Evolutionary Innovations in the Fossil Record: The Intersection of Ecology, Development, and Macroevolution

DAVID JABLONSKI*
Department of Geophysical Sciences, University of Chicago, Chicago, Illinois 60637

ABSTRACT The origins of evolutionary innovations have been intensively studied, but relatively little is known about their large-scale ecological patterns. For post-Paleozoic benthic marine invertebrates, which have the richest and most densely sampled fossil record, order-level taxa tend to appear first in onshore, disturbed habitats, even in groups that are now exclusively deep-water (so that present-day distributions are not reliable indicators of original environments). New results presented here show that the onshore-origination pattern is robust to shifts in taxonomic methods and to new paleontological discoveries, and the few available studies suggest that this pattern can also be seen in terms of excursions in morphospace or the acquisition of derived character states, without reference to taxonomic categories. The environmental pattern at high levels contrasts significantly with the origin of low-level novelties (such as defined genera and families) in crinoids, echinoids, and bryozoans, where first appearances tend to conform to their clade-specific bathymetric diversity gradients. This discordance seems to eliminate potential driving mechanisms that simply scale up within-population genetic or ecological processes. Little is known about the factors that promote the onshore–offshore expansion of orders across the continental shelf, or that drive some clades to abandon ancestral habitats for an exclusively deep-water distribution. The origin of evolutionary innovation must ultimately reside in developmental changes, but the onshore-origination bias could emerge from two different dynamics: the pattern could be primarily genetic and developmental, i.e., innovations truly arise onshore; or primarily ecological, i.e., innovations arise randomly but preferentially survive onshore. Whatever the ultimate driving mechanisms, these macroevolutionary patterns show that theories of large-scale evolutionary novelty must include an ecological dimension. J. Exp. Zool. (Mol. Dev. Evol.) 304B:504–519, 2005.

The past decade has seen spectacular progress in evolutionary developmental biology, but this has rarely been integrated with ecology, the other term in Van Valen’s famous dictum that evolution is “the control of development by ecology” (Van Valen, ’73). Renewed interest in ecological developmental biology, and even an “eco–evo–devo” (e.g., Gilbert, 2001; Gilbert and Bolker, 2003; Givnish, 2003), is a very welcome step towards interdisciplinary synthesis, but the eco–devo enterprise has been implicitly or explicitly microevolutionary in scope (e.g., Haag, 2003). However, many large-scale, macroevolutionary patterns apparent in the fossil record—including the origins of evolutionary innovations and major groups—also appear to have an ecological and biogeographic component. These larger patterns represent a promising basis for the integration of ecology, evolution, and developmental biology.

Much has been written about the most striking temporal pattern in the origin of major groups, with nearly all of the skeletonized animal phyla appearing in a geologically brief pulse near the beginning of the Cambrian Period. This burst of phenotypic variety is important regardless of when the major lineages diverged genealogically (e.g., Erwin, ’99; Knoll and Carroll, ’99; Valentine et al., ’99; Conway Morris, 2000; Valentine, 2002, 2004; Valentine and Jablonski, 2003), and the debate continues on developmental vs. environmental triggers for this event and its termination.

*Correspondence to: Prof. David Jablonski, Department of Geophysical Sciences, University of Chicago, 5734 South Ellis Avenue, Chicago, IL 60637. E-mail: d-jablonski@uchicago.edu

© 2005 WILEY-LISS, INC.
with environmental triggers and feedbacks in favor at present (e.g., Valentine, '95, 2004; Knoll and Carroll, '99; Ciampaglio, 2002, 2004), although definitive tests are elusive (Jablonski (2000, 2003); and see Foote ('99) for a morphological study supporting a role for intrinsic factors, and Erwin ('99) for related arguments). Less profound but distinct episodes of evolutionary innovation also occur in the wake of mass extinctions, and these secondary pulses provide valuable material for comparative analyses (see Erwin et al. ('87) and Foote ('99) for contrasting interpretations of the early Mesozoic recovery from the end-Paleozoic extinction). Bursts of morphological disparity early in the history of major groups have also been the focus of debate on the roles of developmental and ecological factors in promoting or damping innovation (e.g., Gould et al., '87; Foote, '97; Lupia, '99; McGhee, '99; Boyce and Knoll, 2002; Lofgren et al., 2003).

Here I will focus on ecological rather than temporal patterns in the origin of major novelties. First principles suggest that we should not necessarily expect any pervasive ecological patterns: if novelties arise by random mutation—relative to the pressures of the moment—and chance ecological opportunities, then evolutionary innovations should be randomly distributed among environments. For that matter, a palaeontological inventory of evolutionary novelties might simply reflect sampling biases such as the availability of fossiliferous outcrops in time and space. However, neither of these expectations is borne out empirically. The new analyses of the fossil record presented here, and some older work cited below, indicate that major evolutionary novelties in marine invertebrates do not originate randomly with respect to environment, but tend to occur first in onshore, disturbed habitats.

ANALYZING ENVIRONMENTAL HISTORIES OF MAJOR GROUPS

The simplest approach to analyzing environmental patterns in the origin of major groups might seem to be a survey of their present-day occurrences. Sampling and preservation are most complete for living organisms, and if modern distributions were accurate reflections of origination patterns the analysis would be fairly straightforward. However, the fossil record clearly falsifies this assumption. For example, several marine groups today confined to depths exceeding 100 m, including lychniscosidan sponges, isocrinid and millericrinid crinoids, and pedinoid and holasteroid echinoids, clearly first appeared in onshore settings (e.g., Smith, '84, 2004; Bottjer and Jablonski, '88; Jablonski and Bottjer, '90a-c; Smith and Stockley, 2005). As discussed below, an even larger number of marine taxa first appeared in a relatively small set of onshore habitats and subsequently expanded over the entire continental shelf, and often into slope and abyssal environments.

Because the dynamic environmental history of higher taxa rules out the use of present-day distributions as an indicator of their initial habitats, we must turn to the fossil record. The most extensive data are for marine invertebrates, and for these analyses I have focused on the post-Paleozoic fossil record of the past 250 million years (Myr) for several reasons. Temporal resolution and correlation of fossiliferous sequences among regions is more refined; the ecology of ancient environments and the phylogeny of higher taxa are more fully understood; and average fossil preservation for several important major groups tends to be better than in the Paleozoic (e.g., for taxa with skeletons consisting of the less stable forms of calcium carbonate). This paper, then, addresses the environmental origins of the present-day marine fauna, during and after recovery from the massive end-Permian mass extinction.

The ideal starting point for analyzing the environmental origins of major groups would be a fully resolved cladogram of all fossil marine invertebrate taxa. If such extensive phylogenetic data are absent, analyses will be based on an inventory of higher taxa of ordinal rank. Most of the taxa used herein have been verified as monophyletic, but because the focus of the analysis is the origin of functional or ecological novelties, a few paraphyletic groups have been retained when one subclade diverged strongly from the primitive condition. For example, the first occurrence of both the stalked crinoid order Isocrinida and the stalkless crinoids, the Comatulida or feather stars, is of interest even though separating out the comatulids probably renders the Isocrinida paraphyletic.

A rigorous analysis of origination patterns in the fossil record has a number of requirements, in addition to a robust taxonomic or phylogenetic framework, and a review of these protocols is worthwhile.
Criteria for environmental inference that are independent of the presence/absence of particular taxa

Fortunately, the rich literature relating sedimentary and stratigraphic features in the geological record to environmental conditions provides a firm foundation for such an approach (e.g., Reading, '96; Leeder, '99; Dunbar and Barrett, 2005). For the present purposes, the environmental histories of major groups are mapped into a simple five-bin transect of marine benthic habitats, subdivided by geologic evidence for water-energy levels during the deposition of fossil-bearing deposits (see Bottjer and Jablonski ('88) for details). This model has two onshore, shallow-water categories: a nearshore environment always subject to wave action, and an inner-shelf environment below average wavebase but frequently disturbed by storms. Offshore, deep-water settings comprise the middle shelf, which is only rarely affected by major storms, the outer shelf, with no evidence of storm influence, and slope settings beyond the edge of the continental shelf (for a very similar six-bin model for Paleozoic environments, see Sepkoski and Miller, '85). Thus for our purposes, average storm wavebase—the water depth where storms regularly stir up the seafloor—is taken to be the dividing line between onshore and offshore settings, and between frequently disturbed and more stable environments. Although this is a dynamic and fuzzy boundary, most sedimentary deposits are readily placed within these broad categories. When environments cannot be relegated to a single category, stratigraphic resolution for a given fossil is too low to determine its point of origin within an environmentally heterogeneous stratigraphic sequence, or apparently contemporaneous fossils occur in more than one category, first occurrences are split among those categories in the statistical analyses.

A database on the first (geologically oldest) occurrences of each major invertebrate group over the past 250 Myr

These data are drawn from nearly 20 years of work in museum collections and the paleontological literature. To permit analyses on a truly global scale, and the incorporation of older collections where stratigraphic precision is often lacking, first occurrences are placed in 5-Myr time bins, based on the timescale of Berggren et al. ('95). Use of alternative geologic timescales, as in Golonka and Kiessling (2002), has no significant impact on the patterns shown here.

The taxonomic framework used here differs considerably from that of previous papers in this series, providing another check on the robustness of the patterns reported here and earlier. Jablonski and Bottjer ('90a–c, '91; also Jablonski, '93) essentially adopted the taxonomy used by Sepkoski ('82) in his compendium of marine animal families, which in turn was derived mainly from the monumental Treatise on Invertebrate Paleontology. The new analyses discussed here are based on the more cladistic, phylogenetically based classification of Benton ('93), with substantial updates (e.g., Smith, 2001; Hooper and Van Soest, 2002; Mikkelsen, 2002; Williams et al., 2002; Daly et al., 2003; Okusu et al., 2003; Grande et al., 2004). The revised classifications increase the number of orders analyzed from 42 to 53. The oldest known occurrences of some groups have also shifted significantly over the past decade, sometimes because of new discoveries (e.g., Blake et al., 2000), but more frequently by phylogenetic refinements. For example, several polyphyletic sponge taxa have been resolved into multiple orders (Hooper and Van Soest, 2002; Reid, 2003; Finks et al., 2004; Nichols, 2005), the oldest reported pennatulacean corals have proven to be misidentified (Riegraf, '91), and the sole fossil record of the micropygoid echinoids is evidently a diadematoid (Vadet et al., '96, p 54–57). The database has thus been significantly expanded and revised for this analysis.

A method to assess absence of a taxon

This is a problem in any spatially explicit analysis of living or ancient taxa, but is especially acute in the fossil record: does the absence of a taxon from a given time or place reflect real biology, preservation, or sampling? Bottjer and Jablonski ('88) introduced taphonomic control taxa as one approach to this problem: absences would only be considered valid where the control group—a taxon having preservational characteristics similar to those of the focal group—had been recorded. Thus, records of cheilostome bryozoans are lacking from many post-Paleozoic localities, but Jablonski et al. ('97) treated only those sites where cyclostome bryozoans have been recorded as valid cheilostome absences and were able to map the environmental history of the group on that basis.
A null hypothesis for broader sampling biases

Taphonomic control taxa cannot fully take into account the great variation in intrinsic preservation potential among marine metazoans, and temporal variations in the availability of fossiliferous deposits of varying preservation quality. To provide a sampling model, the 53 orders of benthic marine invertebrates can be partitioned into two major preservation categories, based on the potential for recovering identifiable skeletons or skeletal elements (Jablonski and Bottjer, ’90a,c, ’91); this qualitative assessment has been corroborated by an analysis of the distribution of gaps within, and at the base of, the observed temporal ranges of the orders when placed in a phylogenetic framework (Jablonski, in preparation). For example, many authors have emphasized that the seastar fossil record is severely distorted by its poor preservation potential (Blake, 2000, 2002), and this is borne out by the gaps and “ghost lineages” (e.g., Smith, ’94) in their geologic history. Thus, the 25 Myr of the Early Jurassic fossil record contain only 12 species and nine families of asteroids, but Villier and Kutscher (’99) have shown that any reasonable phylogeny requires the presence of at least 23 families at that time, and two orders not yet known from that interval (Villier et al. (2004) add three more Early Jurassic families and four more species, but make the same phylogenetic point). Similarly, echi-nothuroid echinoids are basal to the euechinoid radiation that began in the Early Triassic, but are known from only a handful of fossil occurrences, the oldest in the Middle Jurassic of Switzerland (Smith and Wright, ’90). The roster of marine orders used in this analysis, partitioned according to preservation potential, is provided in the caption to Fig. 1. This dichotomy between good-preservation groups with robust hardparts (N = 31), and poor-preservation groups that rapidly disintegrate or disarticulate after death (N = 22), could be further subdivided but this inevitably yields groups too small for statistical analysis.

Partitioning the orders according to preservation potential creates a powerful tool for assessing paleontological patterns. The poor-preservation group provides a large-scale null model for the quality of the post-Paleozoic fossil record, because the time and environment of the oldest fossil occurrences of these poorly preserved taxa are probably dictated primarily by preservation and sampling. Thus, the environmental patterns in the good-preservation group should only be treated as a biological signal if they differ significantly from the null pattern derived from the poor-preservation group.

ENVIRONMENTAL PATTERNS

Despite the extensive revision of the data, these new analyses corroborate the results of Jablonski and Bottjer (’90a,c, ’91): first appearances of the good-preservation groups are significantly concentrated in onshore habitats, above average storm wavebase (Figs. 1 and 2). About 80% of good-preservation groups first appear in onshore environments, but only 40% of the poor-preservation groups do, a difference significant at P < 0.005 (log-likelihood ratio test; Sokal and Rohlf (’95) (the good-preservation frequencies also differ significantly from the even simpler null model of an even division between onshore and offshore settings, P < 0.02)). This outcome is qualitatively indistinguishable from the results obtained with a different taxonomic framework more than a decade ago (Jablonski and Bottjer, ’90a-c, ’91), suggesting that the patterns are robust to alternative systematic approaches and ongoing paleontological discoveries.

We can further test whether the onshore pattern is an artifact of the quantity of fossils or localities available from different habitats. As shown in Fig. 2C, the environmental distribution of first occurrences of good-preservation orders differs significantly from the first occurrences of species-level records belonging to those orders as compiled by Jablonski and Bottjer (’91) (n = 2,227). Again, orders first appear onshore more often than expected from this measure of sampling intensity. These analyses could still be misleading if taxonomic classifications are truly arbitrary, with ordinal originations unrelated to the generation of evolutionary novelty. Jablonski and Bottjer (’90c) made a crude first step in allaying this concern, which seems unlikely when posed so starkly, by specifying the novel character states associated with ordinal originations in the sea urchins (see Bambach’s (’85) inventory of the functional significance of differences among marine invertebrate classes, a pathfinding survey performed at a higher taxonomic level but in a similar spirit). A fuller evaluation of the mapping of apomorphies against taxonomic rank would be valuable, but Eble (2000) takes a different approach that may be even more promising. He analyzed the first
appearances of two major echinoid clades, the holasteroids and spatangoids, relative to an extinct stem group or grade evocatively named the disasteroids, in a multivariate morphospace. He calculated the nearest-neighbor distances, or disparity (see Foote ‘97), among all the species of the stem group, and resampled the results 1,000 times to produce an expected frequency distribution of

![Fig. 1. Environments of first occurrence for (A) good-preservation orders and (B) poor-preservation orders of benthic marine invertebrates.](image-url)

**Note:** The revised classification of sponges (Phylum Porifera) in the *Treatise on Invertebrate Paleontology* (Finks et al., 2004), which arrived after this manuscript was submitted, is not entirely compatible with the new neontological treatment (Hooper and Van Soest, 2002). However, a preliminary analysis using the *Treatise* classification, and recent molecular phylogenetic results (see Nichols, 2005), finds no significant effect on the analyses reported here: it eliminates one onshore origin (by range extension into the Paleozoic) and adds an early Jurassic offshore origin among the well-preserved orders, but it also adds up to five offshore origins for the poor-preservation orders. Similarly, the hydrozoan orders Stylasterina and Milleporina are subsumed by some recent authors into Order Anthothecata (see Marques and Collins (2004) and references therein), a clade that begins in the earliest Paleozoic or even late Proterozoic (see Nudds and Löser’s (2001) data on their Order Hydroidea), but this taxonomic change also has no significant effect on the statistical results reported here.
morphological divergences within that stem
group. The first members of the two new orders
fell significantly outside of this distribution,
implying that these (onshore) originations are
exceptionally divergent relative to all other inter-
species disparities within the species pool from
which they evolved. In this instance at least, the
onshore novelties indicated by ordinal rank are
significant evolutionary events even when we use
morphological criteria alone to recognize impor-
tant shifts in morphology.

**DISCORDANCE ACROSS LEVELS**

A comparison of the environmental pattern of
ordinal originations with those seen at lower
taxonomic levels further helps to exclude a
sampling hypothesis, and reveals an evolutionary
discordance across hierarchical levels. Genera and
families within even the well-preserved orders
lack a strong onshore bias in first occurrences,
tending instead to appear more evenly across the
continental shelf, or even to originate offshore,
depending upon a given clade’s bathymetric
diversity gradient (see also Bottjer and Jablonski,
’88; Jablonski and Bottjer, ’91).

This discordance across levels can be seen in the
first occurrences of well-preserved crinoid and
echinoid genera (Fig. 3A and B), which differ
significantly from the first occurrences of new
orders—or, more precisely, from the first occur-
cences of genera that initiate new orders. If Smith
(’94) is correct that the global sedimentary record
should be partitioned between the early Mesozoic
and the late Mesozoic–Cenozoic, a more rigorous
test might involve analyses within the first
100 Myr of the post-Paleozoic interval—i.e., from
250 to 150 Myr ago—particularly if some orders
retreat into deep-water distributions after that
time so that late-arising genera are constrained in
their potential environments of origin. Smith’s
(’94) general point is worth addressing but his test
is flawed by his use of calycinid (saleniod)
echinoids as his genus-level sample, because the
group only appeared in the mid-Jurassic ca.
175 Myr ago and took at least 10 Myr to spread
offshore, resulting in relatively brief overlap with
the time of major ordinal production (Jablonski
et al., ’97). A more appropriate test is to compare
the first occurrences of *all* echinoid genera in the
early Mesozoic to the first occurrences of the good-
preservation orders within the same time interval
(echinoid genus data are so far available only
through the first half of the mid-Jurassic, i.e., for
the interval 250–170 Myr ago, but this block is
suitable because it encompasses 19/31 = 61% of
the ordinal originations). The distribution of
genus-level first occurrences differs significantly
from the ordinal pattern (Fig. 3B), refuting
Smith’s (’94) contention that echinoid genera
show the same environmental pattern of origina-
tion as the orders. As Jablonski et al. (’97) note,
numbers are too small for statistical analysis in
the crinoids, but they trend in the same direction,
with 77% of the genera founding new orders in the
interval 250–170 Myr ago starting onshore, in
contrast to only 43% of all other genera in those
clades during that time. Similarly, of the 10
echinoid genera that found new orders during
this interval, 80% first occur in onshore deposits
despite the ~45% onshore occurrences for the
other echinoid genera originating within this
interval (Fig. 3B).

Turning again from taxonomic treatments to
direct analyses of morphological novelties, post-
Paleozoic bryozoans also exhibit a discordance
between the origin of within-clade novelties
and the onshore origination of good-preservation
orders. For both cyclostomes and cheilostomes, within-clade novelties at both zooid and colony level (e.g., the origin of specialized brooding zooids, or of arborescent growth forms, respectively) show no significant tendency to originate onshore (Fig. 3C), but instead tend to appear in the environments with the highest within-clade taxonomic richness (Jablonski et al., '97). Thus for the groups that have been analyzed so far—crinoids, echinoids, and bryozoans—orders preferentially originate onshore but origins of lower taxa and within-clade novelties tend to be diversity-dependent, i.e., tend to be stochastic with respect to environment per se. Further, the fossil record certainly contains examples of genera or families that not only originate offshore but also invade onshore habitats, although the only available whole-clade statistical treatment found that most genera remain in their ancestral habitat, with the relatively few environmental shifts evenly divided between the onshore and offshore directions (Jablonski and Smith, '90).

**DISCLAIMERS**

As a working hypothesis, we can take the onshore first appearances of good-preservation taxa, anchored by taphonomic controls in contemporaneous offshore settings and a variety of tests for sampling and preservation artifacts, to indicate that major groups of marine invertebrates, and the evolutionary innovations that define them, tend to originate in disturbed, onshore environments. This is a statistical statement of course, and a few groups seem to have convincing offshore origins, such as the stylasterine hydrozoans (Jablonski and Bottjer, '90c; Lindner, 2004; although some recent work has demoted them from ordinal to familial status, e.g., Marques and Collins, 2004). The major clade of carnivorous snails in modern seas, the neogastropods, may also have originated in deep water (Jablonski and Bottjer, '90c; Kollmann, '92), although more work is needed and they are no longer accorded ordinal rank. If the neogastropods did arise offshore, the immensely successful onshore expansion of this group, contrary to the spatial dynamics of so many others, would be a valuable target for detailed study.

This is not to claim that the fossil record and the current systematic treatment is perfect or complete. However, the congruence of the results presented here to the outcome of analyses performed 15 yr ago on a rather different and considerably less refined dataset (Jablonski and Bottjer, '91) suggests that the major patterns have stabilized in the face of ongoing sampling and taxonomic revision (see notes in caption to Fig. 1). In light of this stability, and the environmental distribution of poor-preservation orders and species records, the middle shelf appears to be sufficiently well-sampled to rule out a strongly contrary trend, such as large numbers of origins in poorly sampled deep-sea environments followed by an undetected sweep across the shelf to first appear in nearshore settings.

The onshore pattern in order-level originations does not necessarily mean that orders arise via macromutation or other saltations. Eble’s (2000) morphometric analysis shows that the oldest species in his two echinoid orders are divergent from the ancestral clade, but lacks the resolution

![Fig. 3. Environmental patterns in first occurrences of low-level novelties differ significantly from those of good-preservation orders.](image-url)
to track the evolutionary transitions in detail. Morphological analyses are not available for other orders, but the founders of some groups exhibit more dramatic suites of derived character states than do others, suggesting that such analyses would be valuable.

These results should also not be taken to indicate that ordinal rank per se has an objective reality that is perfectly compatible across all phyla. However, the discordant behavior of taxa ranked as genera on the one hand and orders on the other suggests that the origin of novel designs, with the potential to diversify and accumulate additional derived characters, is governed by factors that are different from those that determine the origin of species that simply produce more species and genera. The orders analyzed here certainly were not defined or assigned that rank according to their environment of first occurrence, but knowing that a species is the first member of a good-preservation order allows us to predict its environment, statistically speaking. Even though each major group must be initiated by the origin of a new species, the ecology of the body-plan origination, captured in a crude way by the origin of higher taxa, differs significantly from the ecology of most speciation events.

**POST-ORIGINATION ENVIRONMENTAL SHIFTS**

Each of the 25 good-preservation orders that originated in onshore settings expanded across the continental shelf, albeit at different times and different rates, and at least six of those orders eventually abandoned their ancestral habitats and are now restricted to offshore settings (Lychniscosida, Isocrinida, Millericrinida, Pedinoida, Calycina (=Salenioida), Holasteroidea). The expansion offshore is not terribly surprising, in that marine clades originating onshore can only expand in one direction as they diversify (Jablonski and Bottjer, '83; Gould, '88, 2002), but the abandonment of ancestral onshore habitats is less expected. Even poor-preservation taxa can be informative here, for example, by providing onshore records for taxa today restricted mainly to deep water, as in the notomyoid asteroids (Blake and Reid, '98). However, we know almost nothing about the factors that allow some taxa to maintain their full-shelf distribution but drive others to contract into an exclusively offshore, and even deep-sea, distribution.

The evolutionary-ecologic processes that promote the invasion of offshore habitats by onshore clades have important implications for the organization and dynamics of marine communities (Jablonski and Bottjer, '83, '90a–c; Bottjer and Jablonski, '88; Sepkoski, '88, '91). Biotic interactions such as predation and competition have most often been invoked (Vermeij, '87, '95; Rhind and Allen, '92; Aronson, '94; McNamara, '94; Oji, '96; McClintock et al., '99; Walker and Brett, 2002; Oji et al., 2003), and this makes intuitive sense if novel predators and competitors also originate most frequently in shallow waters. However, little is known about how, or whether, such biotic interactions scale to order-level shifts over tens of millions of years. More work is needed to determine the ecological and functional histories of marine predators and prey over the post-Paleozoic interval (Walker and Brett, 2002), and the potential role of differential extinction probabilities (see Sepkoski, '91), differential origination in the wake of extinction events (Jablonski and Smith, '90), and other macroevolutionary dynamics including the long-term and indirect effects of clade interactions (Jablonski and Sepkoski, '96; Sepkoski, '96). Contraction into deep water appear to be slow and evidently occur at different times in different regions so that stragglers or participants in short-lived reversals can be found in shallow-water deposits (e.g., stalked crinoids and hexactinellid sponges in the Eocene and Oligocene; see Kutscher, '81; Bottjer and Jablonski, '88; Eagle, '93; Meyer and Oji, '93; Stilwell et al., '94; Aronson et al., '97; Gammon et al., 2000). It may be no coincidence that these reversals or holdouts are in high latitudes, but this apparent latitudinal pattern needs to be tested critically against sampling biases (see Clarke et al. (2004) who argue that the seemingly archaic nature of the Antarctic biota actually represents a Cenozoic transformation from more typical modern assemblages).

The offshore shifts observed in post-Paleozoic clades do not mean that the outer shelf or deep sea is simply a museum of ancient lineages. Although some ancient taxa certainly do reside in the deep sea (e.g., Smith ('84, p 137) “All deep-sea echinoid groups can be traced back to shallow-water forms that evolved during the Mesozoic and Tertiary”; see also Rhind and Allen, '92; Wilson, ’99; Richer de Forges et al., 2000; Little and Vrijenhoek, 2003; Smith and Stockley, 2005), we lack a comparative analysis of taxon ages in shallow and deep-sea settings that takes sampling into account, and
many deep-sea taxa appear to be relatively young (see also Warén and Bouchet, 2001). Even the deep-sea members of clades that began in shallow water are often highly derived and represent extensive in situ diversification (Smith, '84; Améziane and Roux, '97; Smith and Stockley, 2005). Like all marine habitats, the deep sea contains an amalgam of in situ origination and invaders. Its greatest contrast with shallow settings may be the relatively low number of higher taxa that it has exported to other biotas, despite potential invasion corridors via the Antarctic (which has probably facilitated movement of lower taxa in both directions; Gage and Tyler, '91, p 257–258; Clarke et al., 2004) and the hydrothermal vent/seep system.

MECHANISMS

The environmental pattern in the origin of major marine groups, and its contrast with dynamics at lower levels, is a phenomenon that lies at the macroevolutionary intersection of ecology and developmental biology. Comparable patterns have long been recorded for Paleozoic marine invertebrates, albeit via somewhat different analytical approaches (Bretsky, '68; Jablonski et al., '83; Sepkoski and Sheehan, '83; Sepkoski and Miller, '85; Miller, '88, '89; Droser et al., '93; Bassett et al., '99a, 2002; Babin, 2000). The evidence for onshore–offshore dynamics seems stronger for the Ordovician diversification than for the Cambrian explosion, but new work is needed in both intervals that more directly accounts for sampling and the availability of important environmental types (see Guensberg and Sprinkle, '92; Westrop et al., '95; Westrop and Adrain, '98; Bassett et al., '99b; Droser and Finnegan, 2003; Novack-Gottshall and Miller, 2003). Landing and Westrop’s (2004) recent analysis of the Cambrian diversification in the relatively complete and fossiliferous sequences of the Avalonian microcontinent (now coastal maritime Canada, eastern Newfoundland, eastern Massachusetts, east-central England) is an important step in this direction. Controlling for sampling, they find a significant onshore–offshore component to this most dramatic of radiations, with all major groups except trilobites first occurring onshore.

Moving out of the invertebrate realm, marine teleosts are also reported to show onshore origination; the fossil record is thin but the data appear to be consistent with evolutionary tree topology (Santini and Tyler, 2003). Even further afield, Paleozoic plants may show an analogous pattern (DiMichele et al., '87, 2001; DiMichele and Aronson, '92); flowering plants appear to have arisen in disturbed, marginal habits (e.g., Hickey and Doyle, '77; Taylor and Hickey, '96; Wing and Boucher, '98; Feild et al., 2004); and Henderson et al. (98) suggested a similar dynamic on neontological grounds for the plants of tropical floodplains. These non-marine events and habitats are difficult to compare to the marine systems, and quantitative analyses such as those presented here are not available, but the arguments are tantalizing.

Rigorous tests of hypotheses on driving mechanisms for these intriguing patterns will require joint efforts by biologists and paleobiologists. Jablonski and Bottjer ('90a–c, '91) used the timing of first occurrences and expansions across the shelf to test, and ultimately reject, an array of potential mechanisms imposed by physical events such as mass extinctions, sea-level fluctuations, and oceanographic changes. Jacobs and Lindberg (‘98) hypothesized a somewhat different environmental driver, with the onshore-origination bias imposed by low oxygen levels in the deep sea during the first three fifths of the Mesozoic. Although an intriguing idea, this hypothesis seems to be framed at an inappropriate scale: the pattern discussed here is determined by dynamics within the continental shelf, and not by differences between the shelf and the deep sea, and in any case the assumption that the deep sea was continuously anoxic or dysoxic through the early Mesozoic and consistently oxic after the mid-Cretaceous is a serious oversimplification (e.g., Kutzbach and Guetter, '90; Horne, '99; Jenkyns, 2003; although the photic zone may have been dysoxic during the three most severe events of the Mesozoic, see Pancost et al. (2004); see Little and Vrijenhoek (2003) for additional critiques, and Wilson ('99) and Smith and Stockley (2005) on survival of deep-sea clades during these events). Further, Jacobs and Lindberg’s analysis omits two important distinctions: between the frequently disturbed inner shelf and the rarely disturbed middle shelf (which they group into onshore settings), and between the good-preservation and poor-preservation taxa and the contrasting signals they capture. Even if we accepted their claim that poor-preservation patterns should be taken literally, the revised database presented here does not support their earlier observation that such taxa did not appear in offshore environments (however defined) until the late Cretaceous. With
this revised database, only 13 orders of any kind first appear after Jacobs and Lindberg’s mid-Cretaceous cutoff (90 Myr ago) for the transition from anoxic/dysoxic to fully oxic deep oceans, so that statistical comparisons between preservation groups are not feasible; ongoing work has shifted so many first occurrences into the early and mid-Mesozoic that the argument is no longer relevant (or testable). The same holds for Smith’s (’94) suggestion that the onshore-origination pattern described above for the 250 Myr of post-Paleozoic time is an artifact of the generally low sea levels of the early Mesozoic (yielding a preponderance of onshore strata) and the generally higher sea levels of the late Mesozoic and early Cenozoic. In any event, nothing in these sparse data indicates contrasting environmental patterns of first appearances for good-preservation orders before and after the early or mid-Cretaceous.

The hypotheses that remain fall into two broad categories: differential novelty production and differential novelty survival in onshore habitats. If novelty is truly generated more frequently onshore, then this is a problem in genetics and developmental biology. The literature is rich in reasons that onshore populations might probe morphospace more actively or more inventively than offshore ones, but none have been tested critically. The onshore bias in evolutionary innovations might be attributable to bathymetric gradients, known or hypothesized, in:

(a) Diversity and variability of selection pressures: The potential evolutionary effects of spatially and temporally variable environments (e.g., Nevo, 2001) may be an intuitively appealing hypothesis for the preferential occurrence of innovations onshore. However, this view does not account for the discordance of environmental patterns across hierarchical levels, which seems to undermine any model involving a simple, per-species probability of novelty production. An exception is Rice’s (’90) model, where populations in more variable environments are more likely to undergo rapid evolution in earlier stages in development but not necessarily in later stages. Developmental burden, i.e., the constraints on early developmental stages owing to their role in determining later characters, may be less important in populations that spend most of their time far from local selective optima (Rice, ’90).

(b) Nutrient inputs or productivity: By supporting larger populations, high productivity may decrease extinction probabilities of novel morphology (e.g., Parsons, ’93). Again, this hypothesis does not seem to address the discordance across levels, nor does it take into account the apparent lack of a secondary locus of innovation associated with shelf-edge upwelling seen on many coasts.

(c) Phenotypic plasticity: Plasticity is clearly the rule in natural populations and is certainly a potential vehicle for evolutionary innovation (e.g., Pigliucci, 2001; West-Eberhard, 2003, and this symposium; Schlichting, 2004), but we still do not know if the type or degree of plasticity varies systematically along environmental gradients. Henderson et al. (’98) suggested that tropical plants exhibit higher plasticity in more disturbed habitats, favored by a combination of high rates of gene flow and low patch persistence, noting in this context that “low speciation is not low macroevolution.”

(d) Population genetics: Perhaps the higher rates of gene flow and chaotic patch dynamics typical of onshore populations—frequently yielding small isolates drawn from large, panmictic populations—elicit rare but more profound developmental perturbations, as Templeton (’86) suggested (see also Jablonski and Bottjer, ’83).

(e) Environmental stress: Many authors have argued for a link between evolutionary innovation and environmental stress, defined in terms of the physiological tolerances of the organism rather than any absolute measure of environmental variability (e.g., Parsons, ’93; Hoffmann and Parsons, ’97; Badyaev and Foresman, 2000; Hoffman and Hercus, 2000; Nevo, 2001). Although many potential mechanisms for the stress–novelty link have been proposed, the most striking have recently come from molecular biologists. For example, transposable elements may be activated or released by environmental stress, by intense inbreeding arising from population bottlenecks, or by hybridization events as might occur in habitats typified by frequent extinction and recolonization (McClintock, ’84; McDonald and Cuticchia, ’93; Kidwell and Lisch, ’97, 2001; Grandbastien, ’98; Capy et al., 2000). Some authors have further argued that such transposon bursts, which might—or might not—be more frequent in onshore settings, are the sources of significant evolutionary innovation and not simply catastrophic disruptions of the smooth functioning of the genome (e.g., McDonald, ’90, ’93; McDonald and Cuticchia, ’93; Shapiro, 2000). Others have suggested that molecular chaperones such as heat-shock proteins act as evolutionary capacitors, so that hidden variation or mutagenic activity is released when the organism is subjected to
temperature or other environmental extremes, and that such events will occasionally yield adaptive phenotypes (e.g., Rutherford and Lundquist, '98; Bergman and Siegal, 2003; Madlung and Comai, 2004; Sangster et al., 2004; but see Ruden et al. (2003) for some caveats). The relation between these mechanisms and the nature of phenotypic plasticity is still unclear, and more generally we need to know whether theoretical findings on the advantages of “a canalization system that tends to be inactivated in extreme environments” (Eshel and Matessi, ’98) are matched by the distribution of phenotypes and their evolutionary behaviors along gradients.

Some of these ideas on the generation of evolutionary novelties along gradients of environmental disturbance are more plausible than others, but all call for real testing. We simply do not know how the developmental and molecular—genetic processes that govern novelty production—what Müller terms (in this symposium) “mechanistic causation”—are arrayed along environmental gradients. The paleontological results presented here strongly suggest that it would be worth finding out.

The alternative to the broad class of novelty-production hypotheses is that the paleontological pattern is actually driven by greater novelty survival onshore, and thus is primarily an ecological phenomenon. Novelty production might actually be stochastically constant across the continental shelf, but the onshore environment might foster greater survival of evolutionary novelties during the perilous early stages of the founding of a new clade, owing perhaps to productivity gradients or metapopulation dynamics. The tendency for onshore species to be more extinction-resistant than offshore ones, observed in several paleontological analyses of background extinction, is consistent with this view (see, for example, Jablonski and Bottjer, ’83). Extinction differentials need not be strong to generate the pattern: these data from marine invertebrates involve only about 30 events over 250 Myr. However, if differential species-extinction rates did shape the onshore-origination pattern, then we might not expect the discordance across hierarchical levels: novelties at all levels should persist at higher rates onshore. More combined modeling and empirical work is needed to test this hypothesis, perhaps by dissecting the within-clade dynamics of several orders having different intrinsic extinction rates (see Jablonski and Smith, ’90; Sepkoski, ’91).

A NOTE ON BIOGEOGRAPHIC PATTERNS

These data are also relevant to one of the major questions in biodiversity dynamics: do the tropics operate as “a cradle or a museum”—as a diversity source or a diversity accumulator (e.g., Stebbins, ’74; Willig et al., 2003)? Virtually all of the post-Paleozoic orders that have maintained their onshore distributions occur today at low latitudes, suggesting that the tropics have functioned as a museum in the origin and maintenance of their latitudinal diversity gradient. However, as Jablonski (’93) found for the earlier version of this dataset, when the strong sampling biases favoring the high latitudes are taken into account, the tropics appear also to function as a cradle, in that they contain more first occurrences than expected from the relatively poor sampling in those regions. This cradle-plus-museum pattern, which is consistent with the dynamics seen among late Cenozoic bivalves at the genus level (Jablonski et al., 2004), will be the subject of a separate paper.

SUMMARY AND CONCLUSIONS

The fossil record shows us that major evolutionary innovations—those commonly afforded ordinal rank—do not arise randomly in time and space. The new analyses presented here corrobrate earlier suggestions that higher taxa of benthic marine invertebrates preferentially originated in onshore, disturbed habitats, at least over the past 250 Myr. This pattern holds even for clades that today occur at all depths, or are restricted to offshore settings, so that present distribution is not a reliable indicator of evolutionary origins. In contrast, the first appearance of genera appears to be diversity-dependent, shaped by clade-specific bathymetric gradients. In the Introduction, I suggested that first principles would not necessarily predict environmental patterns for the origins of higher taxa—and this is especially true if the patterns are absent at lower taxonomic levels, as shown here. Perhaps something is missing from our first principles, or, more precisely, in our understanding of how they operate across environmental gradients over geologic timescales.

Although many potential driving mechanisms for the onshore-origination pattern and the ensuing onshore–offshore dynamic have been ruled out, the many possibilities that remain are difficult to test decisively, and of course more than one may be valid, singly or in concert for a given clade. Whatever the underlying processes,
these unexpected but pervasive patterns in the fossil record challenge our theories of the origin of evolutionary innovation and evolvability, by forcing us to frame these theories in spatially explicit, ecological terms. This challenge demands a strengthened partnership for ecology, evolutionary developmental biology, and paleontology. Such a partnership is sure to be fruitful, and sure to provide us with a steady stream of new surprises.

ACKNOWLEDGMENTS

I thank Susan M. Kidwell, Carlton E. Brett, and Gerd Müller for valuable comments on the manuscript; C. Kevin Boyce, Allen G. Collins, Alberto Lindner, Paula M. Mikkelsen, Dennis M. Opresko, Brian R. Rosen, Andrew B. Smith, and Paul D. Taylor for generous advice and assistance; and Christa Modschiedler, Crerar Library, for help during an improbable bibliographic emergency. I am also grateful to Gerd Müller and Stuart Newman for their invitation to participate in such a stimulating symposium. Much of this research was undertaken as an Honorary Research Associate. I am also grateful to Gerd Müller and Stuart Newman for their invitation to participate in such a stimulating symposium. Much of this research was undertaken as an Honorary Research Associate, Natural History Museum, London.

LITERATURE CITED


