e.g., Newell, 1952; Benton, 1993) and is already providing a foundation for future studies of Phanerozoic diversity and extinction.

The “extinction industry” has flourished, exploding from a handful of papers published in the 1950’s, when Newell (1952; 1963; 1967) first coined the term “mass extinction,” to hundreds in the new millennium (Bambach, 2006)(Fig. 1). The impact of these studies, which reflect the wealth of long-term and large-scale data available in the fossil record, has extended beyond paleontological circles to influence such diverse fields as evolutionary biology, conservation, ecology, sedimentology, geomorphology, astronomy, and physics.

MAJOR THEMES FROM PAST WORK

Rather than describe in detail what we know about the rather lengthy catalogue of Phanerozoic extinction...
events (for recent reviews see Walliser, 1996; Hallam and Wignall, 1997; Jablonski, 2004, 2005; Bambach, 2006), I will provide a short overview of some of the major themes that have emerged from the past thirty years of extinction research, then discuss a range of future research directions. Themes of past work run the gamut from basic descriptions of extinction magnitude, duration, and selectivity, which permeate the literature, to more complex attempts to examine cause and effect. I will focus almost exclusively on the record of marine invertebrates, but the majority of the patterns I highlight and the techniques I describe are equally applicable to other organisms and ecosystems.

What goes extinct?
Differentiating victims from survivors is relatively straightforward, although sometimes complicated by the existence of Lazarus taxa (i.e., taxa that disappear from the record and are thought to go extinct only to reappear later; Jablonski, 1986a) and Elvis taxa (i.e., post-extinction taxa that closely resemble pre-extinction taxa, but are not actually related to them; Erwin and Droser, 1993). Although Lazarus taxa may comprise up to 74% of the survivors for a given event (Erwin, 1996a; but see Nützel, 2005), a variety of methods, including multi-regional approaches to sampling and metrics for assessing the completeness of the record, can be used to identify these taxa and determine why they occur (Wignall and Benton, 1999; Fara, 2001; Rickards and Wright, 2002).

Differentiating victims and survivors
Selectivity, defined as the identification of traits that distinguish victims from survivors, represents a major theme of past work. Selectivity can clarify causal mechanisms, identify which taxa are most vulnerable to extinction in the present day, and act as an important mechanism of evolutionary change across these events. The traits most commonly examined for selectivity are ecological in nature, including body size, geographic range, trophic strategy, and life habit, although species richness and reproductive mode have
also received considerable attention (for recent reviews see McKinney, 1997; Jablonski, 2005). Fewer studies have explicitly tackled morphological, phylogenetic, or physiological aspects of selectivity, although this is certainly changing (e.g., Bambach et al., 2002; Heard and Mooers, 2002; Lockwood, 2004; McGowan, 2004b; Liow, 2006; Knoll et al., 2007; Saunders et al., 2008). Patterns of selectivity vary considerably across extinction events (and across clades across a single extinction event), but a handful of patterns have emerged. Strong evidence suggests that geographic range can be an important predictor of survival, particularly at higher taxonomic levels (for review see Jablonski, 1995; 2005). Similarly, a depositfeeding lifestyle is often hypothesized to promote survivorship (Sheehan et al., 1996; Smith and Jeffery, 1998), although this pattern may be attributable to taxonomic sorting in some cases (i.e., a trait is correlated with survivorship because the clade that possessed it survived, not because the trait itself influenced survivorship; Jablonski and Raup, 1995). Other important traits, such as body size, have yielded disparate and conflicting results (for review see McKinney, 2001) and emphasize the need for meta-analytical approaches to synthesizing these data across clades and events.

Some workers have suggested that traits that promote survivorship during background extinction may not promote survivorship during mass extinction (for K/T see Jablonski, 1986a; Kitchell et al., 1986; Jablonski, 1989; Jablonski and Raup, 1995; Lockwood, 2003) and that the alternation of these selectivity regimes can have a profound impact on evolutionary trends (Gould and Calloway, 1980; Gould, 1985; Jablonski, 1986b; Gould, 2002). In reality, few studies have tested explicitly whether mass extinctions strengthen, weaken, or have no effect on the long-term adaptation of biota. We are in desperate need of studies that take advantage of the repeated nature of extinctions in the fossil record and compare selectivity (during both the extinction and recovery intervals) across a range of events that differ according to magnitude (e.g., Johnson et al., 1995; Smith and Roy, 2006), duration, and causal mechanism.

Quantifying extinction

Methods for assessing extinction magnitude and rate abound, from classic per-taxon rate measures (for review see Foote, 1994), to metrics that count boundary-crossing taxa (Bambach, 1999; Foote, 2000), to clade-based approaches relying on taxonomic survivorship analyses (Foote, 1988), all of which incorporate particular assumptions about extinction processes. Studies applying these methods range from large-scale compilations of global extinction rates throughout the Phanerozoic (e.g., Bambach et al., 2004) to regional analyses (e.g., Patzkowsky and Holland, 1996) to cohort survivorship analyses of a single clade (e.g., Foote, 1988). The results are difficult to compare directly to estimates for the modern diversity crisis, but several studies suggest that we may be experiencing the sixth mass extinction and that the subsequent recovery of diversity will take millions of years (Leakey and Lewin, 1996; Gaston and Spicer, 1998; Myers and Knoll, 2001; Wilson, 2002).

Quantifying extinction magnitude inevitably leads to the question of which events are largest. After three decades, we are still struggling to statistically differentiate mass from background extinctions. Different approaches yield different results, from designation of the “Big Five” extinctions (Raup and Sepkoski, 1982), to the “Big Three” (Bambach et al., 2004), to a continuous distribution of extinction rates (Raup, 1991; Wang, 2003). Wang’s (2003) recent discussion of mass versus background events reminds us that, while we have thoroughly compared the magnitudes of these events, we have surprisingly little comparative information on selectivity (Sheehan, 2001; Lockwood, 2005), effects (Sheehan et al., 1996; Hansen et al., 2004; McGhee et al., 2004), recovery patterns (Hansen et al., 2004), or causal mechanisms.

The above discussion focuses almost exclusively on aspects of patterns of taxonomic extinction (i.e., the loss of species or higher taxonomic units). These events can also be explored using other diversity proxies, including morphological (Foote, 1996, 1999; Hansen et al., 1999; Lupia, 1999; McGowan, 2004a), ecological (Bambach et al., 2002; McGhee et al., 2004), and phylogenetic (McGowan and Smith, 2007) measures of diversity. These measures should be tracked across these events, both the extinction and recovery intervals, as the interplay among them will no doubt provide us with important information regarding the long-term evolutionary consequences of extinction.
Timing of extinction

Turning to the temporal aspects of extinction, high-resolution dating has dramatically altered estimates of extinction duration and rates throughout the Phanerozoic (see methods in Harries, 2003; Gradstein et al., 2004; Erwin, 2006a). For example, detailed examination of Permo-Triassic (P/T) boundary sections from southern China, coupled with dating of numerous ash layers, has substantially decreased our approximations of extinction duration, from estimates of several million years in the 1990s (Erwin, 1993) to current estimates of 10-100,000 years (Jin et al., 2000; Rampino et al., 2000; Twitchett et al., 2001). As these techniques are applied to a wider range of events, radical revision of the timing of extinction events will no doubt continue, with major implications for both the causes of extinctions and their evolutionary effects.

Stratigraphic confidence intervals provide a quantitative means for assessing the Signor-Lipps effect, i.e., the backward “smearing” of last occurrences, due to the incompleteness of the fossil record, that can make a sudden extinction appear gradual (Signor and Lipps, 1982). These techniques, which also evolved independently in conservation biology, can help to determine whether the timing of extinctions was gradual, sudden, and/or stepwise. For some events, the application of confidence intervals has revealed a multi-phased event (KT: Marshall and Ward, 1996) or a more abrupt extinction than was originally recognized (P/T: Jin et al., 2000). For others, it has simply highlighted the need for more closely spaced abundance and occurrence-level data across extinction intervals. Recent improvements to these techniques (Marshall, 1995; Wagner, 1995b; Holland, 2003; Wang and Marshall, 2004) take into account facies shifts and variable sampling within lineages, and make it possible to calculate the statistical likelihood of extinctions of particular magnitudes or timing.

One of the most intriguing temporal patterns to emerge from the fossil record is the apparent periodicity of Phanerozoic extinctions. Cycles of extinction have been postulated for some time (e.g., Newell, 1952; Fischer and Arthur, 1977), but were not fully explored until Raup and Sepkoski (1984; 1986) applied harmonic analysis to the post-Paleozoic record of extinctions and found evidence of a 26 million year periodicity. Although this work has been criticized on the basis of taxonomic and geochronological weaknesses and possible causes for the proposed periodicity are not at all obvious, the patterns remain difficult to ignore (Rohde and Muller, 2005; Lieberman and Melott, 2007). While periodicity is the most-debated secular pattern of extinction in the fossil record, the decline in background extinction rates recorded for both marine families and genera throughout the Phanerozoic is the best-documented one (Raup and Sepkoski, 1982; Van Valen, 1984; Gilinsky, 1994). This pattern cannot be explained simply by improvement in the quality of the fossil record through time (Foote, 2000) and may occur because higher turnover rates in the early history of clades lead to higher background rates in the early Phanerozoic (Sepkoski, 1984; Foote, 1988; Valentine, 1990).

Spatial aspects of extinction

Extinction studies tend to focus on either the global or local scale, with few studies bridging the gap of regional patterns or explicitly considering the effects of bias due to uneven spatial sampling. The spatial aspects of extinction can help us to untangle the relative importance of emigration versus extinction and immigration versus origination, which in turn may have implications for predicting how these processes might work in modern ecosystems (Stigall and Lieberman, 2006). Studies that do parse out global patterns into regional ones find intriguing differences in extinction magnitude, selectivity, and/or rates of recovery (Raymond et al., 1990; Kelley and Raymond, 1991; Jablonski, 1998; Shen and Shi, 2002; Wignall and Newton, 2003; Krug and Patzkowsky, 2004, 2007). For example, Jablonski’s (1998) differentiation of the K/T recovery by region suggested that the rapid diversification of “bloom taxa” (i.e., taxa that exhibit an evolutionary burst of diversification during the recovery) was restricted to the North American Gulf Coast (see also Hansen, 1988), calling into question the role that these taxa play in global recovery patterns.

Causal mechanisms

Our understanding of how and why extinctions occur has, by necessity, lagged behind our attempts to simply describe and quantify events. For the majority of extinctions, the laundry list of possible causes is long and the evidence contradictory (for P/T review see Bambach, 2006; Erwin, 2006b). Causal hypotheses are frequently hampered by extreme approaches,
which postulate either a single cause or a combination of every possible cause for each event (Jablonski, 1980; MacLeod, 1998). It is entirely likely that multiple causes are responsible for these events and only careful hypothesis testing, coupled with fine-scale dating, geochemical tools, and selectivity analyses (see for example Knoll et al., 2007), will help us to differentiate the importance and timing of different mechanisms. Although the stated goal of many selectivity analyses is to identify possible causal mechanisms, few studies have proven successful, in part due to the dearth of explicit hypotheses generated to differentiate among causes, no doubt coupled with few opportunities to integrate vital paleoenvironmental and paleoclimate data.

Recovery
Turning to recovery intervals [i.e., post-extinction intervals characterized by rapid rebound of diversity (Erwin, 2001)], we are still in need of the most basic data when it comes to recoveries, including: (1) rates, (2) durations (for a start see table 1 in Erwin, 1998), and (3) the effects of a range of biases on these patterns. Detailed descriptions of these intervals have shed some light on which taxa are participating in the repopulation and to what extent (Harries et al., 1996). Repeated mentions of bloom and/or opportunistic taxa emphasize the need for a quantitative and/or phylogenetic approach to identifying the key players in recovery, from “disaster forms” (i.e., defined as simple, cosmopolitan, opportunistic generalists: Schubert and Bottjer, 1995) to examples of “dead clade walking” (e.g., defined as survivors that do not participate in the post-extinction diversification: Jablonski, 2002). We need more outcrop-scale studies, combined with regional analyses and phylogenetic tracking (e.g., Rode and Lieberman, 2005; McGowan and Smith, 2007) across both extinction and recovery intervals, to properly differentiate survivors from taxa originating in the post-extinction melee.

PRESERVATION, SAMPLING, AND OTHER FACTORS

In the past decade, our understanding of the extent to which factors such as preservation and sampling affect global estimates of extinction and diversity in the fossil record has grown by leaps and bounds. To generalize, better preservation and greater sampling yield greater sample diversity, which becomes a problem when we try to reconstruct large scale diversity and extinction patterns in the fossil record. To consider the best- and worst-case scenarios, perfect preservation and sampling produce 100% accurate stratigraphic ranges and therefore 100% accurate extinction rates. In contrast, extremely poor preservation and sampling convert all taxa into singletons (i.e., taxa that are restricted to a single time interval) and extinction rates become meaningless. In reality, preservation and sampling lie somewhere between these two scenarios, but they illustrate the importance of estimating preservation rate and standardizing sampling concurrently with extinction rate (Foote and Raup, 1996; Foote, 1997). Peters andAusich (2008) provide a useful framework for categorizing the numerous factors that distort the macroevolutionary record. They subdivide them into intrinsic (i.e., those that are intrinsic to the scientific process of accumulating knowledge, including incomplete sampling and taxonomic errors) versus extrinsic ones (i.e., those that are inherent to the geological record itself, such as rock availability, sequence architecture, and taphonomic factors).

To consider intrinsic factors first, variable sampling (i.e., differences in sample sizes across time intervals) can massively distort estimates of diversity (see Alroy et al., 2001; Bush et al., 2004, among many others) and must be taken into account before any attempt is made to quantify extinction. Sample standardization is often accomplished by applying boundary-crosser metrics (Bambach, 1999; Foote, 2000), removing singletons (Sepkoski, 1996; Foote, 2000), or perhaps more effectively via resampling routines (Alroy et al., 2001) such as rarefaction. It should be noted that sample standardization, which adjusts for variable as opposed to incomplete sampling, does not eliminate the problems associated with the Signor-Lipps effect (see below).

The impact of another intrinsic factor—taxonomic standardization—on global compilations of diversity has been debated for years (Smith and Patterson, 1988; Sepkoski and Kendrick, 1993; Wagner, 1995a; Adrain and Westrop, 2000; Robeck et al., 2000; Ausich and Peters, 2005), but it is unclear whether taxonomic revision minimizes or accentuates peaks in extinction rate. Recently, Wagner et al. (2007) fo-
focused on three molluscan case studies throughout the Phanerozoic and found that taxonomic standardization elevated extinction rates in at least one of their case studies (Paleozoic gastropods). They argued that an overabundance of polyphyletic taxa, coupled with high species richness in paraphyletic taxa, acts to artificially diminish the magnitude of extinction events (see also Uhen, 1996). In contrast, Ausich and Peters (2005) found that substantial revision of crinoid taxonomy yielded significantly lower extinction rates in the Late Ordovician. These results highlight the influence of phylogenetic topology on extinction rates, and emphasize the importance of explicitly controlling for taxonomic bias.

Turning to extrinsic factors, several studies have revealed that the available rock record is strongly correlated with diversity curve shape; decreases in outcrop availability artificially inflate estimates of extinction intensity (Raup, 1972; Peters and Foote, 2001; Smith, 2001; Peters and Foote, 2002; Crampton et al., 2003; Foote, 2003). A number of other extrinsic factors that potentially affect the “completeness” of the record (and hence extinction rates) have been identified, including secular patterns in sequence architecture (Holland, 1995; Smith, 2001; Holland, 2003) and taphonomic and diagenetic factors (Cherns and Wright, 2000; Kidwell and Holland, 2002; Wright et al., 2003; Kowalewski et al., 2006); however, few studies have explicitly quantified the influence of these factors (for example see Bush and Bambach, 2004).

Another extrinsic artifact, “Pull of the Recent,” occurs when increased sampling of the Recent biota extends stratigraphic ranges of fossil taxa, artificially decreasing estimates of extinction towards the present day. Recent analysis of extant and Plio-Pleistocene bivalve subgenera revealed that only 5% of Cenozoic diversity could be explained by “Pull of the Recent” (Jablonski et al., 2003). This result suggests that this particular artifact may not be as much of a problem as originally thought, although analysis of other clades is certainly warranted (see also Foote, 2000).

One useful approach to measuring a combination of extrinsic factors, such the completeness of the fossil record (among many, for review see Foote, 2001), takes advantage of the expected inverse relationship between the number of singletons and preservation quality, and only requires data on the first and last occurrences of taxa (Foote and Raup, 1996; Foote, 1997). This technique, FreqRat (Foote and Raup, 1996), estimates preservation probability per unit time interval given the frequency of taxa with stratigraphic ranges of 1 (i.e., singletons), 2, and 3 intervals. This, in turn, makes it possible to determine whether survivorship among taxa can be explained simply by preservational patterns (Foote and Raup, 1996). Incomplete sampling can also be taken into account by analyzing gaps and occurrences and placing confidence intervals on the end of taxon ranges (Foote, 2001).

A handful of recent studies have attempted to control for both intrinsic and extrinsic factors while calculating extinction rates. For example, Foote (2007) explicitly considered the effect of “back-smearing” caused by incomplete sampling in his reconstruction of Phanerozoic extinction patterns for global marine genera. Controlling for both incomplete and variable sampling yielded an extremely volatile pattern of extinction, with long intervals of quiescence punctuated by even more massive extinctions than originally recognized. At a smaller taxonomic and stratigraphic scale, Peters and Ausich (2008) used a taxonomically- and stratigraphically revised dataset on Paleozoic crinoids to quantify the effects of both variable and incomplete sampling on extinction rates from the Ordovician to early Silurian. They obtained a result similar to that of Foote (2007), namely a substantial increase in the volatility of extinction rates. They also described a positive correlation between extinction (but not origination) and metrics of sedimentary completeness, which they interpreted as evidence that the former may have been controlled by physical environmental changes.

FUTURE DIRECTIONS

As estimates of the rate and magnitude of modern extinction increase, the fossil record becomes fertile ground for predicting the evolutionary effects of the modern biodiversity crisis. The two key advantages of the fossil record—long time scales and large perturbations—allow paleontologists to make important contributions to understanding the role that extinction plays in the history of life. I would argue that, given the fundamental biodiversity and climate changes our world is currently facing, no other question in paleontology is as important (for several examples see NRC, 2005).
The history of life has a sample size of one and as a result, paleontology is generally considered a historical, as opposed to experimental, science. This view, along with the seemingly overwhelming number of differences among extinctions, has limited our ability to systematically identify and synthesize patterns across the Phanerozoic. Although the analogy is far from perfect, viewing extinctions as repeated natural experiments in the history of life would allow us to identify common features characterizing and processes underlying these events. The analogy can be extended further by considering several of the aspects of extinction that are already well-constrained—such as magnitude, duration, tectonic configuration, and climate—as controls for these natural experiments. For example, to target questions concerning the evolutionary effects of extinction according to magnitude, one could select 3-4 extinctions along a gradient of magnitude, all associated with broadly similar causal mechanisms, such as climate change. A plethora of new and newly refined tools, including CONOP-9 [i.e., Constrained Optimization (Kemple et al., 1995)], biomarkers, and GIS [i.e., Geographic Information Systems (Graham et al., 1996; Rode and Lieberman, 2004; Stigall and Lieberman, 2006)], allow us to take a much more holistic approach to these natural experiments. Testing of cause and effect hypotheses requires interdisciplinary, integrated methods, merging disparate fields of paleoclimatology, paleoenvironmental reconstruction, oceanography, evolutionary biology, development, and tectonics (NRC, 2005; Jackson and Erwin, 2006). The call for such an approach is obviously not new [e.g., Geobiology of Critical Intervals (GOCI) and more recently Deep Earth-Time-Life Observatories (DETELOS)] and, in situations in which this multidisciplinary approach has been applied, it has been extremely successful. For example, by mapping diversity data onto probable centers of orogeny for the Ordovician radiation, Miller and Mao (1995) uncovered a relationship between orogenic activity and diversification (see also Miller, 1997, 1998). The time has come to synthesize knowledge of extinction events, across time intervals, disciplines, and ecosystems, in an explicit framework of repeated natural experiments and testable hypotheses. It is in this spirit that I outline three broad areas—selectivity, recovery, and spatial aspects of extinctions—for future research that will greatly aid in elucidating the evolutionary effects of extinction.

**Effects of selectivity**

Extinctions can contribute to evolution by eliminating dominant taxa and allowing subordinate taxa to diversify, they can redirect evolutionary trends by eliminating important innovations, and they can limit the potential evolution of clades by reducing variability. Many of these mechanisms operate via selectivity and one of the major goals of selectivity studies is to clarify how this works. The current approach to selectivity has not changed substantially in the last three decades, but a number of recent studies suggest that broadening the traits examined and developing more effective techniques for identifying selectivity can yield important insights into the macroevolutionary effects of extinctions.

The majority of selectivity studies focus on the mean or dominant traits of a taxon, such as average body size or predominant life habit. This approach ignores the fundamental importance of trait variation, the existence of which is a prerequisite for evolutionary change (Lloyd and Gould, 1993). The breadth of traits observed within a given taxon, such as morphological variance within a species or multiple feeding modes within a genus, is thought to indicate ecological and/or evolutionary potential and may prove just as important as its geographic “breadth” in determining survivorship. A handful of recent studies have explored the important, yet largely unrecognized, role that trait variation can play in extinction (Kolbe et al., 2006; Smith and Roy, 2006). For example, Kolbe et al. (2006) examined selectivity of morphological variation in veneroid bivalves across the Plio-Pleistocene extinction in Florida, focusing on fourteen pairs of closely related species, in which one survived the extinction and the other did not. When they assessed selectivity, while controlling for phylogenetic bias, allometry, and ecophenotypic plasticity (i.e., variation in shape due to variation in environment), they found that morphological variation in veneroid bivalve species was a significantly stronger predictor of survivorship than shell shape, body size, or sampled geographic range (Fig. 2). Veneroid species with higher levels of morphological variation were more likely to survive this extinction event. Although this is the only study I am aware of that has directly assessed
the link between morphological variation and extinction, two recent studies suggest that variation may also influence pathways and rates of diversification. Hunt (2007) empirically examined the interactions among morphological variation and evolutionary divergence in the ostracode genus *Poseidonamicus* and found that evolutionary changes tended to occur in directions of high phenotypic variation within the genus. Similarly, Webster (2007) determined that the incidence of polymorphic characters (i.e., characters which span two or more states of variation in a given taxon) was higher in stratigraphically older and/or phylogenetically basal species of Cambrian trilobites. The possible link between trait variation and extinction is intriguing and leads to a number of hypotheses that can be tested at a variety of taxonomic levels, including the possibility that taxa with more variability should be more likely to survive and more likely to recover more quickly. These hypotheses require testing across multiple clades and events, especially in light of possible implications for the conservation and management of modern biodiversity.

The traditional approach to quantifying selectivity, independently testing a handful of traits in a single clade across a single event, is slowly giving way to much more robust multivariate analyses that take nonlinear covariation among traits explicitly into account. Biological traits, regardless of whether they are life history, ecological, or morphological in nature, are inexorably linked to one another, and these linkages can make it difficult, if not impossible, to determine which traits are being selected for and which traits are simply along for the ride. Multifactorial approaches (e.g., Harnik, 2007; Payne and Finnegan, 2007), such as linear and logistic regression, path analysis, and structural equation modeling (Shipley, 2000), make it possible to identify which traits are most directly related to survivorship. In an analysis of Eocene bivalve species from the U.S. Gulf Coastal Plain, Harnik (2007) reported that, although both geographic range and body size were tied to extinction probability (and to each other), the former exerted a much stronger effect than the latter (see Jablonski, 2008b for the K/T). Similarly, Payne and Finnegan (2007) used binary lo-
logistic regression to assess extinction selectivity during background intervals in Phanerozoic marine invertebrate genera and determined that geographic range remained a significant predictor of survivorship, even after the effects of species richness and occupancy had been removed. Such approaches are long overdue and may even help us to differentiate cause from correlation when it comes to traits that promote extinction in modern taxa.

Studies of minor and background extinction have taken a back seat to mass extinction for decades. We know far too little about how selectivity varies across extinctions of different magnitudes and durations (Johnson et al., 1995; Smith and Roy, 2006). Returning to the analogy of natural experiments, comparisons could preferentially target extinctions with similar causes, magnitudes, or durations, in an effort to reveal differences in selectivity relative to other aspects of the events. A comparison of selectivity according to body size in veneroid bivalves across the K/T, end of the middle Eocene (mid-E), and end Eocene (E/O) events in North America and Europe suggested that the lower magnitude, but longer duration event was associated with statistically stronger selectivity (Lockwood, 2005). Although neither extinction was size selective, the K/T recovery was biased towards smaller veneroids, while the mid-E and E/O recoveries were biased towards larger veneroids (Fig. 3). This result raises the interesting possibility that longer term “press” extinctions, in which the extinction pressure is prolonged, may exhibit stronger selectivity and therefore exert stronger influence on evolutionary trends than short-term “pulse” extinctions (Erwin, 1996b). This interpretation is complicated, in this case, by the extremely different causal mechanisms for these events (i.e., bolide impact for the K/T and climate change for the Eocene events). Payne and Finnegan’s (2007) comparison of selectivity during background and mass extinction intervals for Phanerozoic marine invertebrate genera documented a weak, but intriguing, inverse relationship between extinction magnitude and geographic range selectivity. Selectivity for broad
geographic ranges is weaker during mass versus background extinction intervals, potentially because widespread environmental disturbance during the former simultaneously affects taxa with particular ecological or physiological traits. These studies raise more questions than they answer. Are mass extinctions simply scaled-up versions of minor extinctions? Are minor extinctions scaled up versions of background extinction? Does selectivity differ between extinction- and origination-driven declines in diversity (Bambach et al., 2004)? Is there a threshold below which the direction or strength of selectivity changes?

Finally, turning to the wealth of data already available on selectivity, another way to address these questions, and to facilitate an “experimental” approach to extinction selectivity, is via meta-analyses. Several authors have provided reviews of the direction of selectivity across a range of events and taxonomic levels (see for example McKinney, 2001; Jablonski, 2005). What is missing is a quantitative, meta-analytical approach to this bewildering and often contradictory literature. Meta-analyses have the capacity to statistically detect patterns and consistency in selectivity as well as to help us understand the overall strength of evidence for particular consequences of extinction. Meta-analytical techniques prioritize results on the basis of effect size (i.e., a measure of the magnitude of a treatment effect, Osenberg et al., 1999; Gurevitch and Hedges, 2001) and have recently proven useful in navigating the sizeable literature pertaining to live-dead studies in taphonomy (Kidwell, 2001) and species-energy relationships (Hunt et al., 2005).

**Importance of recovery intervals**

To understand the influence that mass extinctions can exert on evolutionary patterns, it is crucial to examine both the extinction events themselves and the recovery intervals that follow. Similar to extinctions, recoveries can affect the evolutionary history of a biota (Erwin, 1998, 2001). Despite a recent rise in the number of studies focusing on recoveries, we still know relatively little about recolonization and diversification during the post-extinction interval and the processes that control them. This is extremely unfortunate given the potential parallels between post-extinction recovery and restoration ecology. Examining the importance of recoveries in the macroevolutionary arena will require a refocus of selectivity studies on the post-extinction interval, explicit comparisons between radiations and recoveries, and tracking of important macroevolutionary trends across both extinction and recovery intervals.

The evolutionary impact of an extinction event is closely tied to its selectivity; however, too few studies have examined the selectivity that occurs during recovery. Failure to recover can be just as important as failure to survive (Jablonski, 2002). The prolonged duration of many recoveries, particularly relative to the extinctions themselves (Erwin, 1998, 2001), increases the likelihood that these intervals will be important to long-term macroevolutionary trends. Tracking of veneroid bivalves across the K/T mass extinction in morphospace revealed statistically significant selectivity for deeper burrowers (i.e., subgenera with deeper pallial sinuses and more elliptical shell shapes), but only during the recovery, not during the extinction itself (Fig. 4) (Lockwood, 2004). The biased origination of deeper burrowers during the recovery initiated an expansion into deeper burrowing niches that was maintained throughout the Paleogene, emphasizing the important contribution that recovery intervals can make to long-term evolutionary trends. Similarly, biased recovery was more important than selective extinction in driving shifts in veneroid body size throughout the late Mesozoic and early Cenozoic (see above, Lockwood, 2005). These results are limited to a single clade across three extinction events, but they hint at the possibility that changes in other trends across mass extinctions are actually the result of recovery as opposed to extinction dynamics. This in turn raises the interesting question of whether an alternation of selectivity regimes exists during extinction versus recovery or during recovery versus radiation.

The repeated nature of extinctions and their subsequent recoveries makes it possible to test hypotheses of phylogenetic (e.g., development or genetic) versus ecological constraint in the early evolution (or initial radiation) of clades. Perhaps the most obvious events to compare are early Paleozoic radiations (e.g., the Cambrian explosion, the Ordovician radiation) and the P/T recovery. In a classic study, Erwin et al. (1987) tracked the establishment of new phyla, classes, and orders of marine animals across these two intervals and found that the P/T recovery resulted in significantly less diversification at higher taxonomic levels. Erwin et al. (1987) attributed this asymmetry to
the different starting points for each diversification in ecospace—an empty ecospace in the early Paleozoic versus a sparsely occupied ecospace in the Mesozoic. Foote (1996) followed up on this taxonomic examination by comparing the rate at which peak morphological disparity was reached for crinoids radiating in the early Paleozoic versus those recovering after the P/T globally. He documented similar rates during the two events, although further work revealed that the early Paleozoic radiation produced a much greater range of forms (Foote, 1999). These studies provide a useful basis for comparison, but substantially more work is needed, including examination of other proxies of diversity (e.g., ecological and phylogenetic), clades, and time intervals (see Jablonski et al., 1997; Wagner, 1997 for interesting examples).

Remarkably few studies have tackled the question of how evolutionary or ecological trends, from latitudinal diversity gradients to onshore-offshore patterns of origination, shift across recovery intervals. One pattern for which we do have preliminary data involves comparative analyses of taxonomic versus morphological diversity across recovery intervals. These data suggest that the disconnect observed between taxonomic and morphological diversity during background intervals persists during post-extinction intervals. For example, comparisons of taxonomic versus morphological recovery in ammonoids across the P/T mass extinction revealed that morphological diversity (measured as variance of morphological characters) recovered slowly despite increases in taxonomic diversity, in part because of the evolution...
(1) Number of genera

- Variance
- Mean randomized value
- Singletons excluded
- Boundary crossers

(2) Morphological variance

- Variance
- Mean randomized value

Time (Ma)

250 240 230 220 210 200
exhibited by surviving taxa, may contribute to the dynamics of recovery (e.g., Stanley, 1977; Stanley, 1990, 2007). Biotic interactions or the removal of these interactions, including competition or predation, may also be influencing these patterns (for examples see Vermeij, 1987; Miller and Sepkoski, 1988; Jablonski, 2008a), but remain extremely difficult to test. Finally, extrinsic factors, such as ongoing environmental disturbance, are widely blamed for prolonged or stepwise recovery intervals, especially following the P/T (Hallam, 1991) and K/T (Coxall et al., 2006) extinctions, but few studies have considered the effects of environmental factors relative to stochastic, intrinsic, and biotic ones. Comparative analyses are hampered by the complex timing of events during recovery intervals, which often requires high resolution correlation of global environmental signals with local fossil occurrence patterns to sort out. Correlation is not causation, but the timing of “environmental” versus “diversity” recovery across multiple “experiments” of extinction may help to differentiate cause from effect (Vermeij, 2004). For example, the timing of primary productivity collapse versus environmental degradation (e.g., oxygen deprivation, hypercapnia, and methane poisoning) across extinction events will prove crucial in testing Vermeij’s (2004) hypothesis that the latter are merely effects of the former.

**Influence of spatial patterns**

As mentioned above, studies of extinction are often performed at the outcrop or global scale, despite the fact that regional patterns are thought to contribute much to the complexity of extinction and recovery dynamics (Jablonski, 1998, 2002). Too often, we attempt to synthesize general rules of extinction or recovery on the basis of well-sampled records in North America and Europe, with little appreciation for geographic variation. This is particularly unfortunate, given that different faunal responses in different regions can be used as controls in these natural experiments of extinction. Environmental factors that are important in one region may not occur in another, allowing us to test hypotheses of causal mechanisms. For example, extinctions driven by global cooling might be expected to affect tropical regions more severely than temperate regions. Several aspects of the “spatial fabric” (Jablonski, 2005) of extinctions warrant further investigation, including the role of immigration/invasion during recovery and the importance of spatial autocorrelation to both extinction and recovery metrics.

In an ideal world, paleontologists would be able to differentiate emigration from extinction and immigration from origination. Unfortunately, tracking patterns of migration (analogous to gene flow in evolutionary terms) versus in situ origination requires exceptional sampling (Rode and Lieberman, 2004; Stigall and Lieberman, 2006), preferably in combination with a phylogeny (Rode and Lieberman, 2005). The few studies that have quantified extinction and recovery patterns at the regional scale suggest a decoupling of global, regional, and local processes, in part because of the contribution of immigration to the latter two (Raymond et al., 1990; Kelley and Raymond, 1991; Jablonski, 1998; Clemens, 2002; Bowersox, 2005; Krug and Patzkowsky, 2007). Krug and Patzkowsky (2007) parsed generic recovery of brachiopods, bivalves, anthozoans, and trilobites after the Late Ordovician extinction into regional patterns for Laurentia, Baltica, and Avalonia. Sample-standardization of their occurrence data revealed much faster rates of recovery in Laurentia relative to the two other paleocontinents, in part due to a higher proportion of invading taxa in the former, suggesting that immigration played a particularly important part in the recovery (Fig. 6). In a smaller scale study, Bowersox (2005) examined the regional patterns of Pliocene molluscan extinction in the San Joaquin Basin of California. He found twice the magnitude of extinction and substantially slower recovery in the basin relative to the two other paleocontinents, in part because rapid environmental changes restricted flow, and therefore, invaders from the Pacific Ocean. Regional studies such as these, repeated across multiple events, may help us to predict which ecosystems are more or less likely to experience invasion after extinction—a prediction that has all sorts of applications to the modern biodiversity crisis (e.g., Mooney and Hobbs, 2000).

Another important contribution of spatial data pertains to the potential sampling biases associated with geographic patterns. Paleontologists are well aware of the many biases affecting our data and great strides have been made to control for many of these (see above). One potential bias that has been long recognized in ecological circles (see for example Legendre and Legendre, 1998), but little discussed in paleontological circles, is spatial autocorrelation. Spa-
Partial autocorrelation occurs when a pattern observed in one location is affected by patterns at neighboring locations (i.e., when samples are not independent of each other in space). Non-independence of samples in space is a particularly serious problem for extinction and recovery studies, which are often limited to sampling one to two disparate regions. The distribution of samples in space, when considered explicitly, can also highlight ecologically important mechanisms such as source-sink dynamics. Spatial issues in paleontology are commonly assessed by parsing patterns into regions, which often reveals which regions are contributing most strongly to global patterns. The use of “temporal” resampling (i.e., sample standardization of measures such as diversity from one time bin to the next) has become increasingly common in paleontology. A handful of paleontological studies have begun to resample patterns environmentally (e.g., carbonate versus siliciclastic substrates; Kiessling and Aberhan, 2007), but I have yet to see any attempt at “spatial” resampling, i.e., sample standardization of measures from one region to the next. This approach to resampling could prove particularly useful for studies of spatial diversity, such as latitudinal diversity gradients, but they may prove valuable for patterns of extinction as well.

**CONCLUSIONS**

Although our knowledge of the fundamentals of extinction in the fossil record has increased exponentially in the past three decades, we are still grappling with the more complex questions of causal mechanisms and consequences. To understand fully the evolutionary effects of extinction, we need to take advantage of the repeated nature of extinctions and to begin to test explicit hypotheses across these events. Selectivity is one important driver of evolutionary change across extinction events, but we need to consider a broader range of traits, the co-variation among these traits, and comparative patterns across a range of extinctions that differ according to magnitude. Patterns of diversification during recovery intervals can influence evolutionary trends just as strongly as extinction, yet we know surprisingly little about why some taxa rebound and others do not, or what factors control the rate and duration of recovery. Finally, the spatial as-
pects of extinction, including the relative importance of extinction, origination, and migration, during both extinction and recovery intervals, have received relatively little attention thus far.

ACKNOWLEDGMENTS

Ask ten different extinction workers what they think are the most important future research directions in our field, and you’ll get ten different answers. Thank you to R. Bambach and P. Kelley for inviting me to participate in this unique short course and to share my opinion. M. Foote, D. Jablonski, P. Wagner, J. Swaddle, M. Kosnik, P. Kelley, and A. Stigall provided very useful comments that greatly improved the scope and quality of the manuscript. Thank you also to S. Kolbe, Z. Krug and A. McGowan for their willingness to share figures. Acknowledgment is made to the Donors of the American Chemical Society Petroleum Research Fund and the Jeffress Memorial Trust for partial support of this work. This manuscript was developed while a Sabbatical Fellow at the National Center for Ecological Analysis and Synthesis, a Center funded by NSF (Grant #DEB-0553768), the University of California, Santa Barbara, and the State of California.

REFERENCES CITED


CIAMPAGLIO, C. N. 2004. Measuring changes in articulate brachiopod morphology before and after the Permian mass extinction event: do develop-


HALLAM, A. 1991. Why was there a delayed radiation after the end-Palaeozoic extinctions? Historical Biology, 5:257-262.


EXTINCTIONS AS EXPERIMENTS IN THE HISTORY OF LIFE


