The science of life is a superb and dazzlingly lighted hall, which may be reached only by passing through a long and ghastly kitchen.

– Claude Bernard

CHAPTER 1

Macroevolution: The Problem and the Field

The Process and the Field of Macroevolution

The return of macroevolution. The field of macroevolution embraces the excitement of seeking an understanding of the breadth of life. We have long desired to know how best to describe the diversity of life’s forms and to explain how and why this diversity came to be. No mystery is more intriguing than why we have amoebas and horses, or dandelions and palms. The child’s first walk in a meadow, when the child sees flowers and butterflies for the first time, can inspire the same wonder in the most sophisticated biologist walking those same tracks many years later.

We return to this perspective from many quarters of biology and paleontology, after many decades of asking far more restrictive questions that tended to put the process of evolution under a microscope. But now we are stepping back, to take in the broader view. The advances in molecular genetics and developmental biology in recent years have only increased our confidence that the nature of living systems can be understood mechanistically; we can now imagine the possibility of describing the difference between organisms in terms of their genes, gene products, and spatial organization. Such descriptions were beyond our grasp even 10 years ago, but now they are at hand, if still in fragments. The large-scale collation of fossil data and a new understanding of the history of the earth have brought similar increases of confidence among geologists and paleontologists. But we should not overlook some significant changes in fields such as systematics, and the crucial groundwork in population biology established through the advances of the neo-Darwinian movement and the Modern Synthesis. All these place us in position to answer questions that could not even be asked very seriously just a few decades ago.
Definition of the Process of Macroevolution

I define macroevolution to free it from any dependence on specific controversies and, more importantly, to define a field derived from tributaries that have merged from many sources. I define the process of macroevolution to be (Levinton 1983) the sum of those processes that explain the character-state transitions that diagnose evolutionary differences of major taxonomic rank. This definition of macroevolution focuses on character-state differences (defined in chapter 2) rather than on jumps, for example, from one taxon to another of great distance. The definition is noncommittal to any particular taxonomic level. I believe that one should eschew definitions of macroevolution such as (1) evolution above the species level (e.g., Eldredge and Cracraft 1980; Stebbins and Ayala 1981) or (2) evolution caused by speciation and selection among species (e.g., Stanley 1979). These definitions presume that major transitions can be analyzed properly only by examining speciation and other processes occurring at the species level and above, and they restrict our views toward alternative hypotheses. Worse than that, these definitions ignore the forest of organismal phenotypic breadth and focus on the trees of just one component of that breadth.

It is not useful to distinguish sharply between microevolution and macroevolution, as I will show in this volume. The taxonomic rank marking any dichotomy between microevolution and macroevolution would depend on the kind of transition being studied. Our impression of “major” degrees of evolutionary change is inherently qualitative and not fixed at any taxonomic rank across all major taxonomic groups. This is apparent when we consider transitions whose importance may rely on many characters, or just one. For the cichlid fishes, a synarthrosis between the lower pharyngeal jaws, a shift of insertion of the fourth levator externus muscles, and the development of synovial joints between the upper pharyngeal jaws and the basicranium may be necessary (but not sufficient) for the morphological diversification of species with differing food collection devices (Liem 1973). On the other hand, the evolution of the mammals involved a large number of integrated physiological and morphological traits, and these were acquired over a long period of time (Kemp 1982). Yet both fall well within the province of macroevolutionary change, because of the potential at least for evolutionary differences spanning large chasms of taxonomic rank.

A second reason for an unrestricted definition of the taxonomic level required to diagnose macroevolutionary change is the variation in higher level taxonomic splitting among major groups (Van Valen 1973a). There is no simple way of drawing an equivalence between families of mammals and mollusks; comparisons of rates of evolution between groups at “comparable” taxonomic levels (e.g., Stanley 1973a) are therefore usually invalid (Levinton 1983; Van Valen 1973a). This point is illustrated well by qualitative studies on hybridity and genetic and phenotypic distance within groups of species of similar taxonomic distance from different phyla. The taxonomist tends to use a qualitative threshold of phenetic difference to define significant evolutionary distance. Thus the ferret and the stoat were placed in different genera, even though they hybridize and produce fertile offspring. Crosses between congeneric species of frogs, however, do not usually produce viable, let alone fertile, offspring.

Perhaps the most unfortunate influence of taxonomic level in restricting our freedom in studying macroevolution is the presumption that crucial characters define
specific taxonomic levels. This approach is a major organizing force for systematics today, despite the several decades since the 1970s when cladistic approaches have taken a more pluralistic view of the role of characters in defining evolutionary groups (clades) with common ancestry (see Chapter 2). This permeating influence derives from Cuvier’s important notion of subordination of characters, which has survived through the centuries and has led systematists to accept the idea that specific traits define major taxonomic levels. Such thinking leads to unfortunate ideas as the “origin of orders,” even though such a taxonomic level has been defined by an arbitrary character type.

The difficulty of gauging macroevolution by taxonomic distance is exacerbated by our current ignorance of the relationship between morphological and genetic divergence among distantly related taxa. By what proportion of the genome do chimpanzees and humans differ? Despite our available estimates of genetic differentiation from sequenced DNA and protein amino acid sequences, allozymes, and karyotypes, we cannot draw a parallel with our knowledge of morphological differences. We are crippled by this ignorance when seeking to judge how “hard” it is for evolutionary transition to take place. What is our standard of difficulty? Genetic? Functional morphological? Developmental? Worse than that, what if interactions among these three occur? At this point, we cannot even easily inject the notion of time in evolution. We may be able to estimate rates of change of a variety of entities (e.g., DNA sequence, body size, and the like), but we have no idea of whether evolution of a complex morphology, such as the rise of mammals, would be astonishing if it happened in one million years, or dizzyingly slow! If the Cambrian Explosion of eumetazoan life occurred in 10 million years, can we say that this was blazing speed or just an ordinary pace? We do not know.

My last justification for a definition based on genetic and phenotypic breadth is that it permits an expansion of previous evolutionary theory to embrace the larger-scale hierarchical processes (see below) and higher-level taxonomic variations previously ignored by the bulk of evolutionary biologists, except in passing or in gratuitous extrapolation from lower taxonomic levels of concern. It is my hope that my definition will eventually not be needed and that “macroevolution” will merge with “microevolution” to become a discipline without a needless dichotomy. The need for a discipline of macroevolution, in my view, is more to sell the expansion of approaches than to necessarily dismiss any previous theory.

The Scope of Macroevolution

The discipline of macroevolution should include those fields that are needed to elucidate the processes involved in accomplishing the change from one taxonomic state to another of significant distance. Macroevolutionary studies all must be organized around several basic questions:

1. How do we establish the phylogenetic relationships among taxa? What is the nature of evolutionary novelty and how do novel characters define the taxa we delineate?
4

GENETICS, PALEONTOLOGY, AND MACROEVOLUTION

2. How do genetic, developmental, and morphological components channel the course of morphological and genetic evolution?
3. What are the patterns of change and what processes regulate the rate of evolutionary change from one character state to another?
4. What environmental changes regulated the timing of evolutionary radiations and extinctions?
5. What is the role of extinction in the evolutionary potential of newly evolved or surviving groups?
6. What ecological processes regulate morphological and species diversity? To what degree do these effects have evolutionary consequences for any given group?

In the following chapters, I will try to support the following assertions:

1. Systematics is the linchpin of macroevolutionary studies. Without an acceptable network of phylogenetic relationships, it is impossible to investigate the possible paths of major evolutionary change (chapter 2).
2. The nature of evolutionary novelty is probably the most studied and still the most confused element of evolutionary biology. The presence of discontinuity in morphological state can be explained readily using the available data and theory of genetics (chapters 3 and 4). The mechanisms behind the discontinuities are more poorly understood and may relate to a complex interaction between genetic and developmental processes (chapter 4). The epigenetic processes are also subject to genetic control, and thus a spectrum of resultant morphologies can be discontinuous.
3. There is no evidence that morphological evolution is accelerated or associated with speciation, except as an effect of ecologically unique circumstances leading to directional selection. Intraspecific variation during the history of a species is the stuff of interspecific morphological differentiation (chapter 3). When it occurs, intraspecific stasis is affected mainly by gene flow, at a given time and stabilizing selection, over time.
4. Many genetic and epigenetic aspects of development are conserved in evolution. Early development is especially characterized by the use of widely conserved transcription regulators and other regulatory genes. Development, however, is widely labile, as is the order of appearance of expression in developmental genes. Although the expression of developmental genes can be used to trace homologies in closely related forms, developmental genes are a conservative set of elements that can be expressed radically differently in different organisms. Developmental genes are like the musical notes, and the organisms are like rock music, blues, and baroque music. This suggests that there are no profound constraints restricting evolutionary change. Nevertheless, certain early patterns of gene expression were incorporated early in animal evolution and were retained (chapter 5).
5. The nature of form is best understood within the framework of Adolph Seilacher’s concept of Constructional Morphology. Constructional, Phylogenetic-Developmental, and Functional Morphological factors interact to determine form. This combination tends to make evolutionary pathways often eccentric and not conducive to predictions from “ground up” engineering approaches to optimality. Once historical constraints are recognized, however, optimality approaches can be used to gauge the performance of alternative morphotypes. Indeed, without such an approach, studies of adaptation would be vacuous (chapter 5).
6. Having understood the nature of variation, we find little evidence that the fossil record consists of anything more than the standard variation within populations that can be studied by evolutionary biologists. The process of macroevolution need not invoke paroxysmal change in genetics or morphology. The genetic basis of morphological change, nevertheless, involves a considerable variety of mechanisms. Morphological evolution is not the necessary consequence of speciation, though it may be a cause of speciation (chapters 3 and 6).

7. Bauplans are evolved piecemeal. Trends leading to complex forms consist of a large number of specific changes acquired throughout the history of the origin of the derived bauplan (chapter 6). Subsequently, however, stability is common. Some trends, such as a general increase in invertebrate predator defense and reductions in variation of morphologies, are probably due, to a degree, to the selective success and extinction of different taxa. Even though speciation rate is not related causally to the origin of the novelty, intertaxon survival, sometimes due to random extinction, has been a crucial determinant of the present and past complexion of the biotic world (chapter 7).

8. Although earth history has had a clear impact on diversification and standing diversity, patterns of taxonomic longevity may have had a distinctly random component. Major differences in biology may have consequences for rates of morphological evolution and speciation, but patterns of distribution within these groups may reflect random appearance–extinction processes (chapter 7).

9. Mass extinctions and radiations are a fact of the fossil record. But both are more easily recognized by changes in the biota than by any recognizable physical events. Means of distinguishing among current hypotheses of regulation of mass extinction and radiation are equivocal at best (chapter 7).

10. The Cambrian Explosion may have involved two phases. Molecular evidence suggests that the major animal groups diverged, perhaps as small-bodied forms or even as ciliated larvalike forms, about 800 to 1,000 million years ago. The sudden appearance of larger skeletonized body fossils and burrows at the beginning of the Cambrian is probably more of an ecologically driven event reflecting the evolution and radiation of crown groups (the modern phyla), rather than a time when the defining traits of the triploblastic metazoa arose, which was probably long over by Cambrian times (chapter 8).

Is macroevolution something apart from microevolution? Richard Goldschmidt instigated the dichotomous approach to macroevolution when he conceived of hopeful monsters that arose by means of speciation events (see below under Hierarchy and Evolutionary Analysis). The modern version of this beginning pictured a decoupling of microevolution from macroevolution (e.g., Stanley 1975), with the species level being the barrier through which any macroevolutionary change must penetrate. Although the specific notion of macromutations is restricted to only a few macroevolutionists (e.g., Gould 1980a), the notion of an evolutionary breakthrough has been associated with speciation events and their frequency. This point of view has made for an unfortunate battle royal, where victory would mean that the opposing group was irrelevant in evolutionary biology. If the microevolutionists win, then there is no such thing as macroevolution. If the macroevolutionists gain favor, then microevolution exists, but it is a minor part of a much larger set of
evolutionary constructs. Macroevolutionist claims began by relegating microevolution to the ash heap of history (e.g., Gould 1980a). It made for great sound bites. Subsequent arguments have softened, only emphasizing the expansion of evolutionary theory offered by macroevolutionary considerations (Gould 1982a).

Is the dichotomy very useful? For one group to “win” conveniently ensures the irrelevance of the other to major contributions in evolutionary theory. The focus of this argument is at the speciation threshold of evolution. But I hope that the reader realizes already that there is much more to paleontological and neontological macroevolutionary arguments than the nature of speciation.

The focus of macroevolution. Macroevolution must be a field that embraces the ecological theater, including the range of time scales of the ecologist, to the sweeping historical changes available only to paleontological study. It must include the peculiarities of history, which must have had singular effects on the directions that the composition of the world’s biota took (e.g., the splitting of continents, the establishment of land and oceanic isthmuses). It must take the entire network of phylogenetic relationships and superpose a framework of genetic relationships and appearances of character changes. Then the nature of constraint of evolutionary directions and the qualitative transformation of ancestor to descendant over major taxonomic distances must be explained.

The macroevolutionary foci I mention have been largely ignored by the founders of the Modern Synthesis in the past 50 years, who have been devising theories explaining changes in gene frequencies or small-scale evolutionary events, leaving it to someone else to go through the trouble of working in larger time scales and considering the larger historical scale so important to the grand sweep of evolution within sight of the horizon of the paleontologist. The developmental/genetic mechanisms that generate variation (what used to be called physiological genetics) have also been neglected until recently. Population geneticists assume variation but do not study how it is generated nearly as much as they worry about the fate of variation as it is selected, or lost by stochastic processes.

Evolutionary biology and astronomy share the same intellectual problems. Astronomers search the heavens, accumulate logs of stars, analyze various energy spectra, and note motions of bodies in space. A set of physical laws permits interpretations of the present “snapshot of the universe” afforded by the various telescopic techniques available to us. To the degree that the physical laws permit unambiguous interpretations, conclusions can be drawn about the consistency of certain observations with hypotheses. Thus, rapid and cyclical changes in light intensity led to the proof of the reality of pulsars. The large-scale structure of the universe inspired a more historical hypothesis: the big bang origin of the universe.

Does the evolutionary biologist differ very much from this scheme of inference? A set of organisms exists today in a partially measurable state of spatial, morphological, and chemical relationships. We have a set of physical and biological laws that might be used to construct predictions about the outcome of the evolutionary process. But, as we all know, we are not very successful, except at solving problems at small scales. We have plausible explanations for the reason why moths living in
industrialized areas are rich in dark pigment, but we don’t know whether or why life arose more than once or why some groups became extinct (e.g., the dinosaurs) whereas others managed to survive (e.g., horseshoe crabs). Either our laws are inadequate and we have not described the available evidence properly or no laws can be devised to predict uniquely what should have happened in the history of life. It is the field of macroevolution that should consider such issues. For better or worse, macroevolutionary biology is as much historical as is astronomy, perhaps with looser laws and more diverse objectives. If history is bunk, then macroevolutionary studies are … well, draw your own conclusions!

Indeed, the most profound problem in the study of evolution is to understand how poorly repeatable historical events (e.g., the trapping of an endemic radiation in a lake that dries up) can be distinguished from lawlike repeatable processes. A law that states an endemic radiation will become extinct if its structural habitat disappears has no force because it maps to the singularity of a historical event. It is how we identify such events that matters. What we cannot do is infer that all unexplainable phenomena arise from such unique events. For example, if we postulate natural selection as the shaping force of all morphological structures, it is a cop-out to relegate all unexplainable phenomena as arising from unique historical events.

Hierarchy and evolutionary analysis. We need a context within which to study macroevolution. J. W. Valentine (1968, 1969) first suggested to paleontologists that large-scale evolutionary studies should use a hierarchical framework (e.g., Allen and Starr 1982; Eldredge 1985; Gould 1982a; Saltie 1985; Vrba and Eldredge 1984; Vrba and Gould 1986).

I use hierarchy in the sense of a series of nested sets. Higher levels are therefore more inclusive. There are at least two main hierarchies that we must consider: organismic-taxonomic and ecological. The organismic-taxonomic hierarchy can be ordered as:

{molecules → organelle → cell → tissue → organ → organism → population → species → monophyletic group}

A variant of this hierarchy would include the substitution of gene → chromosome → organism at the lower end. The ecological hierarchy would include: organism → population → community. There is no necessary correspondence, however, between levels of the ecological and organismic-taxonomic hierarchies.

Hierarchies can be used either as an epistemological convenience or as a necessary ontological framework for evolutionary thought. Both approaches have been taken in the past, sometimes within the same hierarchy. The standard taxonomic hierarchy is used commonly as a means to examine rates of appearance and extinction. Although different taxonomic levels may change differently over time, such studies do not assign special significance to these levels, as opposed to another set of levels that might also be studied (e.g., studying species, subfamilies, and families, as opposed to species, families, and orders). They are just conveniences whose ascending order of ranking may correlate with differences of response (e.g., Valentine 1969). On the other hand, some regard certain taxonomic levels as fundamental and
of ontological significance. Van Valen (1984) sees the family level as a possible unit of adaptation. The species has been claimed to have great importance (Eldredge and Gould 1972). I and most neo-Darwinians see the organism as a fundamental level of the hierarchy, around which all other processes turn. If a given taxonomic level has meaning, it is because the traits of an organism can be traced to this taxonomic level.

If all processes could be studied exclusively with the smallest units of the hierarchy, then two conclusions would readily follow. First, it would not be necessary to study higher levels (i.e., there would be no macroscopic principles). Second, higher levels would be simple sums of the lower ones, with no unique characteristics of their own. The first principle might lead a geneticist to claim that once genes are understood, the entire evolutionary process could be visualized as gene–environment interactions, with no consideration of the properties of cells, organisms, species, or monophyletic groups. The second might lead a paleontologist to argue that patterns of ordinal standing diversity are a direct reflection of species diversity (e.g., Sepkoski 1978).

Taking the hierarchy as given, we can ask the following questions:

1. Can one learn about the higher levels from the lower?
2. Can one understand processes at a given level without resorting to knowledge of other levels?
3. Is there any principle of interaction among levels, such as unidirectional effects exerted by lower levels on higher levels (e.g., those of genes on individual survival) but not the reverse (the effect of survival of individual organisms on the future presence of the gene)?

The first question raises the issue of reductionism, a major area of controversy in biology (e.g., Ayala and Dobzhansky 1974; Dawkins 1983; Lewontin 1970; papers in Sober 1984a; Vrba and Eldredge 1984; G. C. Williams 1966, 1985; Wimsatt 1980). It is a common belief that all aspects of biological organization can be explained if the entire genome were sequenced and all the nature and sequence of all proteins were known. In parallel with this argument, several biologists have proposed the gene as the unit of selection and the primary target of understanding. A theory at the level of the gene would then be extrapolated to a theory of the entire genome. In one case (G. C. Williams 1966), the claim was a healthy antidote to the proposal that certain forms of evolution can be explained only at another level of the hierarchy, the population (e.g., Wynne-Edwards 1962).

Although reductionism is often an object of scorn among evolutionary biologists (Wimsatt 1980, Gould 1982b), there seems to be much confusion about definitions. At least three concepts are often freely intermixed. First, reductionism may imply a reducing science, which can explain all phenomena in terms of a set of basic laws and units. In this conception of reductionism, biological constructs such as species, cells, and amino acids could be described completely in terms of the language and laws of physics. In evolutionary biology, the language and processes of Mendelian genetics might be substituted by the language and processes of molecular biology (Schaffner 1984). Second, reductionism is often used to imply atomism, where all
phenomena of a science can be described effectively by laws involving the smallest ontological units. Thus, one might claim that the extinction of the dinosaurs could be explained with knowledge of their nucleotide sequences only. This is the type of reductionism often under attack by macroevolutionists (e.g., Gould 1983b; Vrba and Eldredge 1984). Some (e.g., Wimsatt 1980) attack reductionism as an impractical attempt to explain phenomena in terms of the smallest ontological units of a science. This does not imply that it is impossible to do so, only that it is so difficult that higher constructs of a hierarchy are more practical (Nagel 1961). This argument can also be made when, for adequate description of another science the use of a reduced science requires a myriad of complexities in language (e.g., translating Mendelian genetics into molecular genetics [Hull 1974]).

The confusion of these types of reductionism makes debate quite difficult. For example, geneticist Richard Goldschmidt was a reductionist of the reducing science kind (G. E. Allen 1974), even if he is remembered for immortalizing the distinct break of the species level. He believed that chromosomal effects could be reduced to physical laws. Yet, Vrba and Eldredge (1984) placed him on the side of holism. As another example, Wimsatt (1980) criticized the reductionist program, but only because it is impractical to explain many phenomena. From this argument alone, it would not be clear that he would reject the other two types of reductionism, if his objections to workability could be addressed. On the other hand, others find that certain levels have emergent properties, which are irreducible to lower levels of a hierarchy. This opinion, presumably, would also apply if a reducing science were available. In other words, if physics could subsume all biological processes, such individuals would criticize physics if it were atomistic. The attraction of both atomistic and reducing-science reductionism rests in their sweeping approach at explanation. If all scientific explanation could be accomplished with some minimal-level constructs in a single science, then we could achieve an essentially universal language. Keats decried Newton for reducing the poetic elegance of the rainbow to its vulgar prismatic colors. If, however, such a reduction were possible, then grouping concepts such as the rainbow would be superfluous. But can we find such basic elements and a set of relational laws in science? Do we find emergent properties in higher hierarchical levels that cannot be defined in a language derived from the lower levels?

The dream of reductionism has never been achieved, nor does it seem likely that we will explain all by resorting to explanations using only the basic elements (Popper 1974). As we study different geometries, we learn that the detail lost in switching from Euclidean geometry to topology is superseded by whole new concepts that were never previously visible (Medawar 1974). In Euclidean geometry, shape is invariant and transformations and comparisons are based on angles, numbers of sides, and curvature about foci; topology ignores exact shape but maintains a sense of space and linear order. The transition from the former geometry to the latter involves a restriction of detail, but new concepts emerge. Thus, the notion of conic sections appears in the geometry of projection.

In evolutionary biology, the gene is often employed as the smallest unit of consideration, though recent discoveries of molecular genetics muddle this a bit.
Population genetics usually sees the fate of genes in terms of their contributions to fitness and stochastic processes. Complexities of genetic structure, such as epistasis and linkage, greatly complicate population genetic models. Yet it is a legitimate pursuit to ask how genes survive by virtue of their effects on the phenotype, although one might question the power of both our empirical tools and multilocus models to realistically attack population genetic problems (e.g., Lewontin 1974).

Most evolutionary biologists acknowledge a great deal of complexity in the effects of single genes on the phenotype and emphasize the complex interactions among genes. Most adhere to the principle that the organism, and not the gene, is the unit of selection (e.g., Dobzhansky 1970). The integrity of the organism and its internal interactions have been emphasized by Dobzhansky (1951), Lerner (1954), and Stebbins (1974), among others. Consider Stebbins’s statement (1974, p. 302) of the limited evolutionary potential of the incorporation of new alleles:

Mutations that affect these structures and processes have an adaptive value not in direct connection with genotype–environment interactions, but through their interactions with other genes that contribute to the structures or processes involved. In higher organisms, the majority of genes contribute in one way or another to these conserved structures and processes. The adaptive value, and hence the acceptance or rejection by natural selection of most new mutations, depends not upon direct interactions between these mutations and the external environment, but upon their interaction with other genes, and their contribution to the adaptedness of the whole organism.

This is not an appeal to mysticism. Stebbins merely acknowledged that genes serve to determine a functioning phenotype in a complex manner. Genes may very well be retained by virtue of their contributions to fitness, but there is an important hierarchical level, the organism, that also shapes the fabric of genetic organization. The organism is not the simple sum of its parts. It may well be that division of labor in some Hymenoptera serves the purpose of the survival of genes, but the phenomenon of labor division cannot be explained from the genes’ mere presence.

The notion of levels is well entrenched within evolutionary biology, but the exact awareness of levels is not always present when evolutionary hypotheses are formulated. The effects of individual genes on fitness can be overshadowed by other processes, which are best considered as interactions of higher levels of the hierarchy with lower levels. Consider the many studies of regional gene frequency clines discovered by students of allozyme polymorphisms over the past few decades. Typically, one samples over a geographic–environmental gradient and finds a spatially progressive change in allele frequency at a locus (e.g., Adh for Drosophila). The distribution and abundance of the variant alleles have been studied by those interested in the question of natural selection. There is almost universal agreement that if the functional differences among allozymes could be related to fitness, then the problem of geographic variation would be solved. But is this true?

Effects within an evolutionary hierarchical system can be transmitted downward (Campbell 1974). For example, consider a step cline that transects a continent, with allele \( a \) nearly fixed in the east whereas \( b \) is fixed in the west. Suppose that a dramatic change in structural habitat (e.g., loss of the species’ requisite food plant) drives to extinction the entire western part of the species. Owing to stochastic loss, the