Micro- and macroevolution: scale and hierarchy in evolutionary biology and paleobiology

David Jablonski

Abstract. — The study of evolution has increasingly incorporated considerations of history, scale, and hierarchy, in terms of both the origin of variation and the sorting of that variation. Although the macroevolutionary exploration of developmental genetics has just begun, considerable progress has been made in understanding the origin of evolutionary novelty in terms of the potential for coordinated morphological change and the potential for imposing uneven probabilities on different evolutionary directions. Global or whole-organism heterochrony, local heterochrony (affecting single structures, regions, or organ systems) and heterotopies (changes in the location of developmental events), and epigenetic mechanisms (which help to integrate the developing parts of an organism into a functional whole) together contribute to profound nonlinearities between genetic and morphologic change, by permitting the generation and accommodation of evolutionary novelties without pervasive, coordinated genetic changes; the limits of these developmental processes are poorly understood, however. The discordance across hierarchical levels in the production of evolutionary novelties through time, and among latitudes and environments, is an intriguing palaeontological pattern whose explanation is controversial, in part because separating effects of genetics and ecology has proven difficult. At finer scales, species in the fossil record tend to be static over geologic time, although this stasis—to which there are gradualistic exceptions—generally appears to be underlain by extensive, nondirectional change rather than absolute invariance. Only a few studies have met the necessary protocols for the analysis of evolutionary tempo and mode at the species level, and so the distribution of evolutionary patterns among clades, environments, and modes of life remains poorly understood. Sorting among taxa is widely accepted in principle as an evolutionary mechanism, but detailed analyses are scarce; if geographic range or population density can be treated as traits above the organismic level, then the palaeontological and macroecological literature abounds in potential raw material for such analyses. Even if taxon sorting operates on traits that are not emergent at the species level, the differential speciation and extinction rates can shape large-scale evolutionary patterns in ways that are not simple extrapolations from short-term evolution at the organismal level. Changes in origination and extinction rates can evidently be mediated by interactions with other clades, although such interactions need to be studied in a geographically explicit fashion before the relative roles of biotic and physical factors can be assessed. Incumbency effects are important at many scales, with the most dramatic manifestation being the postextinction diversifications that follow the removal of incumbents. However, mass extinctions are evolutionarily important not only for the removal of dominant taxa, which can occur according to rules that differ from those operating during times of lower extinction intensity, but also for the dramatic diversifications that follow upon the removal or depletion of incumbents. Mass extinctions do not entirely reset the evolutionary clock, so survivors can exhibit unbroken evolutionary continuity, trends that suffer setbacks but then resume, or failure to participate in the recovery.

David Jablonski. Department of Geophysical Sciences, University of Chicago, 5734 South Ellis Avenue, Chicago, Illinois 60637. E-mail: djablons@midway.uchicago.edu

Accepted: 18 July 2000

Introduction

The landscape of evolutionary biology has changed significantly over the past quarter-century. Evolutionary and ecological studies now regularly incorporate serious considerations of history, scale, and hierarchy. This expansion of the working toolkit of the discipline—the near-routine application of ideas that were once barely developed, highly abstract, or in some circles outright anathema—heralds a more profound integration of evolutionary biology, paleobiology, systematics, ecology, and developmental biology. Whatever the intellectual forebears of these ideas (and some of their roots are very deep), the growing number and diversity of studies that directly address such factors as intrinsic constraints, phylogenetic effects, differential origination and extinction rates, and local vs. regional effects represent a true operational expansion of evolutionary theory. As a window...
onto a wide range of spatial and temporal scales, including extreme events not accessible to neontological study, paleontology is playing a vital role in this expansion and will certainly continue that role in the coming decades.

This is by no means to declare the demise of the highly successful microevolutionary paradigm, which is alive and well, but rather to say that the study of evolution continues to evolve and expand conceptually, and increasingly incorporates approaches that explicitly emphasize scale and hierarchy. These approaches, even those that can be traced back to Darwin, were barely visible when Tom Schopf and Ralph Johnson were mustering support for a new journal to be titled Paleobiology. Paleontology was one of the fields that provided a phenomenology, a conceptual base, and a battery of new quantitative methods, that fostered the expansion of evolutionary theory in the first quarter-century of this journal.

One way to survey the infusion of hierarchy and scale into evolutionary biology is through the classic darwinian two-step process: the origin and the sorting of variation. Other schemes are possible, of course, from the intrinsic/extrinsic dichotomy of causal mechanisms (e.g., Gould 1977a; Jablonski 2000a) to a nested set of temporal and spatial scales (e.g., Gould 1985; Bennett 1997), and those elements will inevitably appear here as well.

**Origin of Variation**

The origin of heritable variation, the raw material of the evolutionary process, is by definition a matter of genetics. One major issue in bridging scales and hierarchical levels, however, has been the correspondence between genotypic changes and the magnitude and direction of phenotypic transformation. A reasonable assumption underlying most models of evolution has been that the probability of a phenotypic change is inversely related to the complexity or magnitude of the genetic change required to generate it (among other factors, of course). Geneticists have long recognized that mutations of equal magnitude can have strikingly different phenotypic consequences depending on context, but this awareness has been somewhat lost in the empirical evidence from quantitative genetics that many traits are underlain by a large number of genes of small and mainly additive effect. However, intensive study of developmental processes in multicellular organisms has led to a new appreciation of how modest genetic changes can be amplified and channeled developmentally to yield significant variations in the magnitude and direction of phenotypic change. We are only beginning to understand how and when to apply this expanding set of approaches to the origin of novel morphologies (and the absence or rarity of certain forms), but few would deny the potential to both illuminate and be illuminated by the fossil record of morphological evolution—not least because, as discussed below, that record exhibits nonrandom patterns in space and time.

**Genetic Control Pathways and Networks**

At one extreme in the nonlinearities between genotypic and phenotypic change are the high-level regulatory genes now under intense study by molecular developmental biologists. Some of these genes, such as the Hox clusters that help to pattern the body axis and appendages across the entire breadth of metazoan diversity, specify positional information and thereby regulate the transcription of a large number of downstream genes. Others sit near the top of a regulatory cascade, or perhaps more commonly within a regulatory network, that determines a specific tissue or structure. For example, perhaps 2500 genes are involved in building and maintaining the *Drosophila* eye (Halder et al. 1995; 18% of the fly's genome by Adams et al.'s [2000] count, although of course many are also used elsewhere), but experimental manipulation of a single gene—most famously *eyeless*, but also *eyes absent*, *dachshund*, and others—can yield well-formed eyes in the middle of wings and other improbable sites (see reviews by Gerring and Ikeo 1999 and Hodin 2000; and see Chow et al. 1999 for similar results in the frog *Xenopus*). Such high-level regulatory genes and even entire signaling pathways have been recruited as modular units (“cassettes”) in the service of novel morphologies. Thus, the
hedgehog pathway that underlies the generation of limbs also participates in the production of eyespot patterns on butterfly wings (Keys et al. 1999); the Toll pathway has been recruited to the development of chick limbs, the operation of the vertebrate immune system, and the development of ventral structure in Drosophila (Gonzalez-Crespo and Levine 1994; Ghosh et al. 1998), and even the "master control gene" for eyes, Pax-6, plays multiple roles (e.g., Quinn et al. 1996; Duboule and Wilkinson 1998; Hodin 2000; and see Heanue et al. 1999 and Relaix and Buckingham 1999 for an analogous situation involving Pax-3).

Many of the major molecular components of metazoan development appear to have arisen very early and been retained across taxa with very different body plans of varying complexity (reviews in Finnerty and Martindale 1998; Erwin 1999; Valentine et al. 1999; Holland 1999; Knoll and Carroll 1999, among others); although data are sketchier, the same appears to hold for plants (Purugganan 1998; Lawton-Rauh et al. 2000; Theissen et al. 2000). The extraordinary conservation (albeit in a highly dynamic fashion involving gene loss, gain, and duplication) and the multiple expression events and sites of these genes strongly suggests that the evolutionary action at this level is in the enhancer regions that govern the timing and intensity of gene expression. The evolutionary impact of changes in the major regulatory genes will therefore be more complex than the extreme binary effects suggested by the homeotic mutants that abruptly transform segment identities. Instead, that impact should probably be visualized in terms of finer modulations in the expression of major regulatory genes, and therefore should include evolutionary changes that do not depend on the viability of extreme saltations and lone individuals (Gellon and McGinnis 1998; Akam 1998; Duboule and Wilkinson 1998; Purugganan 1998, 2000; Gibson 1999; Li and McGinnis 1999; Ludwig et al. 2000). Possible examples of striking morphological changes that may have involved relatively subtle shifts in Hox expression patterns include the evolutionary differentiation of arthropod appendages (Averof and Patel 1997; Shubin et al. 1997; Weatherbee et al. 1999), morphological transitions of vertebrae along the spinal column (Burke et al. 1995; Belting et al. 1998), the origin and diversification of tetrapod limbs (Shubin et al. 1997; Coates and Cohn 1998), and the suppression of those limbs and the homogenization of vertebral morphology along the body axis of snakes (Cohn and Tickle 1999; Greene and Cundall 2000). Vertebral identity and digit size and number appear to be dose-dependent functions of Hox gene products (Zakany et al. 1997), providing a mechanism for selection on high-level regulatory factors within populations, and thus a pathway for coordinated change effected by major genes but in a polymorphic population rather than a strictly typological context. Similar situations, again underlain by polymorphisms in regulatory genes, have been recognized by plants as well (e.g., Purugganan 1998, 2000; Lawton-Rauh et al. 2000).

These are very early days in the macroevolutionary exploration of developmental genetics. Careful study is needed in the inference of causal, evolutionary relationships between variations in Hox expression and morphological differences among taxa (e.g., Rogers et al. 1997). And the next step, taking insights derived from laboratory populations into natural evolutionary processes, has just begun. Gibson et al. (1999) found that crossing mutants of the Hox genes Ubx and Antp into natural populations of Drosophila could evoke a wide range of extreme phenotypes whose expression depended on the overall genetic context of the mutation, again showing how large-effect genes might impinge on morphological variation generated in the wild. More speculatively, DeSalle and Carew (1992) attribute some of the morphological extremes seen in Hawaiian drosophilid species, such as grotesquely reshaped heads and bizarre mouthpart appendages, to mutations in Hox genes Ubx and Antp into natural populations of Drosophila could evoke a wide range of extreme phenotypes whose expression depended on the overall genetic context of the mutation, again showing how large-effect genes might impinge on morphological variation generated in the wild. More speculatively, DeSalle and Carew (1992) attribute some of the morphological extremes seen in Hawaiian drosophilid species, such as grotesquely reshaped heads and bizarre mouthpart appendages, to mutations in Hox genes because of their resemblance to Antp mutants in laboratory D. melanogaster—an attractive hypothesis that invites direct molecular study. Nevo et al. (1992) reported apparent Hox gene polymorphisms that significantly correlated with a wide array of morphological variables in the mole rat Spalax ehrenbergi, an early, pro-
vocative result that needs further work. On the other hand, Ahn and Gibson (1999a,b) found intraspecific variation in expression domains of Hox genes along the body axis of the three-spined stickleback, but this variation did not correlate with phenotypic variation. Finally, considerable within-species variation has been recorded for a number of major developmental genes in natural and domesticated plant populations, and several dramatic intraspecific variants have proven to be underlain by high-level developmental genes (for overview see Lawton-Rauh et al. 2000).

Evolution via changes in Hox-gene expression is hardly a prerequisite for a strong non-linearity between genetic and phenotypic change, of course, and both theory and evidence are accumulating that genes of large effect—many of them presumably involved in regulating development—do play a role in the origin of adaptations and the divergence of species (e.g., Gottlieb 1984; Orr and Coyne 1992; Palopoli and Patel 1996; Orr 1998). Much empirical work is still needed, of course, but an absolutely micromutational view of evolutionary change in natural populations seems increasingly untenable even for quantitative traits that are often seen as the bastion of the “many genes of equal and infinitesimal effect” approach championed by R. A. Fisher. One recent surprise, for example, has been the number of studies reporting quantitative trait loci (QTLs) of large phenotypic effect in both laboratory and natural populations (Orr 1998: p. 936; see also Paterson et al. 1997; Voss and Shaffer 1997; Doebley and Wang 1997; Bradshaw et al. 1998; Schemske and Bradshaw 1999; Goffinet and Gerber 2000). QTL analyses have their limitations, of course: they are only feasible among species or morphs that can be hybridized, they may not permit an absolute determination of the numbers of genes involved in important traits because minor factors may not be detected, and they can even be biased toward artificially inflating the effects of single QTLs or underestimating the number of genes involved in complex traits (Routman and Cheverud 1997; Lynch and Walsh 1998: p. 474–476; Via and Hawthorne 1998). On the other hand, laboratory selection experiments can be biased toward document-
and McNamara 1991). Further difficulties are imposed by the hierarchy of biological organization: changes in timing at the molecular or cellular level need not produce heterochronic patterns at the whole-organism level, and the heterochrony at the organismal level need not involve changes in the timing of molecular events (e.g., some salamanders owe their paedomorphosis simply to the disabling of the production or reception of the hormone thyroxin, which is closer to a binary switch than a change in timing).

Global or systemic heterochrony—in which the entire phenotype is shifted relative to the ancestral ontogeny—has the greatest evolutionary potential, and is most readily detected, in species that undergo substantial phenotypic changes during ontogeny. Both continuous but nonlinear changes such as ontogenetic allometry and saltational changes such as metamorphosis (as in the classic example of neotenic salamanders that mature in the aquatic larval state) provide ample raw material for dramatic evolutionary events. This implies a little-explored predictive approach to differences in morphological diversification among clades, based on the nature of ancestral ontogenies. This might involve, for example, testing how well the ontogeny of individual organisms can predict the extent of morphological diversification of their descendant clades, a macroevolutionary equivalent of Wayne’s (1986) classic conclusion that the great array of morphologically extreme dog breeds relative to cats is a consequence of their contrasting ontogenetic allometries (see also Wayne and Ostrander 1999).

The preceding example represents just one way in which the evolution of development at the individual level might indirectly shape large-scale patterns. Selection on life-history traits over ecological timescales, which as Gould (1977b) notes can often result in heterochrony, can also have far-reaching, indirect macroevolutionary effects, for example via changes in genetic population structures. For example, populations of obligate paedomorphic ambystomatid salamanders are genetically more distinct from one another than are metamorphosing populations (Shaffer 1984). Presumably the obligate paedomorphs tend to stay in their aquatic neighborhoods, yielding lower rates of gene flow, and thus potentially higher speciation rates compared with metamorphosing relatives free to travel among ponds. Similarly, selection for small body size per se, energy economy, or short generation time in benthic invertebrates can evidently give rise to paedomorphs that are so small that they must evolve a low-dispersal, non-planktotrophic mode of development due to fecundity constraints (see Jablonski and Lutz 1983; Lindberg 1988); such a change in developmental mode would in turn be likely to result in high speciation and high extinction rates at the clade level, as discussed below (and see experiments showing that selection on egg size can in turn alter other aspects of larval biology, such as those of Sinervo and McEdward [1988] and Emlet and Hoegh-Guldberg [1997]). Again, this intriguing intersection of micro- and macroevolution, where life-history theory meets heterochrony (an important but neglected insight in Gould 1977b; see also McKinney and McNamara 1991; McKinney and Gittleman 1995), would produce dramatic shifts not only in morphology but in evolutionary dynamics.

Although global or systemic heterochrony has enjoyed the most attention, local or specific heterochrony has been much more common (e.g., McKinney and McNamara 1991). Such changes in the rate or timing of development of particular structures within an organ can break allometric relationships and generate new and coordinated morphologies, again by drawing on established developmental interactions but this time within a localized region or developmental field. For example, the most successful crinoids in modern oceans, the stemless, mobile comatulids, were derived from the stemmed crinoid order Isocrinida; intermediates such as Eocomatula and Paracomatula suggest progressively earlier offset of stem formation, with the substratum-gripping cirri, which adorn the long stem in the isocrinids, finally arising from a single centro-dorsal ossicle on the base (Simms 1988a,b, 1994; Hagdorn and Campbell 1993).

In mammals, the ossification of facial bones is accelerated relative to the central nervous system in marsupials compared with both mono-
tremes and placentals, correlated with the particular demands of neonate survival, including head movements during migration to the pouch, attachment to the teat, and suckling (Smith 1996, 1997; Nunn and Smith 1998). Such localized developmental changes, from the elongation of pterosaur digits to support a wing to the elongation of echinoid plates to create a protruding rostrum, represent coordinated changes in suites of complex tissues. Such dissociation and reintegration among local growth fields is still poorly understood but permits an enormous range of morphologies to be tapped by alterations of local growth fields. This does not mean that phenotypes are infinitely malleable, but that the scope for evolution via heterochronies of existing morphologies is far wider than implied by global heterochrony alone.

Calibration of ontogenetic trajectories against better metrics than the convenient but unreliable size criterion will be important for rigorous analysis of local and global heterochrony, particularly if life-history parameters are potential targets or by-products of selection. Fortunately, this can be achieved in palaeontological material where accretionary growth leaves internal growth lines and a stable isotope record that can be calibrated at the level of monthly, seasonal, and annual periodicities (e.g., Jones 1988, 1998). In some taxa these can be tied to onset and perhaps frequency of reproduction, because reproductive growth interruptions can often be distinguished from disturbance checks (Kennish 1980; Harrington 1987; Sato 1995, 1999). Such skeletal chronometers have been used extensively in ecological studies of bivalves, but will be extremely valuable in evolutionary applications (for an exemplary study, see Jones and Gould 1999).

A final mechanism for coordinated morphological change is heterotopy, or alteration in the location of a developmental event. Heterotopy has received much less attention than heterochrony, but some authors have argued that spatial changes in development may prove to have greater evolutionary impact than temporal ones (e.g., Raff 1996; Zelditch and Fink 1996; Hall 1999). Although this perceived frequency is in part a consequence of an expanded definition that may include virtually every developmental change that is not narrow-definition heterochrony (see Klingenberg 1998: p. 83), spatial changes in developmental events can certainly give rise to novel morphologies, as in the shifting of the scapula from outside to inside the ribcage in turtles (Burke 1989, 1991). Hall (1999: p. 388) goes so far as to say, “heterochrony tinkers, but heterotopy creates” and to anticipate that “heterotopy may be about to come into its own as heterochrony wanes and our knowledge of developmental mechanisms increases.” This may be selling heterochrony a bit short (Hall goes on to note the intimate connections between heterochrony and heterotopy), but the recognition that the evolution of development involves more than just heterochrony—also being driven home by molecular work on gene regulation—is welcome. And, continuing one of the themes of this section, heterotopy is an effective evolutionary mechanism because, like heterochrony, it draws on preexisting developmental pathways and components: the spatial reordering of the turtle skeleton is striking by any measure, but the turtle tucks its scapula, identifiable as such during morphogenesis, under its ribcage rather than evolving a novel structure.

Epigenetics

Local heterochronies and heterotopies are effective evolutionary agents because of another aspect of developmental systems that contributes to the nonlinearities between genetic change and morphological effect: epigenetics in the classical sense, i.e., the local cell and tissue interactions that help to integrate the developing parts of an organism into a functional whole (for example the induction of vertebral cartilage formation by contact with the spinal cord [Hall 1983] or the growth response of both embryonic and postnatal bone to mechanical loading [Carter et al. 1998]). By drawing on a set of preprogrammed responses to local signals, such interactions allow the developing embryo to accommodate evolutionary changes in particular morphological elements without a host of independent but mutually beneficial mutations. We have little detailed knowledge of mechanisms, but strik-
ing epigenetic responses to experimental alterations in morphology have been described by many workers, and their evolutionary implications abundantly discussed (reviews include Rachootin and Thomson 1981; Raff and Kaufman 1983; Thomson 1988; Atchley and Hall 1991; McKinney and McNamara 1991; Hall 1999). Those experiments show that evolutionary changes in, for example, eye size need not be accompanied by independent mutations in genes governing bones, nerves, muscles, or blood vessels in the skull (see Twitty's famous transplant experiments reviewed by Müller 1990), and changes in vertebral posture need not involve independent mutations in genes that collectively control thoracic cross-sectional shape or the sites of muscle insertions on limb bones (see Slijper's bipedal goat, reviewed by Rachootin and Thomson 1981), because epigenetic interactions yield accommodations to those morphological changes.

The epigenetic interactions that help generate complex forms do have limits to the changes they can accommodate, as attested by many "failed" embryological manipulations or the more bizarre gene regulation experiments that haunt the pages of Nature, Cell, and similar journals. The macroevolutionarily important questions revolve around where those limits lie for particular clades or particular kinds of changes, and how evenly the potential directions of permissible change are distributed in morphospace. An approach along those lines will help to illuminate how two of the seemingly conflicting themes of development, modularity, and integration—at several hierarchical levels within the organism from molecular pathways to tissue inductions—conspire to produce evolutionary novelty.

Wanted: An Integrated Genetics

All of these aspects of genetics strongly underscore the nonrandom and nonlinear nature of variation that is the raw material for evolution. Any probability distribution of potential changes around a genotype (phenotype) will inevitably be inhomogeneous, reflecting evolutionary lines of least resistance that are conditioned by the underlying structure of developmental pathways at the molecular, cellular, and tissue levels. The challenge is to take these new insights from the organismal level to the population, species, and clade level to forge a better understanding of the links and discontinuities among those levels.

A synthesis of developmental and population genetics will not be easy, as R. A. Fisher himself recognized ("I can no longer calculate it," he said when confronted with early evidence for nonadditive genetic variation [Mayr 1992]). A general mathematical theory may not much resemble our textbook versions of microevolutionary population genetics or quantitative genetics but will contain elements of both. Various, rather disparate, pioneering attempts suggest some potential components of such a program (e.g., Arthur 1988, 1997; Atchley and Hall 1991; Atchley et al. 1994; Schluter 1996; Nijhout and Paulsen 1997; Wagner et al. 1997; Orr 1998; Rice 1998). Given that our knowledge of the genotype—phenotype relation is still heavily biased toward laboratory strains of a few model organisms, far more extensive analysis of natural populations will be a critical step whose early phases were noted in the section on genetic hierarchies and networks, above.

A discouraging possibility is that so much taxon-specific information will be required that quantitative models, or even qualitative predictions, at the macroevolutionary scale will have little power. However, the strong commonalities among phyla in basic developmental genetics hold out some hope, and certain differences among clades suggest some simple hypotheses that might be explored. For example, protostomes have evolved the genetic machinery for generating complex morphologies by keeping genomes relatively small and increasing the number of times a gene is used during development, whereas deuterostomes have enlarged the genome by two rounds of duplication so that multiple copies of a given gene are available to diverge functionally (Akam 1998; Holland 1999; Valentine 2000; among many others). Do these differences in genetic architecture impinge on phenotypic evolution to produce differences in evolutionary style between the two main metazoan branches (e.g., in dissociability of structures—that is, the relative degree of
developmental independence of structures or organs within bodies, styles of heterochrony or heterotopy, etc.)? The two architectures may prove to be functionally equivalent from a macroevolutionary perspective, but demonstrating that they generate qualitatively as well as quantitatively similar variation would be equally interesting. Whether phenotypic evolution is gradual or discontinuous when regulatory changes give rise to major morphological novelties is still utterly uncertain, and probably depends on the types of characters involved. Reasonable arguments are available for either evolutionary dynamic and so data rather than theory will be required to determine relative frequencies.

Bridging the empirical and conceptual gap between developmental biology and macroevolution remains a challenge. Living and fossil organisms are all the products of developmental sequences that themselves had to evolve to yield the diverse forms of past and present taxa, but hypotheses on the nature and evolutionary impact of those developmental changes are difficult to test, and inferences on developmental changes in extinct clades can rarely be precise. Using our growing knowledge of gene-expression events within, for example, a Drosophila or crustacean embryo as a basis for understanding the morphological diversification of the arthropod clade (and vice-versa) represents a truly daunting shift in spatial and temporal scale. But here as elsewhere, paleontology has much to gain and much to offer. For example, paleontological comparisons across clades, habitats, or time intervals of the density distributions and sequences of novel morphologies around their ancestral starting points can make substantial contributions to understanding the large-scale consequences of the organization of developmental systems. This work will of course be most powerful if done in concert with actual developmental information from the clade in question, but paleontology’s spectacular array of morphologies viewed in historical perspective can lead the way for a host of new questions and help to target organisms for developmental analyses likely to yield macroevolutionary insights.

Temporal and Spatial Variation in the Origin of Novelties

Given the many ways in which developmental changes can generate evolutionary novelties, the macroevolutionary expectation might be that novelties arise stochastically according to clade-specific pressures and opportunities. The highly nonrandom first appearances of novelties in the fossil record thus present an intriguing challenge. Although debate continues on the mechanisms behind these patterns, and more empirical work is sorely needed, these large-scale patterns in time and space hint at the potential paleontological contributions to a theory of evolutionary novelty.

Temporal Patterns.—The most striking burst of evolutionary creativity in the animal fossil record comes early in the Phanerozoic, with the Cambrian explosion of metazoan bodyplans. This extraordinary interval, which saw the first appearance of all but one of the present-day skeletonized phyla (along with an array of less familiar forms) in an interval of less than 15 m.y., has received considerable attention recently from both geological and developmental perspectives. The standing of the two major explanatory models, one involving intrinsic, developmental or genomic controls and the other involving extrinsic, environmental or ecological controls, has varied over the past decade, but resolution has been difficult (see Erwin et al. 1987; Jablonski and Bottjer 1990a; Valentine 1995; Conway Morris 1998; Erwin 1999, 2000a; Knoll and Carroll 1999; Valentine et al. 1999; Jablonski 2000a). The rival hypotheses need not be mutually exclusive, of course, for as Erwin (2000a) notes, “successful innovations require ecological opportunity, developmental possibility and an appropriate environmental setting.” The problem therefore becomes a matter of assessing which of Erwin’s triad of requirements—for example, open ecospace as set by the biota, intrinsic morphogenetic potential, or by favorable oxygen levels or other physical limiting factors—was the immediate trigger of the explosion. An equally intriguing and no less elusive problem is to determine which of them damped the production of bodyplan after-
wards, without actually damping diversification at lower levels—including further developmental modifications—so that the production of phyla ceased but lower-level taxonomic diversity soared to new heights during the Ordovician, and again in the post-Paleozoic. And of course the trigger might have been different from the damper.

Secondary pulses of evolutionary innovation occur in the wake of mass extinctions, adding further temporal structure to the origin of novelties. These repeated pulses, measured either as the first appearances of high-ranking taxa or increases in morphological disparity, strongly suggest a role for ecological opportunity in the origin of novelties (see also below). But is this sufficient to account for the Cambrian explosion, or were genomic factors also involved? Erwin et al. (1987) reasoned that taxonomic diversity but not genomic flexibility should have approached Cambrian levels after the Permo-Triassic extinction, allowing a test of the genomic and ecological hypotheses. However, they found that even the end-Permian debacle, although cutting deep into taxonomic diversity, failed to remove major functional groups from the marine biosphere and thus did not sufficiently mimic the ecological character of Cambrian seas for a definitive test.

Another approach is to sidestep taxonomic rank and examine the evolution of morphology more directly. The rapid filling of a quantitatively defined morphospace has been documented for a number of clades originating in the early Paleozoic (see Foote's 1997 review). More daunting is the task of quantifying larger and more heterogeneous groups such as the deuterostomes or the bilaterians for the purpose of gaining an overall picture of the deployment of morphological diversity through time—but see Thomas et al.'s (2000) analysis showing that 80% of the theoretical skeleton designs available to living and extinct metazoans were occupied by the Middle Cambrian.

Comparative approaches to morphospace occupation can be used to address the rival hypotheses for temporal trends in the origin of novelties. For example, Foote (1999) found that the post-Paleozoic rediversification of cironoids, although representing a rapid increase in morphological disparity, yielded a narrower set of architectural novelties than those established in their initial, early Paleozoic radiation. He took this as evidence for less constrained developmental inputs to the Cambrian explosion than in later times. Wagner (1995) also argued for the lability of different kinds of traits during the initial and later phases of Paleozoic gastropod diversification. These are intriguing results that appear to support a genomic component to the Cambrian explosion, but they are just a first step, in part because they lack an explicitly generative or developmental component. For example, a developmentally based partitioning of traits to compare the long-term behavior of those that might have been more subject to initial freedom and later limitation, relative to those that might have been subject to earlier entrenchment (as attempted by Jacobs 1990) would be valuable. Pinpointing appropriate characters for comparative analysis may not be straightforward, however. For example, Hughes et al.’s (1999) finding that the variability of thoracic segment number in trilobites depends on the number of segments in a given species rather than its phylogenetic position or geologic age undermines a particular argument (McNamara 1983) but cannot fully address the kinds of regulatory networks that were involved in the radiation of arthropod and related clades in the Early Cambrian.

Instead of focusing on particular characters, the role of genomic changes in the damping of the Cambrian explosion might be tested by tracking levels of morphological integration through the early Paleozoic (Erwin 1994)—that is, the morphometric correlation patterns within an organism (e.g., Olson and Miller 1958; Cheverud 1996). This may require a better understanding of developmental mechanisms than presently available, because phenotypic integration can be stable even when genotypic covariances are not (e.g., Turelli 1988; Shaw et al. 1995; Schlücht and Pigliucci 1998). However, Nemeschkal’s (1999) finding that avian morphometric correlation patterns correspond to the expression domains of Hox and other developmental control genes is encouraging in its implication...
that morphometric matrices can reflect developmental architecture (see also Leamy et al. 1999). Operational questions aside, the evolutionary role of morphological integration needs further theoretical and empirical study: maximal evolutionary lability may come at an intermediate level of integration, where the body is composed of locally integrated units that can behave as modules, as discussed above (e.g., G. P. Wagner 1996; Kirschner and Gerhart 1997). G. P. Wagner (1996) argues that the origin of multicellularity led to a decrease in integration as regional specialization of morphology led to differential gene expression, but developmental “burden” is held to increase through time as developmental interdependencies accumulate (e.g., Riedl 1978; Donoghue 1989). The two perspectives are of course potentially compatible, as they may represent very different scales (multicellularity vs. the developmental genetics of a single feature), but they do suggest that this issue would be worth exploring in greater depth.

Although direct comparisons are difficult, a major burst of novel plant architectures is associated with the invasion of land (e.g., Niklas 1997; Bateman et al. 1998). We need to develop criteria to determine whether the establishment of the major land plant designs was as profound an evolutionary event as the Cambrian explosion, as sometimes suggested. If so, we will probably need to confront the intrinsic/extrinsic debate here as well: were plant developmental systems more labile as late as the Devonian (in contrast to animal systems), or is the ecological opportunity afforded by the assembly of traits permitting terrestrial existence a sufficient explanation? Recent work suggests that the MADS-box genetic architecture that orchestrates plant development in ways reminiscent of Hox and other high-level regulatory genes in animals were in place early in land plant evolution (Theissen et al. 2000).

Evolutionary novelties at lower levels exhibit very different temporal patterns from major novelties, and may depend on different variables. As made clear by Valentine’s (1969, 1973, 1980, 1990) seminal work on evolutionary and ecological hierarchies in the fossil record, this is immediately evident from the discordance between the diversity dynamics of marine taxa ranked as phyla and classes relative to the very different dynamics of those ranked as families and genera (see also Erwin et al. 1987). Such discordances are not simply an artifact of the greater inclusiveness of higher taxa (as suggested by Smith [1994], among others), because similar patterns emerge from taxon-free analyses of multivariate morphological data (Foote 1993, 1996a, 1997, 1999; and see Lupia 1999 on an early burst of disparity followed by stability in angiosperm pollen, discordant with species-level diversity). Raup (1983) essentially recognized this as well by noting that the requirements of tree topology alone could not account for morphological divergences like the Cambrian explosion (that is, the topology of the tree cannot account for the autapomorphies on each branch).

The dynamics of novelty production at lower levels can be quite unexpected when examined in detail. For example, Jablonski et al. (1997) found the production of morphological novelties within the bryozoan orders Cyclostomata and Cheilostomata to be opposite in timing to that expected from the ecological-opportunity hypothesis that is the chief contender for explaining high-level originations. Cyclostomes generated novelties in a steady trickle despite their occurrence in the relatively low-diversity early Mesozoic world, whereas the cheilostomes produced novelties in a burst despite being embedded in the presumably more crowded, predator- and competitor-rich mid-Cretaceous environment. In this comparative study, ecological context mattered less than the relative speciation rates (and competitive abilities?) of the respective clades.

Spatial Patterns.—Evolutionary novelties also show spatial patterns in their first appearance in the fossil record. Like the major temporal patterns, onshore–offshore patterns show discordances between the first appearances of higher taxa and patterns of origination, extinction, and diversity accumulation at lower taxonomic levels: benthic marine orders tend to originate in onshore, disturbed habitats, regardless of the diversity dynamics of their constituent species, genera, and families (Jablonski and Bottjer 1990a,b, 1991; Droser et
Although much more work is needed, this pattern can also be seen in terms of derived characters (Jablonski and Bottjer 1990b is only a crude beginning) and in the morphological divergence of the founding species for two echinoid orders relative to the disparity among species in the ancestral group, showing how the onshore initiators of these orders "broke away from the bounds of [their ancestor's] morphospace from the very start" (Eble 2000: p. 68). In apparent agreement with these patterns, major land plant originations also appear to be concentrated in disturbed habitats, in both Paleozoic and Mesozoic settings (e.g., DiMichele and Aronson 1992 and Wing and Boucher 1998, respectively), although the botanical data have not been analyzed for the marine discordance across hierarchical levels.

The geography of first occurrences of major groups is particularly subject to sampling bias, but an attempt to take these into account found disproportionate appearances of marine invertebrate orders in tropical latitudes (Jablonski 1993). Similarly, phylogenetic analysis suggests that major plant lineages tend to originate in the Tropics and spread poleward, in that primitive members of clades tend to be tropical and derived taxa tend to be restricted to or best developed in the temperate zones (e.g., Judd et al. 1994, and in a very different tradition, Meyen 1992), although of course different dynamics could underlie this pattern (but see also Askin and Spicer 1995 on paleontological data). More work is needed, however, to test whether the latitudinal trend in higher-level originations significantly exceeds the probabilistic, per-taxon expectation, given that the latitudinal trend in species richness appears to have been present over most of the Phanerozoic, albeit with varying slope and subject to considerable sampling and preservation bias (e.g., Stethi et al. 1969; Humphreys and Bambach 1979; Kelley et al. 1990; Crame 1996; Walsh 1996).

Species and genera show less striking latitudinal origination patterns, and many clearly started in high latitudes (e.g., macroinvertebrates [Feldmann et al. 1993; Crame 1997], Neogene Foraminifera [Buzas and Culver 1986; Wei and Kennett 1986; Spencer-Cervato et al. 1994], terrestrial plants [Wen 1999]). The relative fraction of high-latitude origination at low taxonomic levels has not been sufficiently quantified to compare with the preferential low-latitude appearance of the much smaller number of higher taxa, however. Plotting the mean or median geologic age, or alternatively the estimated net diversification rates, of extinct lower taxa against latitude has given apparently conflicting results (e.g., for marine taxa [Wei and Kennett 1986; Flessa and Jablonski 1996 and references therein; Crame and Clarke 1997], terrestrial birds [Gaston and Blackburn 1996], birds and butterflies [Cardillo 1999]). Violation of assumed time-homogeneous dynamics may be one source of the conflict, but this entire area deserves more extensive study.

Potential mechanisms for spatial biases in evolutionary innovation are plentiful but difficult to test, and as with the temporal bias the jury is still out on whether this is a genetic or an environmental problem, or even whether the pattern is underlain by differential production or differential persistence of evolutionary novelties (or both). Environmental possibilities include high rates of local extinction and invasion and thus local opportunity for innovation, more variable and/or intense selection pressures, or the potential extinction-resistance of novelty-bearing species, in more disturbed habitats relative to more stable ones (e.g., Jablonski and Bottjer 1983, 1990b; Hoffmann and Parsons 1997). Latitudinal patterns might simply represent diversity-dependent novelty production at this spatial scale, or more complex combinations of novelty production and survival to produce the net effect recorded paleontologically (Jablonski 1993).

Genetic explanations for these patterns include the greater potential of small peripheral isolates—which would arguably be more frequently formed in disturbed, heterogeneous habitats—to have lower developmental stability or to break developmental canalization (Levin 1970; Jablonski and Bottjer 1983; Clarke 1993; Hoffmann and Parsons 1997). Most recently, Rice's (1998) model suggests that canalization might most readily be broken by strong truncation selection, when significantly
less than 50% of a population is selected to produce the next generation, a situation that might obtain most often with the enormous mortality of propagules in shallow-marine benthos (although the selective basis of that mortality and its overall impact relative to mortality in other populations is poorly understood [e.g., Rumrill 1990; Pechenik 1999]). The potential role of highly variable environments in fostering evolutionary innovation (e.g., Parsons 1993, 1994; Hoffmann and Parsons 1997; Hoffmann and Hercus 2000) has gained a new developmental wrinkle with the finding that some regulatory molecules that suppress phenotypic variation can be disabled not only by mutation but by environmental extremes such as high temperatures (Rutherford and Lindquist 1998; Wagner et al. 1999).

The connection between any of these mechanisms and the empirical macroevolutionary patterns remains highly speculative. However, they do suggest testable ways of making mechanistic sense of evolutionary patterns that are not smooth extrapolations up the taxonomic hierarchy. Comparative analyses of morphological divergences within and among clades, structured along environmental gradients, would be a valuable—and tractable—step in this area, particularly because the first appearances of most orders are not as divergent from likely ancestors as are the phyla of the Cambrian explosion. Complementary approaches that examine the first appearance of major groups in terms of derived characters or multivariate morphometrics relative to patterns within clades and across environments (along the lines of Eble 2000) would be valuable.

The discordance between high and low taxonomic levels in temporal and spatial patterns of origination, and between morphological diversification at different levels within nested clades and subclades, thus provides an intriguing set of patterns that require a hierarchical approach. The spatial component also demonstrates that large-scale evolutionary processes cannot be analyzed exclusively at the global scale, because unexpected—or at least previously undetected—structure resides at the regional scale and across environmental gradients (see also Miller 1998; Jablonski 2000a). These aspects of the evolutionary process can also be seen in the sorting of variation.

### Sorting of Variation

#### Species-Level Stasis and Change

**Stasis and Its Causes.**—As many authors have pointed out, microevolution can occur as rapidly as needed to account for virtually any speed observed in the fossil record (e.g., Charlesworth et al. 1982; Gingerich 1993; Kirkpatrick 1996; Hendry and Kinnison 1999; and many others). This has been abundantly demonstrated not only in the laboratory but in natural populations from Galapagos finches (Grant 1986; Grant and Grant 1995) to Bahamian anoles (Losos et al. 1997) to mosquitoes in the London Underground (Byrne and Nichols 1999), although as discussed above simple extrapolations of these changes may not provide the best model for all of the inhomogeneities in the origins of major novelties. The more challenging question then becomes, why are evolutionary rates generally so slow in the fossil record? This question pertains both to the species level, which is the domain of punctuated equilibrium and its alternatives (Gould 1982; Gould and Eldredge 1993), and to the clade level, where large-scale evolutionary trends often unfold with excruciating slowness when viewed on microevolutionary time-scales (e.g., Stanley 1979; McShea 1994). The relation between potential mechanisms at the different levels has been discussed mainly in broad generalities, but few workers have attempted to address whether the factors that cause, for example, species-level stasis seen in many members of the horse lineage (e.g., Prothero and Shubin 1989) are also responsible for the slow rate of body size increase in the clade. Averaged over the duration of the entire clade, this size increase was so slow as to be virtually indistinguishable from drift (see Lande 1976, and Stanley's [1979, 1982] punctuational reinterpretation, seconded by Stebbins 1982; and also Lieberman et al. 1994, who found rates so slow in a Devonian brachiopod lineage that they would have involved only three selective deaths per 10 million individuals if treated as a continuous trend).
Over the past quarter-century, evolutionary stasis has proven to be a pervasive morphological pattern in the fossil record (reviewed in Erwin and Anstey 1995a; Gould and Eldredge 1993; Hallam 1998; Jackson and Cheetham 1999). However, few of the hypotheses on the forces that maintain this stasis at the species and higher levels have been conclusively tested and again, different mechanisms may obtain in different clades. The research questions have shifted to testing for among-clade and among-habitat differences in frequencies of evolutionary tempo (abrupt vs. gradual change) and mode (anagenetic vs. cladogenetic, sometimes termed phyletic vs. branching), the roles of intrinsic and extrinsic factors that might govern those differences, and whether the direction of phenotypic change during sustained anagenesis or cladogenesis is related to the morphologic behavior of the species or its constituent populations during preceding intervals.

Given pervasive stasis, the stunning diversity and subtlety of biological adaptations must often arise episodically, in the punctuations between stable species, either in single punctuational episodes—which, of course may encompass tens or hundreds of thousands of years (e.g., Jackson and Cheetham 1999: Table 1)—or in cumulative series. This process need not rely entirely on isolation itself as the trigger for adaptive change, but may also draw on geographic variation within established species. Consider, for example, Futuyama's (1987) very attractive but still untested suggestion that speciation events cordon off local adaptations into discrete gene pools, thereby packaging ordinarily ephemeral characters into more stable evolutionary units (see also Eldredge 1989, 1995). The apparently episodic nature of this process, at least in terms of the morphologies accessible in the fossil record, underscores the need to understand stasis.

At the species level, stasis over geologic timescales has been attributed to variation in both rate and direction of change. Variation in the rate of change involves truly slow evolutionary rates between the punctuations, with temporal stability generally attributed to constant selection for intermediate phenotypes, interrupted by rapid anagenetic or cladogenetic shifts (maximum observed rates of change may also be artificially reduced by the size of the time bin encompassing both stasis and directional change). Less often considered is the possibility that directional selection fluctuates so rapidly that populations cannot respond, with the net effect of stasis at the mean phenotype; another alternative would be time-averaging of samples rather than selection pressures, detectable if not geologically then perhaps by exceptional apparent population variances (see Kidwell and Aigner 1985).

Sustained stabilizing selection must be the force behind habitat tracking as a mechanism for stasis (Eldredge's 1985, 1989, 1995 hypothesis), in which species remain morphologically static as they move with a favorable environment during climatic and other changes. The tracking process seems well supported in the Pleistocene (e.g., Valentine and Jablonski 1993; Coope 1995; Clark 1998; see also discussions in Price et al. 1997 and Jackson and Ov erpeck this volume). Intrinsic differences among taxa in their ability to keep up with shifting environments have not been explored as an explanation for differences in evolutionary tempo and mode; this may be unimportant, however, if we can generalize from rates of movement in Pleistocene plants and animals (e.g., Clark 1998). For widespread species, a more realistic model might be cline translocation (coined by Koch 1986), in which a set of populations that vary along an environmental gradient shift in and out of a sampling area to give the appearance of oscillatory or even directional change as the species overall maintains a constant morphology (see for example Stanley and Yang's 1987 extensive study of late Cenozoic bivalves, in which the total range of multivariate oscillations through the history of each lineage was very similar to its present-day geographic variation).

Most species-level lineages appear to lack directionality rather than evolutionary lability; that is, they show high total rates of evolution while accumulating little net change. The most frequently invoked model is that of oscillating directional selection (e.g., Ginger-
ich 1993; Sheldon 1996; Hendry and Kinnison 1999), a process well documented in some modern populations (e.g., Grant and Grant 1995; Via and Shaw 1996; among many others). Such evolutionary dynamics can be modeled and in principle tested against drift and other forces, although paleontological applications are still being developed and in some instances seem highly model-dependent (see Bookstein 1987, 1988; and Roopnarine et al. 1999; but see Cheetham and Jackson’s [1995] overview of their superb multidisciplinary analysis of Neogene bryozoan evolution). A less-explored, explicitly hierarchical alternative involves the spatial structure within species: gene flow among highly dynamic local populations within a species might allow little net overall change (e.g., Eldredge 1985, 1989; Lieberman et al. 1995). Although molecular and other data suggest that few species lack some internal spatial structure (e.g., Hanski 1999; Avise 2000), it is not clear whether the particular metapopulation dynamics required by this model for stasis are truly pervasive in nature (e.g., S. Harrison 1998; Maurer and Nott 1998).

Also controversial is whether the apparent bounds on oscillatory stasis represent intrinsic limits of the organisms or reversals in selection pressure. I am not going to venture into the dense and tangled literature of evolutionary constraints, but the widespread existence of evolutionary trade-offs (as, for example, between age and size at first reproduction, when selection favors both early reproduction and large size [e.g., Stearns 1992]) seems to be a strong endorsement for some form of intrinsic constraint, at least in the short run (for morphological examples, see Nijhout and Emlen 1998). The detection of such trade-offs, however, generally carries little information on mechanisms underlying constraints, and, as with genotypic and phenotypic variance–covariance matrices (e.g., Shaw et al. 1995; Arnold and Phillips 1999), we do not know how stable they are over evolutionary time. Some must be nearly absolute, others may be quite transient and context-dependent. Plant and animal breeders hit limits all the time, and the failure to break the egg-a-day barrier in chickens (Lerner 1953), or to increase thoroughbred racing speeds significantly over the past 70 years (Gaffney and Cunningham 1988), is not for lack of intense directional selection or high heritability of relevant traits.

Experimental work on host specificity in phytophagous insects suggests that intrinsic factors may be important in wild populations as well. Many insects exploit a restricted diet, presumably owing to plants’ defensive compounds, but experiments in some groups have detected no significant relation (or, less often, a positive relation) between insects’ performance on their host plants and their performance on other species, undermining a trade-offs hypothesis; a lack of genetic variation may actually be a limiting evolutionary factor in this instance (Futuyma et al. 1995; but see Keese 1998). This is not a trivial issue, given Farrell’s (1998) contention that the overwhelming species richness of beetles is related to the macroevolutionary consequences of host shifts in phytophagous clades. The general relation between trade-offs, genotypic covariances, and other apparent limitations to evolutionary responsiveness on the one hand, and patterns of morphologic change in species over geologic timescales on the other, is clearly an attractive target for combined paleontological/neontological analysis of particular clades. To cut through the terminological morass, all of these features can be put under the rubric of developmental constraints, which might be defined as the resistance, owing to the configuration of developmental networks and pathways, of the phenotype to selection in certain directions. In principle this can be distinguished from canalization, which might be defined as the resistance, owing to the buffering or redundancy of developmental processes, of the phenotype to mutation or to environmental variation (and see Gibson and Wagner 2000 for a valuable overview).

**Distribution of Stasis and Change.**—The distribution of evolutionary tempos and modes at the species level remains poorly known, not least because rigorous research in this area is such a daunting task. Few studies have fully addressed all of the issues, but, drawing on the discussions of Gould and Eldredge (1977), Fortey (1985), Clarkson (1988), Erwin and An-
stey (1995a), and Jackson and Cheetham (1999), an appropriate protocol would include

1. Large samples in a closely spaced time-series
2. Objective delimitation of species as operational units
3. Stratigraphic control independent of the target clade
4. Independent evidence on sedimentation and preservation rates that might vary to create artificial punctuations or protracted transitions
5. An assessment of within-species geographic variation
6. A phylogenetic hypothesis

The characterization of morphospecies has become increasingly rigorous with the availability of multivariate morphometric methods. An encouraging development has been the generally good correspondence between biological units and the morphospecies of the shelly macroinvertebrates used in most analyses of evolutionary tempo and mode at these scales. This is not the place for an extensive discussion of species concepts, but from an evolutionary perspective species-level units are most useful if they are essentially independent lineages (e.g., Simpson 1961; Wiley 1981; Mayden 1997; de Queiroz 1998, 1999, and references therein). For the outcrossing biparental species that provide most of the animal and protistan fossil record and a sizeable but unknown fraction of the plant record, that independence often involves reproductive isolation or genetic cohesion, and so coincides with any broadly defined biological species concept that can accommodate isolation, recognition, cohesion and related viewpoints (e.g., Templeton 1989, 1998; Ghiselin 1997; Coyne and Orr 1998; R. G. Harrison 1998; de Queiroz 1998). However, that evolutionary independence need not be compromised even if those barriers are not absolute (to give just two examples, the fossil record shows that cottonwoods and balsam poplars have been generating hybrids in western North America since the Miocene but have remained distinct entities [Eckenwalder 1984] and that two lineages of Neogene cyprinid fishes hybridized for 2 m.y. without subverting the evolutionary identities of the parent lineages [Smith 1992]), and other processes besides the traditional isolating barriers may impose or contribute to evolutionary independence as well (e.g., Van Valen 1976; Hull 1997; among many others).

Analyses within paleontologically important phyla where morphometrically defined species correspond closely to biologically, usually genetically defined ones include: the cheilostome bryozoans Stylopoma, Steginoporella, and Parasmittina (Jackson and Cheetham 1990, 1994, who also used an extraordinary set of breeding experiments, and see also Hageman et al. 1999); the benthic foraminifera Glabratella (Kitazato et al. 2000, also based on breeding experiments; but see below for the uncertain situation with planktic Foraminifera); the gastropods Amalda (Michaux 1987, 1989, who also found congruent phylogenies using both data sets), Nucella (Collins et al. 1996, albeit with considerable intraspecific shell variation), Littorina (Rugh 1997, who compared shell morphology with such biological species indicators as genital and egg-capsule features), and Lacuna (Langan-Cranford and Pearse 1995, again using breeding experiments); the corals Porites (Potts et al. 1993; Budd et al. 1994, again with congruent phylogenies) and Montastraea (Weil and Knowlton 1994; Knowlton et al. 1997); the decapod Synalpheus (Duffy 1996, using allozymes); the articulate brachiopod Terebratulina (Cohen et al. 1991, using both allozymes and mtDNA); and even the notoriously nondescript inarticulate brachiopod Glottidia (Kowalewski et al. 1997, who lacked genetic data and relied on previous biospecies definitions). One can only hope for a steady stream of such studies, including a new round on vertebrates such as Steppan's work (1998) on the rodent Phyllotis, using mtDNA versus skeletal morphometrics.

These accumulating results suggest that the paleontologist need be at no greater remove from biological units than any other systematist lacking a full molecular treatment of the taxonomic units under study. And by providing concrete support for the biological reality of the morphological differences between related fossil species, they imply that the morphological punctuations in fossil lineages—an empirical pattern open to multiple interpretations—do tend to correspond to speciation
events. Anagenetically evolving lineages lacking speciation-scale punctuations can be more problematic, of course, and when broken into arbitrary taxonomic segments may imply an artificially punctuational pattern (see Sheldon 1993). However, this pitfall will be avoided as long as phenotypic change is the final arbiter on questions of evolutionary tempo and mode at the species level, as seen in most recent studies including the examples cited here and by Jackson and Cheetham (1999).

On the other hand, sibling or cryptic species—that is, biological species that are virtually undetectable morphologically—are common in many taxa, both terrestrial and marine (e.g., Knowlton 1993; Avise 2000). To some authors, this imperfect correspondence between morphospecies and biologically discrete species dictates the collapse of the entire enterprise (e.g., Levinton 1988; Hoffman 1989), but this simply is not true, so long as the questions are posed appropriately. For example, a lineage is punctuational if most morphological change occurs in the context of speciation when viewed over geologic timescales. But this does not require that the converse be true, that all speciation events are accompanied by morphological change. More problematic is the generation of temporally and spatially persistent, discrete morphotypes that can arise abruptly but are not reproductively isolated, that is, are not evolutionarily independent entities (e.g., Palmer 1985; Trussell and Smith 2000). The examples cited above suggest that paleontologists are becoming adept at partitioning their morphological units in ways that are genealogically significant, but the ranking of discrete morphologies remains a potential problem and needs more attention. The same is true for character-based neontological species concepts, of course, particularly those based on “smallest diagnosably distinct units” (Cracraft 1989, 1997; Nixon and Wheeler 1990; Davis and Nixon 1992; Luckow 1995), where the taxonomic ranking and evolutionary roles of those units also can be controversial (e.g., Theriot 1992; Hull 1997; Knowlton and Weigt 1997; R. G. Harrison 1998).

The most robust analyses will be those that compare rates and patterns of morphospecies production among clades (particularly within the same geological arena, so that many potential taphonomic biases are held constant), rather than depending on absolute values. Significant differences detected in comparative analyses will be misleading only if the frequency of sibling species has a strong inverse relation to the frequency of morphospecies origination. Little or no evidence of such a relation exists, although a formal analysis would be valuable. The sparse literature on important components of the fossil record, such as marine invertebrates and terrestrial vertebrates, conveys the general impression that the numbers of morphospecies and sibling species are, if anything, positively correlated among clades. If this is true, or if the relation is random so that no systematic bias is introduced, then cryptic species will not be a serious problem for comparative studies of evolutionary tempo and mode, at least in large data sets.

Geographic variation has been the Achilles’ heel of many paleontological studies of evolution at the species level. Analyses centered on one or a few closely spaced stratigraphic sections or cores risk confounding the lateral movements of trends in intraspecific variation with evolutionary change, the methodological pitfall created by the cline translocations mentioned above (and the potential for local populations to exhibit independent morphological trajectories without net species-level directionality adds another hierarchical level to be considered [see Lieberman et al. 1995; Bralower and Parrow 1996]). This problem was recognized over 40 years ago (Newell 1956; Arnold 1966), but its remedy is generally so labor intensive that only a few studies have risen to the challenge (but see, gratefully, Stanley and Yang 1987; Cheetham and Jackson 1995, 1996). A formidable obstacle is the inverse relation between acuity of stratigraphic resolution and geographic distance, particularly along environmental gradients or among disjunct regions: the temporal acuity often achieved by closely spaced samples in a single section declines significantly when correlating among sections (see Behrensmeyer and Hook 1992 and Behrensmeyer et al. 2000 [this volume] on analytical time-averaging). This Paleontologi-
cal Uncertainty Principle—the trade-off between temporal resolution and geographic coverage—seems to be little appreciated outside the field but has implications for virtually every kind of paleobiological analysis. Quantitative stratigraphic methods, significant refinements in radiometric dating techniques, and tuning of correlations to Milankovitch cycles (e.g., Shackleton et al. 1999) will yield increasingly fine correlations, but resolution will tend to approach a limit on the order of thousands of years, if only because natural time-averaging operates at about this scale for most micro- and macrofossil records (see Kidwell and Flessa 1995; Martin 1999).

All of the end-member combinations of evolutionary tempo and mode have now been observed in fossil species transitions, and so the challenge is to assess the frequencies of the different patterns, and to test for the influence of biological traits, environmental factors, and other potential controlling variables. This effort is complicated by the strong imbalance in the evidence required to demonstrate gradualism versus stasis (see Fortey 1985, 1988; Clarkson 1988; Sheldon 1993, 1996; Pearson 1995; Wagner and Erwin 1995). Stasis can often be convincingly documented by samples from a succession of discrete sedimentary packages, even when the packages are separated by depositional hiatuses or unfavorable environments. Further, quantifying geographic and other intraspecific variation is less critical if even the local pattern is one of temporal stability. Stasis is unlikely to be artificially generated or removed by time-averaging, where successive populations are homogenized within a single sedimentary bed. Short-term directional changes can be collapsed into a single artificially variable assemblage, but trends extending over more than 10,000 years (depending on depositional environments, of course) and thus significant relative to the average duration of morphospecies, will generally be retained, and situations that would obliterate them can be recognized by independent evidence (e.g., Kidwell and Aigner 1985; Bell et al. 1987; Kidwell and Flessa 1995).

Ironically, then, gradualism is more difficult to demonstrate conclusively in the fossil record than the alternatives, even though it was long taken to be the dominant style of evolutionary change! That said, distinguishing between true punctuated equilibrium, i.e., punctuated cladogenesis, and punctuated anagenesis, in which morphological change occurs episodically but without lineage branching, is not always straightforward either. This distinction cuts to the heart of the question of speciation’s role in evolutionary change: the anagenetic mode can accommodate a broad range of intraspecific evolutionary processes (e.g., Gould 1982; Wright 1982a,b; Lande 1986; and a host of others since then). As noted above, however, establishing the coexistence of ancestor and descendent species, or of multiple sister species, requires a detailed phylogeny and well-resolved stratigraphic range endpoints. Like all paleontological analyses it also hinges on the ranking of morphologically defined units: taxonomic lumpers will tend to reduce the number of branching events, while splitters are more likely to convert anagenetic patterns to cladogenetic ones by increasing the number of contemporaneous taxonomic units. The growing inventory of studies linking morphology to genetically defined species suggests that the splitters have been closer to the biological reality (with past excesses and missteps, of course). Although some cladists have rejected the possibility of identifying ancestral species on theoretical grounds (e.g., Englemann and Wiley 1977; Frost and Kluge 1994; Norell 1996), increasingly rigorous protocols have become available for the recognition of potential ancestors for both fossil and living organisms (e.g., Paul 1992; Theriot 1992; Fisher 1994; Smith 1994; Marshall 1995; P. J. Wagner 1995, 1996a; Foote 1996b; Omland 1997). The data are still sparse but suggest that ancestral species can be detected and that temporal overlap with descendants, as expected for punctuated cladogenesis, is not uncommon. The challenge now is to refine and apply methods that will permit a quantitative assessment of when, where, and how often the different evolutionary patterns obtain in nature. A vast and nearly uncharted territory is open for modeling the interplay of sampling and paleobiological pattern (see Holland and Patzkowsky 1999), but most urgently needed
is a new battery of carefully designed and selected empirical studies.

Attempts to assess the relative frequency of evolutionary tempo and mode are premature, but some possibilities and problems can be defined. As already noted, stasis and punctuation appear to be the pervasive phenotypic patterns in marine macrofossils, although the relative proportions of anagenesis and cladogenesis remain unclear (e.g., Hallam 1998; Jackson and Cheetham 1999). Although more rigorous quantification would be valuable, there is little reason to doubt Forney’s (1985) report that gradualism occurs in fewer than 10% of the 88 trilobite species that have a meaningful stratigraphic range in the Ordovician Valhallfonna Formation, Spitsbergen, or Johnson’s (1985) assessment, backed up by his data-rich monograph (Johnson 1984), that only one of the 34 scallop lineages in the northern European Jurassic shows possible gradual change in morphology. On the other hand, the famous Jurassic oyster *Gryphaea* shows a more complex mixture of stasis and gradualism (Johnson and Lennon 1990; Johnson 1993, 1994), and whether this complexity—and contrast with other contemporaneous bivalves—reflects the intensity of research prompted by *Gryphaea*’s notoriety as a classic evolutionary exemplar, difficulties of phylogenetic analysis in a morphologically difficult and heterochrony-prone group, or a true biological difference, remains uncertain.

Sheldon (1993, 1996) made the intuitively appealing suggestion that benthic species in more stable offshore environments might be more subject to gradual change, but empirical evidence is slim: Sheldon’s trilobite study involves parallel changes in a single character in a set of lineages from a single restricted area in which the environment is changing upsection, albeit subtly (see Sheldon 1987, 1988). Better documented is the long-standing observation that pelagic species are more likely to show gradual change than benthic ones (Johnson 1982; Fortey 1985; Clarkson 1988; Jackson and Cheetham 1999). Fortey (1985) contrasts the evolution of a pelagic trilobite with that of co-occurring benthic species, but the richest data for gradualistic change come from microfossils. Three caveats obtain here: first, geographic coverage remains a weakness of many analyses of pelagic organisms, although this is becoming less true; second, stasis and punctuations do occur in many microfossil lineages (see tabulations in Erwin and Anstey 1995a and Jackson and Cheetham 1999), even when hiatuses are taken into account (see MacLeod 1991); and third, so little is known about the population genetics, or even how individuals are packaged into species, in these unicellular groups that interpretation of paleontological patterns is doubly difficult (e.g., Tabachnick and Bookstein 1990; Norris et al. 1996; Huber et al. 1997; Darling et al. 1999, 2000; de Vargas et al. 1999; but see Kitazato et al. 2000 for encouraging results on a genus of benthic foraminifers, and recall that some genetic analyses are finding that detailed morphometry of, for example, test porosity may help to capture genetic units [e.g., Huber et al. 1997; de Vargas et al. 1999]).

The record for land vertebrates is difficult to interpret because many studies lack one or more of the elements enumerated above (for cautionary notes see, for example, Schankler 1981; Heaton 1993). Some mammal lineages do appear to present robust examples of gradualistic change at the species level, however (see reviews by Barnosky [1987], Martin [1993], Chaline et al. [1993], and Carroll [1997]). For example, Chaline and Laurin (1986) found gradualism in a Plio-Pleistocene vole lineage over a broad geographic area, with quantitative data on cheek-tooth morphology in a series of time planes extending over an area from Spain and Britain to northern Italy, Poland, and the Czech Republic, with additional qualitative data from localities as far east as Moldova and western Siberia. But as with microfossils, mammals are not purely gradualistic in evolutionary tempo; indeed analyses of entire faunas or assemblages of clades suggests that stasis and punctuation is pervasive and perhaps prevalent (e.g., Barnosky 1987; Flynn et al. 1995; Prothero and Heaton 1996). Again the key issue is relative frequency and the factors that impose different frequencies among clades.

Attempts to assess the frequency of different types of speciation based exclusively on modern species have their own pitfalls. As
Wagner and Erwin (1995) note, phylogenetic tree topology alone cannot reliably distinguish evolutionary tempo and mode. Inferences based on molecular data as a source of temporal estimates show considerable promise but remain model-dependent, not only in terms of molecular-rate constancy but in assumptions about the pattern of morphological change between nodes (e.g., Garland et al. 1999). Finally, estimates of the relative frequency of allopatric and other types of speciation based on the present-day deployment of modern species (e.g., Lynch 1989; Barraclough and Vogler 2000) are undermined by the geographic volatility of species in the recent geologic past and by extinction. Only species that have split since the last glaciation, say in the last 10,000 years, are likely to capture the relative spatial distributions of sister species at the time of speciation. Species that split, say, 2 m.y. ago have been subject to perhaps 20 episodes of geographic shuffling with the waxing and waning of Pleistocene glaciation (e.g., Valentine and Jablonski 1993; Roy et al. 1996; Jackson and Overpeck this volume), so that the relative frequencies of geographic range overlap today probably say more about competitive interactions between close relatives than about speciation events (see also Chesser and Zink 1994). Taxa separated by major geographic barriers like the Rocky Mountains or the Isthmus of Panama are reasonable candidates for allopatric speciation, of course, but these more ancient splits are subject to the problems of extinct species more closely related to one or the other living ones—i.e., of intervening speciation events that represent the true spatial and temporal pattern of lineage splitting (e.g., Schneider 1995; Jackson and Budd 1996).

With all of these caveats, and in light of the sparse and uneven nature of the data, it is unsurprising that no clear taxonomic or environmental pattern has emerged for the distribution of evolutionary tempo and mode at the species level. Perhaps, in an obvious if unsatisfyingly context-specific hypothesis, species histories depend on their geographic extent and genetic population structure—i.e., on scale and hierarchy. If gradual anagenesis is simply the expected paleontological outcome of homogeneous directional selection, in other words, true Fisherian mass selection, then this sets some requirements on the spatial scale of gene flow relative to that of environmental variation and thus makes predictions on the distribution and genetic structure of gradualistic taxa. On the other hand, for those species that maintain genetic cohesion over different environments, or among regions with disparate selective pressures through time, the interplay of local adaptation and gene flow—intermittent or regular—will tend to impose fluctuations around a mean rather than directionality (an argument raised by Eldredge 1985, 1989; and also consistent with Futuyma 1987). Such a return to the textbook basics could explain why lineages on islands (e.g., Lister 1989) and in isolated basins (e.g., Geary 1995; and Povel 1993 in part) exhibit gradualism while related taxa in more extensive or scattered habitats often show stasis and punctuation. It also provides an approach to the presence of contrasting evolutionary patterns in co-occurring lineages, which would be unexpected if the physical environment alone (e.g., habitat stability [Sheldon 1993, 1996]) determined tempo and mode. In our present state of ignorance it may even explain the gradualistic evolution of many planktic microfossils, which may often evolve as enormous populations that occupy different depth zones in one or more otherwise relatively homogeneous oceanic water masses (e.g., Lashrus et al. 1995; but see Norris et al. 1996 and other foraminiferal references cited above). The Plio-Pleistocene vole data are, however, an apparent counterexample: Chaline and Laurin (1986) note with surprise the gradualistic trajectory of their lineage despite its likely subdivision into semi-isolated populations. This may be the exception that proves the rule, however, if the particular phenotypic changes they measured, involving increasing hypsodonty and elaboration of enamel patterning on the tooth crown, can be attributed to selection imposed by long-term vegetation changes throughout the study area.

As already noted, the relative frequency of anagenesis and cladogenesis has yet to be established. Intuitively, even excluding "pseudoextinction" (i.e., anagenetic transformation
obscured by taxonomy), species extinction rates seem to be sufficiently high that frequent branching is required for lineages to persist over geologic timescales. A number of paleontological analyses of tempo and mode that consider clades of sufficient size and phylogenetic resolution for analysis do show significant numbers of species arising cladogenetically, with stratigraphic range overlaps between putative ancestors and descendants, or between sister species (see Erwin and Anstey 1995a; Jackson and Cheetham 1999; also Stanley et al. 1988; Wagner 1998). Nonetheless, all of these references, and many more besides, also contain examples of punctuated anagenesis, so that the apparent prevalence of stasis in many situations may or may not be matched by the prevalence of cladogenesis, as required by the punctuated equilibrium model. Clearly, analyses modeled on the Cheetham and Jackson (1995) studies and focused on other groups well represented in the fossils record, say bivalves or gastropods, would be valuable. Especially useful in light of the potential role of gene flow and its relation to the spatial scale of environmental variation would be to track lineages with contrasting evolutionary tempo and mode through the Neogene fossil record to their present-day populations.

Taxon Sorting

The prevalence of intraspecific oscillatory evolution and of evolutionary stasis means that the direction of speciation is difficult to predict from within-species evolutionary trajectories. Further, wherever punctuated cladogenesis is prevalent, long-term evolutionary trends will not be simple extrapolations of intraspecific evolution but instead must involve some form of sorting among species (stepwise, punctuated anagenesis patterns are less clear-cut and might also involve sorting among populations or even highly episodic, species-wide changes propelled entirely at the organismic level). That such differential speciation and extinction rates among clades might in principle shape large-scale evolutionary patterns appears to be generally accepted (e.g., Sober 1984; Maynard Smith 1989; Williams 1992). Equally important, as Slatkin (1981) noted, differential rates can drive taxon sorting even in gradualistic systems depending on the extent of variation generated by cladogenesis and anagenesis. As in so many macroevolutionary questions the issues are the frequency of this sorting among species, the circumstances under which it occurs, and the nature of dynamics across hierarchical levels, i.e., identification of focal levels and upward and downward causation (e.g., Vrba and Gould 1986; and see Grantham 1995 for an especially thoughtful and clear review).

Species Selection and Related Processes

The term “species selection” has been used in both broad and narrow senses, sometimes by a single author. One approach, drawing on the insights of Lewontin (1970) and Hull (1980) and advocated by Vrba and Gould (1986) among others, is to maintain the neutral term “species sorting” for any pattern shaped by differential origination and extinction. Others would apply the term “species selection” here instead because fitness, i.e., differential birth and death, is being expressed at the species level, as the “emergent fitness” of species—speciation and extinction rates—within clades (e.g., Lloyd and Gould 1993; Stidd and Wade 1995; Gould and Lloyd 1999). Alternatively, species sorting can be divided into two categories depending on the hierarchical level of the characters that influence speciation and extinction rates. Then, in “effect macroevolution” differential rates are governed by organismal-level traits such as body size or habitat preferences, while in species selection the differential rates are governed by emergent, heritable properties at the species level (see Vrba 1984, 1989; Jablonski 1987; Grantham 1995).

Emergence and Heritability.—The concept of emergence in evolutionary biology has been difficult, but a simple operational approach is to recognize a feature as emergent at a given level if its evolutionary consequences do not depend on how the feature is generated at lower levels. (This approach is similar to Brandon’s 1982, 1988 application of Salmon’s 1971 statistical concept of “screening-off,” and to a parallel view, “multiple realizability,” that recently has been criticized as insufficiently pre-
cise in some circumstances; for discussion see Sober 1999; Sterelny and Griffiths 1999.) A classic example at the organismal level involves selection experiments in Drosophila where Robertson (1959) concluded that equivalent changes in wing size could be achieved either by changes in cell size or by changes in cell number, with variance in wild populations usually owing mainly to cell number, and in his experimental groups mainly to cell size (see also Stevenson et al. 1995). As the organism was the focal level of the experiment, the large-winged phenotype was the emergent property under selection, and not the cellular or genetic levels underpinning the evolutionary changes. Outside the lab, evolution of the emergent organismal property of DDT resistance is underlain by many alternative responses at the cellular level, from changes in cell walls that exclude the DDT molecule, to changes in cell metabolism that neutralize DDT when it penetrates the cell, to changes in cell physiology that sequester DDT before it can be effective (e.g., McKenzie and Batterham 1994; Feyereisen 1995).

By the same token, geographic range is an emergent property at the species level, not simply because most geographic ranges are determined by the overall distribution of conspecifics rather than by the movements of single bodies, but also because the evolutionary consequences of broad or narrow geographic ranges tend to be similar regardless of how those ranges are mediated at the organismal level (at least within broad groups, such as benthic marine invertebrates). For example, widespread species of marine gastropods are geologically longer-lived than restricted species, and the establishment and maintenance of these different ranges are statistically related to modes of larval development—an organismal trait—that differ in dispersal capabilities (Hansen 1978, 1982; Jablonski and Lutz 1983; Jablonski 1986a, 1987, 1995; Scheltema 1989, 1992; Gili and Martinell 1994; Kohn and Perron 1994). Jablonski (1987) found geographic range to be heritable at the species level (that is, closely related species showed significant correlations in the magnitudes in their geographic ranges), completing the requirements for evolution by selection at any level: the existence of heritable variation in a feature that, by interaction with the environment, imparts differential success. Cheetham and Jackson (1996) also found widespread species of bryozoans to be geologically long-lived relative to restricted species; in fact their widespread species, taken as occupying >4 regions, have a median duration of about 7.5 m.y. while the narrowly distributed species a median duration of about 2 m.y., each remarkably close to the high- and low-dispersal molluscan species, respectively, as cited above. But here the differences in geographic ranges presumably derive from the rafting of adults (e.g., Watts et al. 1998). Thus, differential taxonomic survival depends on the emergent, species-level property, i.e., the scale of the species’ range and not on the underlying organismal traits.

Genetic population structures, again not a property of single organisms, can be viewed in the same way. Jablonski (1986a, 1995) attributed high per-taxon speciation rates seen in gastropod lineages having low larval dispersal ability, as inferred from their larval shells, to their genetically fragmented populations (an argument broadly supported by genetic analyses of benthic marine invertebrates [see Pechenik 1999; Bohonak 1999]). Similarly, Wilson et al. (1975) suggested that mammals with complex social structures should have genetically more fragmented populations and thus higher speciation rates than those with more open breeding systems. And more recently, Belliure et al. (2000) found that natal dispersal ability in birds is inversely related to population differentiation and therefore, they argued, to speciation propensity. If these very different routes to highly subdivided populations yield similar macroevolutionary dynamics, this again would argue for genetic population structure as an emergent property at the clade level. The consistent relationship between dispersal ability and genetic population structure in plants (Govindaraju 1988) and animals (Bohonak 1999, in his valuable meta-analysis of 333 species across all animal groups and environments) suggests that this will be a profitable avenue for macroevolutionary research. Perhaps this general mechanism underlies the decrease in speciation rate
observed by Dodd et al. (1999) when angiosperm lineages switch from animal pollination to wind pollination, for example.

Recent molecular work has shown that even widely dispersing marine species can sometimes, perhaps usually, have subdivided rather than panmictic populations (e.g., Palumbi 1996; Geller 1998; Benzie 1999a; Avise 2000). This does not mean, however, that high-dispersal species are as readily subdivided as low-dispersal ones. The key issue is the stability and long-term evolutionary effects of that population structure relative to taxa with low dispersal abilities. The consistent relationships among larval type, geographic extent, and speciation/extinction rates in Cretaceous, Paleogene, and Neogene taxa (which appear to be robust to sampling [Jablonski 1988; Marshall 1991]) suggest that in at least some settings the population structures detected by mtDNA analysis may be transient or in any case do not have predictable macroevolutionary effects (see also the diversity of analyses tabulated by Bohonak 1999). An intriguing pattern that needs a more detailed evolutionary perspective is the discovery that genetic connectedness among Pacific populations of benthic invertebrates does not conform to present-day ocean circulation patterns but may be a Pleistocene holdover (Benzie and Williams 1997; Palumbi et al. 1997; Benzie 1999a,b). Spatial scale may also be important here: the vast but highly discontinuous environments of the Indo-West Pacific may impart a different evolutionary dynamic from that documented in the more linear shelves and the more continuous two-dimensional epicontinental seas that provided the paleontological data (see Valentine and Jablonski 1983).

Clearly, further analyses of evolutionary sorting of taxa would benefit greatly from a more detailed phylogenetic framework. Duda and Palumbi (1999) rightly note that the further analyses of such patterns in a phylogenetic context would be valuable. However, their emphasizing an evolutionary bias toward the production of species having low-dispersal larvae, rather than species sorting for the larval modes for Pacific Conus, is difficult to interpret because they lack data on extinct species and their model does not take into account empirical evidence for higher extinction and origination rates in low-dispersal lineages.

Another unresolved problem is that marine bivalves do not exhibit the same relationship between larval types and species-level dynamics as the co-occurring gastropods (e.g., Jablonski and Lutz 1983; Stanley 1990). Perhaps this is because modes of larval development in bivalves are more tightly linked phylogenetically to feeding types, body sizes, and other factors that might also influence evolutionary rates. Jablonski (1986a, 1995) showed that larval modes in marine gastropods override those of adult feeding types, and if the opposite is true for bivalves then the two groups in tandem might provide a valuable system for exploring the interplay of rate-determining traits at different hierarchical levels.

The heritability of species-level traits remains a neglected area. Jablonski (1987) and Ricklefs and Latham (1992) found geographic ranges to be heritable in marine mollusks and terrestrial plants, respectively. Their comparisons of closely related species were designed as a phylogenetic analogue to the sib-sib comparisons of quantitative genetics (and see also Peterson et al. 1999, who successfully predict geographic distributions of sister species based on a model of ecological niche conservatism). Gaston and Blackburn (1997) did not find strong species-level heritability in birds using nested ANOVAs, a very different design that also has precedents in quantitative genetics but lacks a detailed phylogenetic framework, necessarily omits extinct species and Pleistocene range adjustments, and compares taxa in different geographic situations, unlike Jablonski’s analysis, which is a macroevolutionary analogue of a common-garden experiment. More work is needed to assess the strengths and weaknesses of the different approaches and where they might be applied most robustly.

The Limits of Species Selection and Species Sorting.—The domain of strict-sense species selection, which depends on emergent characters, is much narrower than broad-sense species selection, which depends only on emergent fitnesses (i.e., differential origination and extinction rates regardless of the hierarchical level at
which they are determined [see Vrba and Gould 1986; Lloyd and Gould 1993; Grantham 1995; Stidd and Wade 1995)]. Beyond that, we simply do not know the relative frequencies of different sorting processes, overall or among clades. The theoretical literature has outstripped the empirical database, in part simply because of the scale of the databases required for rigorous analyses. However, if geographic range is arguably a species-level character, then the macroecological literature is rich in potential examples that might fit the species-selection paradigm, because so many features of living organisms can be related to geographic range and thus are candidates for hitchhiking on species sorting processes (see for example Brown 1995; Brown et al. 1996; Gaston 1998). Other components of rarity as classified by Rabinowitz (1981; Rabinowitz et al. 1986) might also be examined in this list: population sizes or densities may be emergent properties (e.g., Vrba and Eldredge 1984; and here too a large macroecological literature exists, ripe for macroevolutionary analysis, e.g., Brown 1995; Blackburn and Gaston 1997, 1999), whereas habitat specificity may reside more fully at the organismal level (e.g., Vrba 1987).

One important issue needing more attention is the stability of such species-level characteristics. Jablonski (1986b, 1987) gave evidence that marine species achieve their geographic-range magnitudes rapidly relative to their geologic durations. Tracking the magnitude, rather than the position, of geographic ranges during Pleistocene or other environmental oscillations would be interesting, as would testing for evolutionary rate differences among taxa that differ in the amplitude of their range-size changes over time. Population density should also be tested more fully for long-term stability (e.g., Arneberg et al. 1997). Both exciting and daunting is the loose covariation of geographic range, abundance, and body size (Brown 1995, 1999; Gaston and Blackburn 1999; Lawton 1999), and the question of how these effects spanning hierarchical levels and spatial scales interact, and become linked or decoupled on ecological and evolutionary timescales.

Given that the history of most lineages is evidently dominated by stasis and punctuation, other potential species-level features that might be heritable owing to factors like population sizes or genetic population structure include relative morphological inertia (and so the average duration of stasis in the phyletic mode, or the amplitude of oscillations within stasis, if these are set intrinsically) and perhaps even the size-frequency distribution of morphological divergences of daughter isolates. Sex ratios may be another example of higher-level trait, although possibly played out at an intermediate focal level if interdemic differences in sex ratios are common in some groups (e.g., Delph 1990; Graff 1999). Is the relative genetic or morphological variability of species an emergent species-level trait, or is it simply the summation of organismic properties and therefore an aggregate trait as argued by Lloyd and Gould (1993)? It depends on how that variation arises, and how sorting processes operate on that trait, and empirical work is needed here.

The potential for species sorting (= broad-sense species selection) seems extensive, given the abundant evidence for differences in intrinsic extinction and origination rates among clades (e.g., McKinney 1997; Kammer et al. 1998; Sepkoski 1998; see also “extinction risk” studies on extant organisms, e.g., Bennett and Owens 1997 and references therein). Here, too, phylogenetic hypotheses can provide a valuable framework for rigorous analysis, and methods are being developed for rigorous testing of differential origination and extinction rates in a phylogenetic context (e.g., Kirkpatrick and Slatkin 1993; Slowinski and Gayer 1993; Sanderson and Donoghue 1996; Harvey and Rambaut 1998; Paradis 1998). These methods have mostly been applied to extant taxa, where the estimation of evolutionary dynamics is made difficult by unrecorded extinction that must be ignored or assumed to be constant through time, but some also show promise for the testing of species-sorting hypotheses in the fossil record.

Species sorting, including narrow-sense species selection, will generally play a different evolutionary role from the microevolutionary sorting of organisms within populations: it will tend to determine diversity differentials
among clades rather than shape adaptations. Species sorting may not construct a complex eye or a long neck, but it may determine how many species possess complex eyes or long necks over evolutionary timescales. This has two immediate implications. First, the setting of species sorting and microevolution as rival hypotheses or mechanisms is often inappropriate. And second, the mapping of species densities in morphospace need not reflect the topology of the adaptive landscape. That is, the frequency distribution of morphologies within or among clades may not be a simple indication of relative fitnesses at the organismal level.

But is sorting at the species level unequivocally ruled out as a mechanism for organismal adaptation? Rice (1995) showed that the efficacy of species sorting in character evolution depends on the speciation rate per generation, the mutation rate, and the survival rate of reproductive isolates (and the genetic complexity of the trait under consideration). Thus, species sorting would be more likely to influence character evolution directly in organisms with long generation times; for example Asian elephants fit this model, given reasonable mutation rates and selection coefficients, and a very broad taxonomic domain is possible if speciation rates are sufficiently high (Rice 1995). Perhaps a more pervasive way for species sorting to influence the evolution of complex characters is by determining the persistence and proliferation of taxa bearing characters shaped by individual selection, so that the process of assembling those characters can proceed beyond the duration of individual species (Raup 1994; Rice 1995). This means that rate differentials among clades might be important even if, as Hallam (1998) argues, many individual clades generate only a few species per unit time.

Hierarchical analyses of large-scale changes in morphology would be valuable for improving our understanding of the relative roles of processes at different levels and the potentially complex interactions between them. The infaunalization of bivalves, the evolution of tree-like growth habits with the initial assembly of terrestrial forests, and the escalation of defensive morphologies in marine benthos during the increase in predation pressure described by Vermeij (1987) as the Mesozoic Marine Revolution (MMR) would all be possible subjects. To use the last of these biotic transitions as an example (and see also Signor and Brett 1984), long-lived bivalve or gastropod genera that appear to persist through much of the MMR, such as Aphrodina, Cyprimeria, or Solariella, might be targeted to test whether any given species shows anagenetic change in antipredatory adaptations independent of local environmental changes. (Kelley 1989, 1991, did find intraspecific trends, but within a much later, Neogene, interval.) If species tend to be static, several alternative dynamics should be tested. A mechanism below the species level might involve directional, stepwise anagenesis, shaped by widely spaced episodes of directional selection or within-species shifting-balance processes. Next up the hierarchy would be the sorting of species to convert clades to better-defended morphologies. Operating at an even higher level would be clade sorting, in which preferential extinction of poorly defended clades or preferential origin of well-defended ones might shape the overall biota. As many authors have noted (e.g., Stanley 1973; McShea 1994; P. J. Wagner 1996b), such analyses cannot focus only on mean morphologies or the appearance of derived character states; an important component of the sorting dynamics may involve the retention of less heavily defended morphologies. Finally, spatial scale will be important in understanding mechanisms, in terms of both the establishment of well-defended species—which Roy (1996) found to vary regionally in timing and which may have been promoted by environmental perturbations (see also Miller 1998)—and the fates of less derived forms (which Vermeij [1987] argues have shifted geographically or environmentally through time).

Such large-scale evolutionary interactions of predators and prey raise the broader issue of clade interactions and how they might drive taxon sorting. As Sepkoski (1996) made clear, negative clade interactions need not produce straightforward reciprocal diversity patterns (the "double wedge pattern" of Benton 1987), but instead can be manifest in com-
plex coupled logistic patterns, and perhaps most importantly depressed but still positive diversification rates (see also Miller and Sepkoski 1988). Still open to investigation is exactly how local ecological interactions—as many of these clade interactions surely must entail—actually scale up to determine origination and extinction rates. Simple intuitive arguments are easy: competition decreases population sizes and/or intrinsic growth rates, thereby making species more vulnerable to stochastic extinction; origination rates might be depressed in the same way, if biotic interactions reduce the population sizes of isolates and thus decrease their probability of surviving to achieve speciation. On the other hand, clade interactions via predation could arguably operate in the same fashion, although Vermeij (1994) explicitly rejects the role of escalation in extinction on mainly theoretical grounds. How negative ecological interactions actually determine per-taxon origination and extinction rates seems an important area for research.

Such analyses in the fossil record are complicated by the highly diffuse nature of many of the interactions under consideration: escalation in molluscan defenses is a response to the collective diversification of durophagous arthropods, mollusks, vertebrates, and other groups, so that simple responses to specific enemies will be unusual, difficult to detect, and probably fleeting in geologic terms. Targeting competitive interactions for combined micro/macrolevel analysis will also be complicated by the well-established observation that taxonomic distance is not a reliable proxy for interaction intensity (e.g., Brown and Davidson 1977; Brown et al. 1979a,b on seed-eating ants vs. rodents; Kodric-Brown and Brown 1979 on hummingbirds vs. insects; Schluter 1986 on finches vs. bees; Jackson and Hughes 1985 on spatial competition in marine encrusting communities among several phyla). Paleontologists have been aware of this—Bambach’s (1983) “megaguilds” are decidedly polyphyletic, for example—but the study of the large-scale evolutionary effects of biotic interactions remains a difficult task. One powerful approach deserving more extensive application is the projection back into the fossil record of interactions that have been experimentally dissected in modern communities (e.g., Aronson 1992, 1994; Lidgard et al. 1993; Sepkoski et al. 2000).

Incumbency.—Macroevolutionary incumbency effects, such as the damping of diversification among mammals for the first two-thirds of their history (presumably by incumbent archosaurs), are perhaps the most compelling evidence of the macroevolutionary effects of competitive interactions (Jablonski and Sepkoski 1996; Jablonski 2000a). Such effects represent very different dynamics from those underlying the reciprocal diversity patterns discussed above, of course: in double-wedge or coupled logistic models, taxon A progressively excludes taxon B by virtue of A’s competitive superiority; in incumbency interactions, taxon A excludes B by virtue of A’s ecological preemption of resources, which need not reflect competitive superiority on a level playing field (as emphasized by Rosenzweig and McCord [1991]). The double-wedge pattern is not exclusively the hallmark of the rise of a competitive dominant, however: piecemeal extinction of the dominant incumbent may create cumulative opportunities for surviving taxa, for example. Other kinds of progressive replacements might also be underlain by incumbency effects, as Valentine (1990) suggested for the Phanerozoic decline in evolutionary rates in the marine biota: over the long run, extinction-resistant clades will tend to preempt high-volatility clades (see Sepkoski 1998 for a somewhat different view).

The acceleration in evolutionary rates, taxonomically and morphologically, among previously established clades following an extinction event (e.g., Miller and Sepkoski 1988; Patzkowsky 1995; Foote 1997; McKinney 1998; Sepkoski 1998) is the primary macroevolutionary measure of incumbency effects. Considerable evidence supports a spatial analogue, in which asymmetries in biotic interactions appear to reflect regional differences in extinction intensities, both in the fossil record (e.g., Vermeij 1991) and in present-day biotas (e.g., Case 1996). However, not all invasions are mediated by prior extinction (see for example Williamson 1996; Lonsdale 1999), and at least for the end-Cretaceous extinction,
invasion intensities need not correlate to local extinction intensities (Jablonski 1998). More work is needed that tackles the difficult task of separating the effects of diversity loss per se from regional differences in environmental change; in other words, is differential diversity loss the promoter of asymmetrical biotic interchanges, or another symptom of the effects of differential environmental perturbation? Incumbency effects may represent one of the strongest bridges between ecology and macroevolution, and cross-scale and spatially explicit analyses of how those effects are maintained, broken, or circumvented would be valuable, not only for theoretical reasons, but in order to address the pressing issues of present-day biotic interchange.

Extinction Events As Filters and Promoters

So much has been written lately on the evolutionary role of extinction events that I will touch on only a few points here (see Raup 1994; Jablonski 1995, 2000a,b; Erwin 1998, 2000b for an entry into this literature). First and foremost, we still have much to learn about the role of extinction events in evolution. Perturbations occur at all intensities and spatial scales, and as discussed above they promote biotic change in important ways, but effects are difficult to predict from magnitude alone (see Miller 1998). The differential response of large-scale biotic units like clades and regional biotas to seemingly similar perturbations at different times—whether asteroid impacts or climate changes—reflects the fundamental nonlinearities that typify most complex systems. Thresholds, and especially the importance of antecedent events, are probably an essential component to a system where four sea-level oscillations cause little change in marine community composition but a fifth brings significant turnover; asteroid impacts that form craters of 50 km have no perceptible effect on the global biota but a crater of 300 km corresponds to a Cretaceous/Tertiary (K/T) scale event; and even apparently similar mass extinction intensities can have differing effects on biosphere structure and function (Droser et al. 1997).

On the other hand, as Levin (1999: pp. 180–184) notes, there is little evidence that large-scale biotic systems, let alone global biotas, evolve to the perpetual edge of collapse like the canonical sandpile, as suggested by Bak and colleagues (e.g., Sneppan et al. 1995; Solé et al. 1999). The resilience of these biotic systems to many perturbations, and their relatively loose organization as manifested in the relatively fluid composition of Pleistocene and many earlier communities, argues against a state of self-organized criticality, and more generally argues that the major turnover events in the history of life were externally driven (rather than the product of internal dynamics). Such avalanche models may apply on local scales and the short term, but hierarchy and scale defeat their universality, because the necessary biotic interconnectedness falls away rapidly at larger scales. Spatially heterogeneous dynamics can be seen, for example, in the geographic complexities of macroevolutionary and macroecological patterns like the Ordovician radiations, the marine Mesozoic revolution, the recovery from the end-Cretaceous mass extinction, and the demise of the Pleistocene megafauna (Roy 1996; Miller 1997a,b, 1998; Jablonski 1998; Martin and Steadman 1999).

The major mass extinctions probably account for less than 5% of the species turnover in the geologic past, and their disproportionate evolutionary effects probably derive from their removal of not just minor constituents of the biota, but also established incumbents. Analyses are still sparse, but contrasts between survivorship patterns during mass extinctions and those prevalent during times of low extinction intensity have been recorded for each of the major extinction events (Jablonski 1995, 2000a,b). This does not mean that certain features favored during times of “background” extinction could never also be advantageous during mass extinction events, but even partial discordances in survivorship can have profound and lasting effects, given the intensities involved. The monotonic rather than bimodal frequency distribution of extinction intensities in the geologic past (e.g., Raup 1991), and some of the similarities in survivorship among mass extinctions despite apparently different triggers, suggests that the scale of the perturbation and the operation of
threshold effects were important factors in the observed changes in selectivity.

Patterns of selectivity are difficult to assess, however, because traits can be lost as a by-product of selection on other features (or, of course, purely stochastic survivorship [Raup 1994]). For example, what was it about the end-Ordovician extinction that selected against broad selenizones in snails (P. J. Wagner 1996b), or about the end-Cretaceous extinction that selected against schizodont hinges in bivalves, elongate rostra in echinoids, or complex sutures in cephalopods? All of these losses or declines are more likely to represent correlations rather than direct causation, but they had long-term morphological effects, and additional examples are plentiful.

Finally, the long-term effects of mass extinctions are set not only by the victims and survivors of an event, but also by the dynamics of the recovery process. Here, of course, is where the incumbency effects are most vividly illustrated, by the diversifications that unfold with the removal of previously dominant taxa. As discussed by Jablonski (2000b), however, clade dynamics across extinction events can exhibit several different patterns, including (1) unbroken continuity, as in the escalation of antipredatory defenses in marine bivalves across the K/T boundary (Hansen et al. 1999); (2) continuity with setbacks, as in the increase in cheilostome bryozoan abundance relative to cyclostomes across the K/T boundary (McKinney et al. 1998) or the increase in suture complexity in Paleozoic ammonoids (Saunders et al. 1999); (3) failure to rebound and eventual extinction, as in the demise of the prolecanitid ammonoids in the Early Triassic after surviving the end-Permian debacle (Page 1996); or (4) unbridled diversification, as seen most famously in the radiation of mammals after the end-Cretaceous demise of the dinosaurs and other reptilian groups. The apportionment of survivors among these different trajectories needs much more analysis before we can understand the evolutionary roles of extinction events.

That said, mass extinctions have lasting effects across many scales. As Erwin (1998) argues, the Permo-Triassic extinction permanently restructured marine and terrestrial ecosystems, so that the raw material for microevolution and the web of potential biotic interactions was profoundly shaped by the taxonomic losses of that time (although Sepkoski [1996, 1998] held that those changes were inevitable given differential turnover among groups, and were merely hastened by the mass extinction). Similarly, the end-Cretaceous extinction removed the rudists and nearly exterminated the trigonioid bivalves, but its lasting influence can also be seen in the age distribution of living marine bivalve genera, which shows a secondary peak corresponding to the early Cenozoic. This 60- m.y. evolutionary echo is less dramatic than the mammalian rise to dominance that began at the same time, but it reflects a similar evolutionary process (Flessa and Jablonski 1996).

**Conclusion**

The relation between micro- and macroevolution is complementary and not mutually exclusive, with effects cascading both upwards and downwards over long timescales. The conceptual expansion of evolutionary biology with the advances in our understanding of the origin and sorting of variation has benefited many disciplines and is promoting a fuller integration across those scales and hierarchical levels. I have touched only indirectly on many of the important methodological advances discussed elsewhere in this issue: in phylogeny estimation, in dissection of developmental processes, in calculation of confidence limits on diversification rates and patterns, and in quantifying the relation of paleontological patterns to the fabric of the stratigraphic record. These developing methodologies, and the host of new questions that are emerging at the interface of biology and geology, will provide rich research opportunities for the next 25 years of Paleobiology.

**Acknowledgments**

This paper benefited from discussions with many colleagues, most recently A. K. Behrensmeier, B. Chernoff, D. H. Erwin, M. Foote, T. A. Grantham, S. M. Kidwell, K. Roy, A. B. Smith, J. W. Valentine, P. J. Wagner, S. L. Wing, and, in the final throes, the NCEAS Working Group on Ecological Processes and Evolution-
ary Rates (P. M. Brakefield, N. Eldredge, S. Gavrilets, J. B. C. Jackson, R. E. Lenski, B. S. Lieberman, M. A. McPeek, W. Miller, and J. N. Thompson); few if any of these people agree fully with the views expressed here. D. J. Futuyma, S. M. Kidwell, D. W. McShea, and J. W. Valentine provided valuable reviews. I thank the editors of Paleobiology for inviting me to write this paper, and for their patience during its elephantine gestation period; Natasha Atkins significantly improved the clarity and accuracy of the text. Supported by the National Science Foundation and a John Simon Guggenheim Memorial Fellowship.

Literature Cited

Bradshaw, H. D., K. G. Otto, B. E. Frewen, J. K. McKay, and D.
raminiferal small subunit ribosomal RNA genotypes and their potential as tracers of present and past ocean circulations. Paleoceneanography 14:3–12.


—. 2000a. Macroevolution is more than repeated rounds of microevolution. Evolution and Development 2:78–84.


Frost, D. R., and A. G. Kluge. 1994. A consideration of episte-


MICRO- AND MACROEVOLUTION


Stanley, S. M., K. L. Wetmore, and J. P. Kennett. 1988. Mac-roevo-

dolutionary differences between two major clades of Neo-


versity Press, Oxford.

Stebbins, G. L. 1982. Perspectives in evolutionary theory. Evo-

lication 36:1109–1119.


within Phylloctis (Rodentia: Sigmodontinae): concordance be-

tween mtDNA sequence and morphology. Journal of Mamm-

alogy 79:573–593.

Sterelny, K., and P. E. Griffiths. 1999. Sex and death: an intro-

duction to philosophy of biology. University of Chicago Press,

Chicago.

Stevenson, R. D., M. F. Hill, and P. J. Bryant. 1995. Organ and

cell allometry in Hawaiian Drosophila: how to make a big fly.


Stidd, B. M., and D. L. Wade. 1995. Is species selection depen-

dent upon emergent characters? Biology and Philosophy 10:

55–76.

Tabachnick, R. E., and F. L. Bookstein. 1990. The structure of in-

dividual variation in Mioocene Globorotalia. Paleobiology 44:

416–434.

Templeton, A. R. 1989. The meaning of species and speciation:


———. 1998. Species and speciation: Geography, population

structure, ecology, and gene trees. Pp. 32–43 in Howard and

Berlocher 1998.

Theissen, G., A. Becker, A. Di Rosa, A. Kanno, J. T. Kim, T. Mün-

ster, K.-U. Winter, and H. Saelder. 2000. A short history of

MADS-box genes in plants. Plant Molecular Biology 42:115–

149.

Theriot, E. 1992. Clusters, species concepts, and morpho-


olution of exploitation of design options by the first animals


versity Press, New York.

Trussell, G. C., and L. D. Smith. 2000. Induced defenses in re-

sponse to an invading crab predator: an explanation of his-

torical and geographic phenotypic change. Proceedings of the

National Academy of Sciences USA 97:2123–2127.

Turelli, M. 1988. Phenotypic evolution, constant covariances,

and the maintenance of additive variance. Evolution 42:1342–

1347.

Valentine, J. W. 1969. Taxonomic and ecological structure of the

shelf benthos during Phanerozoic time. Palaeontology 12:

684–709.

———. 1973. Evolutionary paleoecology of the marine bio-

sphere. Prentice-Hall, Englewood Cliffs, NJ.

———. 1980. Determinants of diversity in higher taxonomic cat-


———. 1990. The fossil record: a sampler of life’s diversity. Phil-

osophical Transactions of the Royal Society of London B 330:

251–268.

———. 1995. Why no new phyla after the Cambrian? Genome


sea: General patterns and biogeographic controls. In R. W.

Simms, J. H. Price, and P. E. S. Whalley, eds. Evolution, time


———. 1993. Fossil communities: compositional variation at

many time scales. Pp. 341–349 in R. E. Ricklefs and D. Schlu-

ter, eds. Species diversity in ecological communities: histori-

cal and geographical perspectives. University of Chicago Press,

Chicago.


ecules and embryos: new perspectives on the Cambrian exp-


Vermeei, G. J. 1987. Evolution and escalation. Princeton Univer-

sity Press, Princeton, N.J.

———. 1991. When biotas meet: understanding biotic inter-


———. 1994. The evolutionary interaction among species: se-

lection, escalation, and coevolution. Annual Review of Ecol-


Via, S., and D. J. Hawthorne. 1998. The genetics of speciation:

promises and prospects of Quantitative Trait Locus mapping.


Via, S., and A. J. Shaw. 1996. Short-term evolution in size and


jor gene effect: paedomorphosis in the Mexican axolotl. Pro-

ceedings of the National Academy of Sciences USA 94:14185–

14189.

Vrba, E. S. 1984. What is species selection? Systematic Zoology

33:318–328.


histories of Miocene-Recent mammal clades. Evolutionary

Ecology 1:283–300.

———. 1989. Levels of selection and sorting with special ref-

erence to the species level. Oxford Surveys in Evolutionary Bi-


Vrba, E. S., and N. Eldredge. 1984. Individuals, hierarchies and

processes: towards a more complete evolutionary theory. Pa-

leobiology 10:146–171.

Vrba, E. S., and S. J. Gould. 1986. The hierarchical expansion of

sorting and selection: sorting and selection cannot be equat-


Wagner, G. P. 1996. Homologues, natural kinds and the evolu-


regulator of evolvability? Journal of Experimental Zoology

(Molecular and Developmental Evolution) 285:116–118.

Wagner, P. J. 1995. Testing evolutionary constraint hypotheses:

examples with early Paleozoic gastropods. Paleobiology 21:

248–272.

———. 1996a. Ghost taxa, ancestors, and assumptions: A reply


———. 1996b. Contrasting the underlying patterns of active

———. 1998. A likelihood approach for evaluating estimates of

phylogenetic relationships among fossil taxa. Paleobiology

24:430–449.

tests of speciation models. Pp. 87–122 in Erwin and Anstey

1995b.

Walis, J. A. 1996. No second chances? New perspectives on bi-


281–288 in P. Copper and J. Jin, eds. Brachiopods. Balkema,

Rotterdam.


anthropogenic dispersal mechanisms in the marine environ-

ment: a study using cheilostome Bryozoa. Philosophical