

Chance and necessity: the evolution of morphological complexity and diversity

Sean B. Carroll

Howard Hughes Medical Institute and Laboratory of Molecular Biology, University of Wisconsin–Madison, 1525 Linden Drive, Madison, Wisconsin 53706-1596, USA (e-mail sbcarroll@facstaff.wisc.edu)

The primary foundation for contemplating the possible forms of life elsewhere in the Universe is the evolutionary trends that have marked life on Earth. For its first three billion years, life on Earth was a world of microscopic forms, rarely achieving a size greater than a millimetre or a complexity beyond two or three cell types. But in the past 600 million years, the evolution of much larger and more complex organisms has transformed the biosphere. Despite their disparate forms and physiologies, the evolution and diversification of plants, animals, fungi and other macroforms has followed similar global trends. One of the most important features underlying evolutionary increases in animal and plant size, complexity and diversity has been their modular construction from reiterated parts. Although simple filamentous and spherical forms may evolve wherever cellular life exists, the evolution of motile, modular mega-organisms might not be a universal pattern.

“Drawn out of the realm of pure chance, the accident enters into that of necessity, of the most implacable certainties.”
J. Monod¹

It is widely accepted that the evolution of any particular organism or form is a product of the interplay of a great number of historical contingencies¹. Rewind and replay the tape of life again and again, as the now familiar argument goes, and there is no predicting (or reproducing) the outcomes. Roses and redwoods, humans and hummingbirds, trilobites and dinosaurs each owe their existence (or demise) to unfathomable combinations of innumerable rolls of the ecological and genetic dice.

Life's contingent history could be viewed as an argument against any direction or pattern in the course of evolution or the shape of life. But it is obvious that larger and more complex life forms have evolved from simple unicellular ancestors and that various innovations were necessary for the evolution of new means of living. This raises the possibility that there are trends within evolutionary history that might reflect the existence of general principles governing the evolution of increasingly larger and more complex forms. The first task of this review is to examine the degree to which the evolution of the shapes of life are a matter of chance — a random walk in morphospace — or of necessity — borne from the demands of natural selection and the constraints imposed by physics, genetics and development. The second task is to extrapolate from the evolutionary trends on Earth to assess what they might portend for the evolution of life elsewhere.

There is a long history of support for the general notion of overall evolutionary trends towards increases in size², complexity^{2,3} and diversity^{4–6}. However, there are two fundamentally distinct mechanisms that have been proposed to explain these trends⁷. One is a random, passive tendency to evolve away from the initial minima of organismal size⁸, complexity and diversity through an overall increase in variance (that is, there is “nowhere to go but up”)¹⁰. The second is a non-random, active or ‘driven’ process that biases evolution towards increased size or complexity¹¹.

There are relationships between size and complexity and between complexity and diversity that are intuitive. Increases in organismal size through increases in cell number create the potential for increases in diversity of cell type and, as a result,

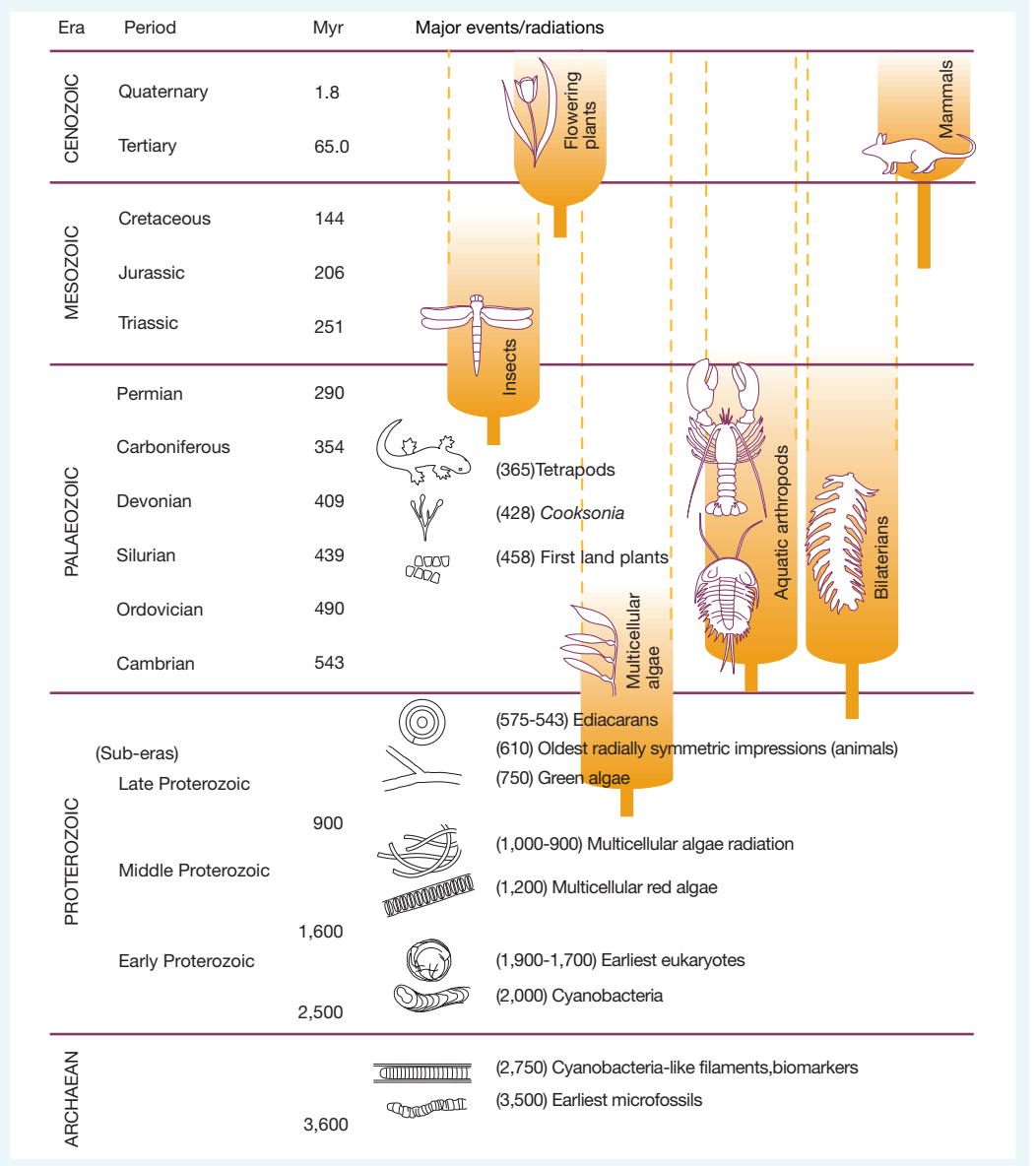
anatomical complexity. Increases in morphological complexity may lead to expansions into previously unoccupied ‘ecospace’ and accompanying expansions of species diversity.

Testing these relationships, searching for trends, and identifying potential causes requires analyses far beyond the mere description of the global history from bacteria to whales, to the consideration of detailed histories of specific lineages. Singularities are dangerous territory for the formulation of general trends and principles¹²; fortunately, however, aspects of evolutionary history have been repeated in different lineages. Here, I will focus on events that have occurred and trends that are manifest in a variety of macroforms with the aim of identifying some of the trends and potential underlying forces that have shaped the size, complexity and diversity of macroscopic life. I will first review the evidence for global trends and then consider the histories of particular lineages where the mechanisms underlying the generation of morphological complexity and the evolution of diversity are beginning to be better understood. In particular, I will focus on the relationship between genomic and organismal complexity that can now be addressed by recent advances in the analysis of genes, genomes and development. I will develop the argument that one of the most important features that has facilitated the evolution of plant and animal complexity and diversity is the modularity of their construction from reiterated, differentiated parts. Finally, I will discuss which trends in the evolution of morphology are likely to apply wherever life may be found.

Milestones in morphological evolution

The principal events of interest here are the major changes in organismal size, form and complexity, and the major expansions in diversity, that have produced the many shapes of macroscopic life. The foundations for inferences about the sequence and direction of evolution are the fossil record and the phylogenetic tree of life. Integration of palaeontological and systematic data is required to establish the number of times particular events occurred, the order in which important sets of traits evolved, and to identify the possible sister groups of major taxonomic groups. The fossil record is also a primary source of data on the time of origin of taxa. One must

Figure 1 History of major evolutionary events from the fossil record. The earliest records of particular groups and the radiations of selected taxa are shown. There is earlier biomarker evidence for eukaryotes around 2,700 Myr (ref. 91). For sources, see text.



bear in mind that initial appearances in the fossil record set only a minimum for the age of clades. Many of the most challenging and controversial questions in evolutionary history concern the origin of major clades (for example, multicellular eukaryotes, animals, land plants, insects and flowering plants), which are also a focus of this review.

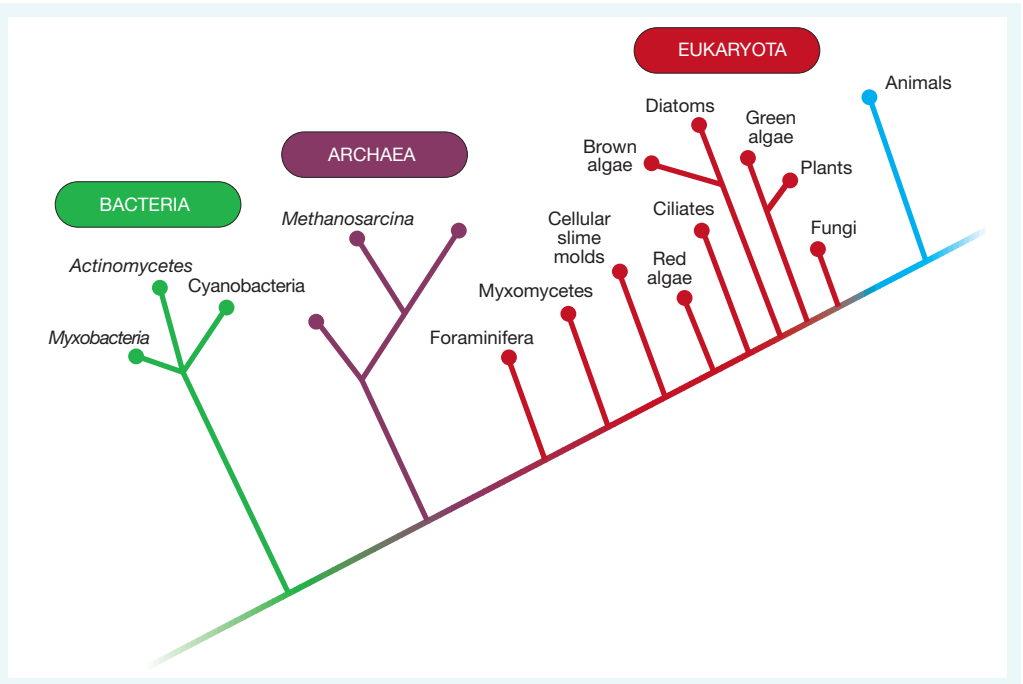
The initial appearances and major radiations of selected taxa documented in the fossil record are summarized in Fig. 1. When considered in the light of the phylogenetic relationships of the major multicellular taxa (Fig. 2), there are three trends evident in the fossil record that I will examine in greater detail. First, multicellularity evolved independently many times and in all three domains of life^{2,13}. Second, following the evolution of multicellularity from different unicellular ancestors, macroscopic forms with new body plans or physiologies and representing higher grades of morphological complexity (for example, multicellular protists, animals and land plants) arose. And third, the emergence of new forms was often followed (after sometimes considerable delays) by periods of rapid diversification (for example, the Cambrian explosion of animals, the rise of insects in the Devonian and Carboniferous, the radiation of flowering plants in the late Cretaceous, and the mammalian radiation in the early Tertiary; Fig. 1). I shall analyse these global trends in the evolution of size, complexity and diversity in more detail and then consider some potential explanations.

Size and multicellularity

For the first 2,500 million years of life on Earth, most species rarely exceeded 1 mm in size and were generally much smaller. The earliest reported bacterial microfossils from about 3,500 million years (Myr) averaged about 5 µm in diameter¹⁴. Early eukaryotic microfossils (acritarchs), while considerably larger, still ranged generally from about 40 to 200 µm in size (with a few larger exceptions, see ref. 15) for much of their first 600–800-Myr history¹⁵. Organismal size increased appreciably with the evolution of multicellular forms. In bacterial and algal forms with cell walls, one of the simplest ways to become multicellular was for the products of cell division to remain together to form long filaments¹³. Many early multicellular eukaryotes were millimetre-scale, linear or branched, filamentous forms^{15,16}.

The size and shape of life did not expand appreciably until the late Proterozoic (Fig. 1). Radially symmetric impressions and trace fossils indicate the presence of millimetre-scale metazoans around 550 Myr (ref. 17; reviewed in ref. 18). The enigmatic Ediacaran fauna comprised of tubular, frond-like, radially symmetric forms generally reached several centimetres in size (although some, such as *Dickinsonia*, approached 1 m), as did macroscopic algae. Organismal sizes expanded considerably in the Cambrian, including bilaterians up to 50 cm in size, as well as sponges and algae up to 5–10 cm (ref. 19). Maximal body lengths of animals increased subsequently

Figure 2 The phylogenetic relationships of multicellular taxa. Multicellularity has evolved independently many times in each domain of life. (Modified from ref. 13.)



by another two orders of magnitude, as did algal sizes (for example, kelp).

The largest extant organisms, giant fungi and trees, evolved from independent small ancestors. Land plants are believed to have evolved from charophyte green algae, and both green algae and plants evolved from a unicellular flagellate ancestor^{20–22}. Fossil spores indicating the earliest evidence of plant life date from the mid-Ordovician. The oldest plant-body fossil (*Cooksonia*) suggests that early land plants were small^{20,21} and, on the basis of molecular phylogenetic analyses, are believed to be comparable in organization and life cycle to liverworts²³. Many of the principal groups of land plants have evolved large (>10 m) species at some point in their history.

Thus, increases in both mean and maximal organismal size occurred in the evolution of multicellular bacteria, eukaryotes and multicellular eukaryotes, and within the algal, fungal, plant and animal lineages.

Complexity

Complexity is one of those problematic terms that has been used to describe so many objects and phenomena as to have lost any generally recognized precision or meaning. In describing organisms, two of the most common usages are in reference to the number of different cell types^{2,13,24} or the number or functional specialization of parts. McShea³ has suggested four distinct categories of complexity that include the specific case of cell-type number but can be used to describe compositions and processes at different levels of biological organization, from molecular to ecological organization. Specifically, these four types of complexity include: (1) the number of different physical parts (for example, genes, cells, organs or organisms) in a system; (2) the number of different interactions among these parts; (3) the number of levels in a causal specification hierarchy; and (4) the number of parts or interactions at a given spatial or temporal scale. These definitions are particularly appropriate in two aspects. First, they allow one to frame questions about the evolution of complexity more narrowly and specifically. Second, they enable the complexity of two independent variables to be compared in order to search for or to refute correlations. For example, it is becoming possible to make comparisons of morphological complexity with the complexity of the genetic systems and developmental programs that generate it.

In each sense of these criteria, the evolution of life has ascended increased grades of complexity. This is most obvious from simple measures of cell number and type (Table 1)^{2,24}. From unicellular

ancestors, multicellular forms have evolved many times in different lineages (Fig. 2). Thus, there have been both global (for example, bacteria to vertebrates) and within-lineage (for example, animals and the green algae/plant clade) increases in the number of cell types. But the maximum number of cell types in general plateaus in bacteria (at 3), in protists (about 4), in protostomes (about 50), and perhaps in vertebrates as well (there is strong suspicion that cell-type number is underestimated in vertebrates but also better studied in these animals).

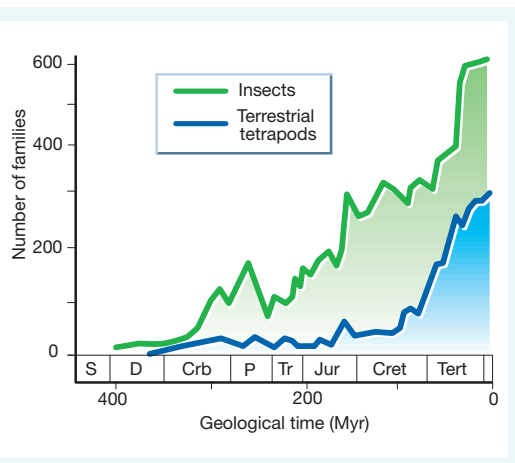
Increases in cell-type number obviously increase the potential physiological and anatomical complexity of organisms through a division of labour among cells and the formation of specialized tissues and organs. Among the bacteria, for example, large cyanobacteria (blue-green algae) have evolved specialized heterocysts that differentiate at regular intervals within a filamentous chain. These heterocysts enable cyanobacteria to segregate the processes of nitrogen fixation (in heterocysts) from photosynthesis (in surrounding vegetative cells)². Spore-forming bacteria are also prokaryotic forms that generate more than one cell type. There are clear advantages to the evolution of hardy spores whose production is under the influence of environmental cues. The shapes of forms with few cell types, from bacteria to slime moulds, are limited generally to filamentous or spherical morphologies. There is little internal morphology in organisms before the evolution of plants, fungi and animals. The evolution of internal complexity accompanied the evolution of greater numbers of cell types and the organization of cells into tissues and organs.

The evolution of cell-type number and internal complexity has been addressed most frequently in the Metazoa, within which the different grades of cellular and anatomical complexity are readily apparent. Placozoans bear only four cell types while the Porifera (sponges) and Cnidaria (including jellyfish and sea anemones) possess 10–12 (ref. 24). Furthermore, cnidarians have only two distinct germ layers (that is, they are ‘diploblastic’), whereas bilaterians possess a third, mesodermal germ layer and considerably more cell types. The evolution of the mesoderm and its derivatives had profound consequences for the evolution of animal body cavities, locomotion and overall size. Among the bilaterians, vertebrates possess the greatest number of cell types, some of which can be attributed to the evolution of the neural crest²⁵.

Another potential index of complexity is gene number. The number of biochemical events within a cell can increase with and therefore bear some correlation to gene number. However, the relationship between gene number and the number of different cell types in multicellular

Figure 3

Patterns of diversification. Terrestrial tetrapods (blue line) and insects (green line) underwent exponential phases of diversification over periods of 100 Myr or more. (Adapted from ref. 30.)



organisms, if any, is not well understood. Recent inventories of the complete genomes of model organisms used for studies of gene regulation, cell differentiation and development offer the opportunity to search for correlations between gene number and cell-type number as indices of complexity.

Among unicellular organisms, total gene numbers range from as few as 470 genes in *Mycoplasma genitalium* to more than 4,000 in *Escherichia coli* (Table 1). There is no apparent relationship between gene number and phylogeny among unicellular organisms. Two members of the Archaea have gene numbers in the middle range of unicellular genomes, as do various bacteria. The smallest genomes are thought to derive from larger genomes through gene loss.

Gene number is not greater in the spore-forming bacterium *Bacillus subtilis* than in other bacteria lacking cell differentiation. However, gene number is considerably greater in the yeast *Saccharomyces cerevisiae* (about 50% greater than the bacterial maximum), which is capable of forming three different cell types (two haploid forms and one diploid form). In the plant *Arabidopsis thaliana* and in two protostomes, the fruitfly *Drosophila melanogaster* and the nematode *Caenorhabditis elegans*, gene number is 2–4 times greater than in *S. cerevisiae*. And in vertebrates, gene number is estimated to be on the order of 4–6 times greater than in the two reported protostome genomes.

From these surveys, we can readily conclude that gene number and cell-type number have increased in the evolution of macroscopic forms from unicellular ancestors. However, the quantitative relationship between the two indices is not at all clear. Gene number varies more than eightfold among unicellular organisms. Although multicellular eukaryotes uniformly possess more genes than bacteria or archaea, organisms with fewer cell types may possess more genes (for example, *A. thaliana* compared with *D. melanogaster*) and the reasons for differences in gene number between species of similar cell-type complexity (for example, *C. elegans* compared with *D. melanogaster*) are unknown. *C. elegans* and *D. melanogaster* belong to the same major protostome clade, the Ecdysozoa, and it is known from analysis of principal developmental genes that the nematode has lost genes that were present in its common ancestors²⁶. In spite of this, the total gene number in *C. elegans* exceeds that of *D. melanogaster* by 5,000 genes. One important contribution to the differences in gene content is the extent and pattern of gene duplications and losses. These events are lineage-specific so the sampling of a few species chosen for reasons other than their phylogenetic relevance may obscure a relationship, if any, between gene number and morphological complexity. However, while total gene number may not be all that informative in regard to complexity, the number of genes with particular developmental functions may be relevant.

Diversity

It is a given that life's diversity has expanded from its origin. The more pertinent question is whether this expansion represents a continuous increase? The answer is most definitely no. Major extinctions have

caused marked reductions in the diversity of the global biota in many episodes of life's history. Furthermore, the dynamics of species diversification and extinction are well studied only for a few groups that have left a long and rich fossil record. For many organisms, particularly those made entirely of soft tissues or of small size, we just cannot say whether total diversity increased or decreased over long periods of time.

What we can say is that there are many episodes in the fossil record of 'bursts' of diversification within lineages. Eukaryotes in the Proterozoic and early Cambrian²⁷, animals in the Cambrian^{28,29}, insects in the Carboniferous, flowering plants in the Tertiary, and other groups experienced periods of rapid radiation (Fig. 1). The geological and ecological settings and the potential catalysts of these periods of accelerated change differ immensely in their particulars.

Perhaps one general theme is that many of these radiations reflect the release from or the surmounting of some environmental or structural constraint(s), or a new way of life. The most obvious of these changes is the transition from an aquatic to a gaseous environment. The invasions of the land by plants and animals were accomplished by enormous changes in physiology and anatomy, which enabled exploitation of new 'ecospace'. The subsequent radiations of land plants, terrestrial tetrapods and insects were explosive and their diversifications followed largely exponential patterns for 100 Myr or longer (Fig. 3)³⁰.

Passive or active global trends?

There are two central questions when evaluating long-term trends in evolution. The first is whether the trend is passive (that is, due to the increase in the total variance in a clade, with the direction of change imposed by the boundary of some initial minimum value), or active (that is, due to the biased replacement of primitive forms with more derived forms). The plot of change in morphology over time distinguishes these two trends (Fig. 4a, b). Given the initial conditions of life (that is, a low minimum boundary), and that simple, small unicellular forms have not been replaced, the global trends in organismal size, complexity and diversity described above must at least in part be passive and due to an overall increase in variance (Fig. 4a). However, the evolution of new traits (such as multicellularity, cell differentiation, internal complexity, support structures and modularity) can establish new levels of complexity and enable subsequent bursts of diversification (through further increases in variance). Thus, whereas global trends may be passive, there may be active, directional trends nested within the overall arc of evolutionary history⁴. To identify these active trends and to consider how they have influenced the shapes of life, we have to look at the histories of individual clades.

The second question in evaluating long-term evolutionary trends, active or passive, concerns whether the mechanism involved is external (affected by selection, ecology or environment) or internal (under genetic, developmental or biomechanical control). It is important to note that the distinction between active and passive trends bears on the question of the potential mechanisms responsible, not on the existence of a trend. Passive trends may well have interesting origins¹⁰. For example, the tendency for mammals to evolve at small sizes ('Cope's rule', see below) begs explanation.

Active trends within clades

Most empirical work on evolutionary trends within clades has focused on animals because of their richer fossil record. Considerable emphasis has been placed upon methodology because different approaches may yield different conclusions regarding the same phenomenon. Three tests have been devised to distinguish passive from active increases: (1) the test of the behaviour of the minimum, which should increase if the system is driven; (2) the ancestor-descendent pair test, which should reveal increases in random samples of ancestor-descendent pairs that are selected away from the boundary of the minimum; and (3) the subclade test, which should reveal a skew in the mean of subclades sampled from the tail of a distribution⁷. These tests have revealed some active trends in the evolution of animal size and complexity.

The evolution of size and Cope's rule

One of the most scrutinized trends is that noted by E. D. Cope³¹ regarding the size increase in mammalian fauna during the Cenozoic (65 Myr–present). Cope attributed the pattern to a tendency for new lineages to evolve at small sizes and an active drive towards increasing size. An active drive has been attributed to various advantages of larger size (for example, evasion of predators, increased brain size or increased longevity; see ref. 8), but rigorous evidence for such a trend was lacking. Stanley⁸ attributed Cope's rule to be due to the tendency for new groups to evolve at small size (that is, an initial minimum value) relative to their optima such that there is a passive drift towards larger mean body size in descendants through an increase in variance. However, Alroy³², in an analysis of body mass estimates for over 1,500 North American fossil mammal species, found a consistent increase between matched pairs of younger and older species. Similarly, MacFadden³³ found a driven trend in the evolution of body size of horses within this era. These within-lineage comparisons indicate that active trends do operate within overall passive trends.

A general applicability of Cope's rule to other taxa and eras has not been found. Neither Cretaceous molluscs³⁴ nor planktonic foraminifera¹⁰ show an active trend. Rather, trends of increasing size in both of these taxa have been attributed to an overall increase in variance.

Complexity of animal morphology

One of the most often analysed trends is the evolution of morphological complexity in the Metazoa. As discussed earlier, cell-type number has increased in the evolution of diploblasts, bilaterians and vertebrates, respectively. As a long-term trend, however, it also seems that cell-type number plateaus in these groups, so there may have been increases in the stem lineages of these groups, but not in the subsequent radiations of these lineages. The discovery of many genetic and developmental similarities in different bilaterian clades has led to the inference of a common ancestor of these clades that was much more anatomically complex than once thought^{18,35,36}. Although likely to have been small in comparison to its Cambrian and later descendants, cell-type number in the last common ancestor of bilaterians was probably comparable to that of modern protostomes and basal deuterostomes³⁷. However, minimum cell numbers probably increased with body size in certain clades in the Cambrian, indicating perhaps a driven trend in the Cambrian radiation.

Complexity with respect to other characters has been scrutinized only for taxa with rich fossil records such as the brachiopods^{3,38}, ammonoids³⁹, aquatic arthropods⁴⁰ and vertebrates⁴¹. Active trends have been identified in the complexity of brachiopod geometry³⁸, ammonoid septal sutures³⁹ and arthropod limb types⁴⁰.

One particular type of complexity of special interest is that of serially repeated structures. Body segments in annelids and arthropods, vertebrate in vertebrates, limbs in many taxa, and teeth are serially homologous structures. Compared to other structures, these are easily quantified and differentiated. Complexity of serially repeated parts is a function then of both the overall total and the number of individual types of structures. In both vertebrates and arthropods there has been a clear increase in the maximum number of distinguishable individual types of repeated structures³. In the evolution of diverse arthropods from trilobitomorphic or lobe-podan⁴² ancestors, the mean and maximum number of distinct limb-pair types increased⁴⁰ (Fig. 5), as did the minimum (temporarily). This suggests a trend that may be driven in part. All three measures have remained static for the past 250 Myr of arthropod evolution (Fig. 5).

In vertebrates as well, the maximum and mean number of differentiated vertebrae has increased at a high taxonomic level in the transition from fish to mammals. Vertebral columns of fossil and modern fish are relatively uniform whereas those of birds and mammals are more complex. Detailed analyses indicate that this is a largely passive trend as the minima have not changed and there is no clear trend of increase in ancestor-descendant pairs⁴¹.

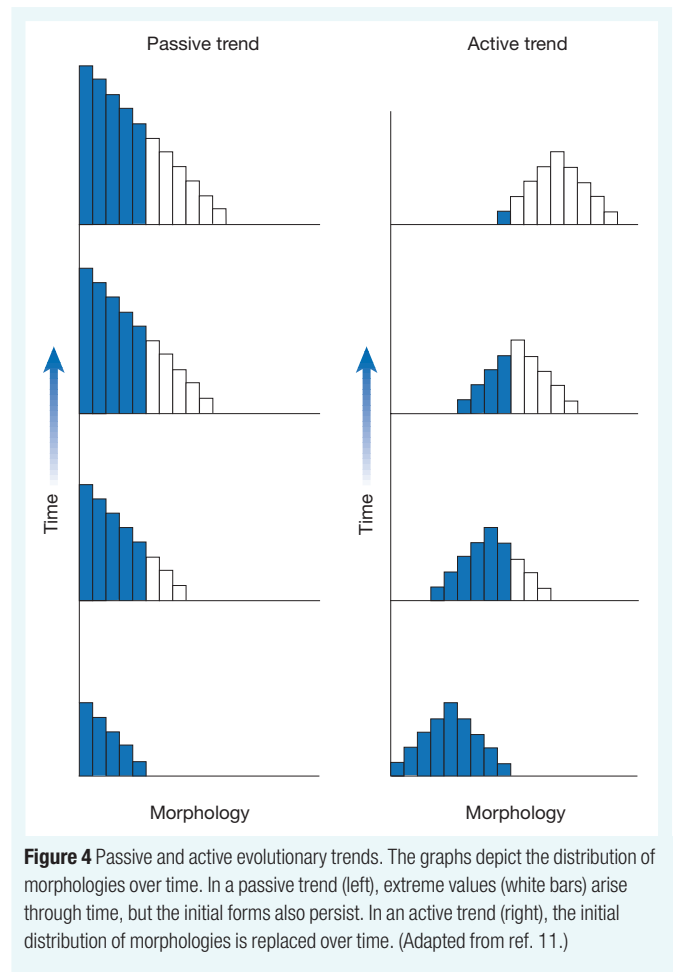


Figure 4 Passive and active evolutionary trends. The graphs depict the distribution of morphologies over time. In a passive trend (left), extreme values (white bars) arise through time, but the initial forms also persist. In an active trend (right), the initial distribution of morphologies is replaced over time. (Adapted from ref. 11.)

Diversity

There are two different elements of diversity that are distinguishable, the morphological and the taxonomic. Morphological diversity is a function of the occupancy of multidimensional morphospace, whereas taxonomic diversity is a function of net speciation events. In principle, these two functions can vary independently of each other. There is empirical evidence that morphological variety and taxonomic diversity can increase together during the initial diversification of a clade⁴³. This trend would be consistent with a passive, diffusive process such that as variance increases, morphospace and ecospace are filled from initial boundary minima.

That the filling of morphospace is initially more rapid is well illustrated by analysis of the evolution of skeletal designs. The 'skeleton space' is a theoretical morphospace against which the actual skeletal designs of fossil and extant organisms have been compared⁴⁴. Of the roughly 180 designs that have been used by all phyla that bear skeletal elements (internal or external, rigid or pliable, uni- or multicomponent, and of various geometries), 146 were exploited by the time of the early Middle Cambrian (in Burgess Shale fauna)⁴⁵. Thus, more than 80% of all designs that ever evolved appeared within the first 6% of overt animal history. Among the earliest and most frequently exploited designs were single-element rods and multielement, metameric exoskeletons. These structures are correlated with the simplest possible strategies for increasing body size. Underrepresented in Cambrian fauna are structures whose growth involves continuous remodelling, such as those of animals with internal skeletons.

Modularity provides the key link in macroform evolution

There is an intuitive relationship between complexity and diversity. As new traits evolve that enable organisms to surmount prior

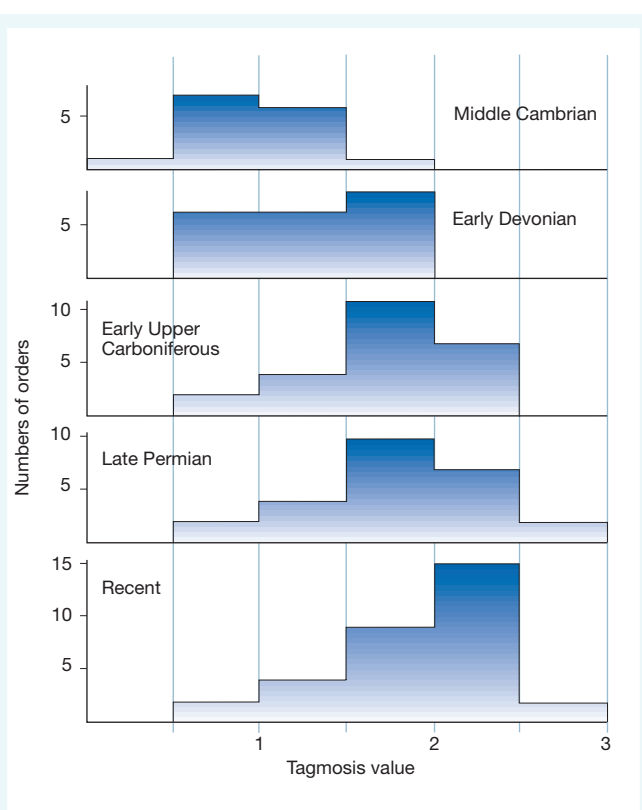


Figure 5 The evolution of limb-type complexity in aquatic arthropods. The number of orders is plotted against tagmosis, a value computed by considering both the total number and different types of arthropod limbs. The minimum, mean, and maximum tagmosis value increased from the Cambrian to the late Permian, suggesting an active trend. (Adapted from ref. 37.)

limitations, such increases in complexity may allow the exploitation of new morphospace and ecospace^{4,46}. In the case of skeletons described above, the evolution of the biochemistry for forming hard parts (chitinous cuticles and mineralized internal or external elements) opened up new ways of life that enabled bursts of diversification. This theme appears repeatedly in evolutionary history. The invasion of the land by plants, tetrapods and insects triggered radiations that transformed terrestrial ecosystems. The early phase of diversification of these groups (considering higher taxonomic levels) was greater than that of later phases, until later innovations (for example, flight in vertebrates or flowers in plants) led to further bursts of diversification. Once again, the overall trends seem to be passive (and those at higher taxonomic levels are to some degree the consequence of how the topology of evolutionary trees are determined^{47,48}), but that does not at all diminish the importance of mechanisms underlying the increases in complexity or diversity in these lineages, the most complex and diverse of all macroscopic forms.

Can we draw any generalities about deeper mechanisms underlying these overall trends, perhaps even the active trends that may be nested within them? I argue here that one of the most critical features underlying the evolution of large and complex animals and plants, and a key to their diversity, is their modular construction. The significance of the construction of animals from repeated parts has long been recognized. Cope³¹, Bateson⁴⁹, Gregory⁵⁰, Rensch⁵¹, Bonner², as well as Darwin (see ref. 52) have suggested various advantages of modular construction, including the facilitation of greater size and efficiency, and the evolution of greater complexity and adaptation through the functional differentiation of repeated parts. Similarly, in plants, modular construction allows for greater size and the differentiation of functional roles among leaf and reproductive structures⁵³. Modularity in plants

and animals can be viewed as being analogous conceptually to the division of labour between cell types of simpler organisms such as cyanobacteria and *B. subtilis*, except that whole body parts are the building blocks rather than cells. The main innovation that enabled large, modular organisms to evolve was the evolution of regional specification systems that subdivide growing embryos into semiautonomous units⁵⁴. We now understand enough about some of the developmental genetic mechanisms for the construction of modular animals (for example, arthropods and vertebrates⁵⁵) and of plant body plans and parts^{56,57} that statements can be made about the mechanistic bases of the morphological diversification of forms at higher taxonomic levels.

For an illustration of the relationship between the evolution of modular body plans, complexity, diversity and the architecture of the underlying genetic systems that differentiate forms, we can return to the example of the evolution of arthropod limb-pair types (Fig. 5). Cisne⁴⁰ documented a trend of an increasing mean and maximum number of limb-pair types in the first half of the Phanerozoic. The functional significance of this potentially driven trend is obvious. Arthropod limbs are “the tools of its trade”⁴⁰ and the morphological specialization of limbs for feeding, locomotion, sensation, copulation, brooding young, burrowing and defence is reflected in the variety of limb types (that is, complexity). The most specialized orders, those with the greatest number of different limb types, are also the most diverse in terms of the number of species. This is illustrated by the malacostracan crustacean lineage that evolved filtering maxilla as a new means of feeding. The evolution of maxillopods freed the serially similar pairs of trunk limbs to become specialized for walking, swimming and burrowing⁴⁰.

The diversification of the modular arthropod body plan and of arthropod limb types required the diversification of the genetic regulatory system that specifies regional and segmental identities in arthropods, namely, the *Hox* genes. The products of the *Hox* genes are expressed in domains that subdivide the anterior–posterior axis of bilaterians, and these proteins regulate the expression of many genes within these domains⁵⁸. The differentiation of serially homologous limbs in arthropods is regulated by different *Hox* genes expressed in different developing limb primordia (reviewed in refs 16, 55 and 59). Comparisons of representatives of all arthropod classes and many different orders have revealed strong correlations between the regional deployment of *Hox* genes and the patterns of limb differentiation (reviewed in refs. 16, 55 and 60). Thus, in taxa with similar sets of repeating limbs (for example, centipedes and brachiopod crustaceans), the same *Hox* genes or combination of *Hox* genes are expressed in all of the limbs of the same type. Whereas in taxa with different numbers and kinds of limbs, relative shifts in *Hox*-gene expression domains are correlated with differences in limb number and identity.

Arthropod limb-type diversity has evolved from ancestors with very similar complements of *Hox* genes⁶¹. This is contrary to initial expectations that gene duplication and divergence would correlate with the duplication and increasing diversity of arthropod segment and limb types^{62,63}. Instead, it is apparent that the great diversification of limb types is due to the evolution of regulatory mechanisms operating at two major hierarchical levels in arthropod development. First, at the level of the regulation of *Hox* genes along the anterior–posterior axis; and second, within the *Hox*-regulated hierarchies of genes that pattern individual limbs. The evolution of axial diversity within the modular body plans of other bilaterian phyla such as annelids and vertebrates has followed similar themes (reviewed in ref. 55).

Modularity, constraints and evolvability

These recent discoveries in developmental genetics and comparative biology illustrate an important property bestowed upon evolving organisms by modularity, that is, the ability to dissociate developmental processes in one part of the body from another^{64,65}. The regional specification systems that subdivide developing animal and plant embryos into discrete territories make it possible for the development and morphology of one territory to evolve independently of

Table 1 Evolution of cell type and gene number

Number of cell-types*	Species	Number of genes in genome	Reference
1	<i>Mycoplasma genitalium</i>	470	76
	<i>Rickettsia prowazekii</i> (intracellular parasite)	834	77
	<i>Haemophilus influenzae</i>	1,709	78
	<i>Escherichia coli</i>	4,288	79
	<i>Campylobacter jejuni</i>	1,654	80
	<i>Aquifex aeolicus</i> (thermophile)	1,512	81
	<i>Neisseria meningitidis</i>	2,121	82
	<i>Archaeoglobus fulgidus</i> (Archaea)	2,436	83
	<i>Methanococcus jannaschii</i> (Archaea)	1,738	84
	<i>Synechocystis</i> sp. (cyanobacterium)	3,168	85
2	<i>Bacillus subtilis</i>	~4,100	86
	<i>Caulobacter crescentus</i>		
3	<i>Saccharomyces cerevisiae</i>	6,241	87
4	<i>Volvox</i>		
	<i>Ulva</i> (sea lettuce) placozoans		
7	Mushrooms		
	Kelp		
~11	Sponge, cnidarians		
~30	<i>Arabidopsis thaliana</i> (plant)	~24,000	88,89
~50	<i>Caenorhabditis elegans</i> (nematode)	18,424	90
	<i>Drosophila melanogaster</i> (fruitfly)	13,601	70
~120	Zebrafish	>80,000–100,000	
~120	Human	~80,000–100,000	

*From refs 2, 24.

another. It has been suggested that modularity (also called compartmentation^{66,67}) facilitates change by conferring upon organisms a greater ability to escape internal constraints on morphology⁶⁷. These constraints include the physical limits imposed by biomechanics on organismal size and shape and genetic and developmental constraints that limit the range of variation that is tolerated and available within species (see ref. 64 for a discussion of constraints). If modularity and dissociation enhance the capacity to generate variation (that is, evolvability), then this may confer a selective advantage on modular clades that possess it^{66,67}. Exploitation of this dissociability is illustrated by the great variety of forms and functions seen in the serially repeated parts of animals and plants.

Genomic complexity and regulatory evolution

The generation of cell-type and body-part diversity depends upon transcriptional regulatory proteins that control the cell- or region-specific expression of target genes. The differences in gene expression between cell types within an organism can number in the hundreds to thousands, but these differences are often controlled by just a small set of regulatory proteins. In *B. subtilis* for example, a small number of regulators control the differential expression of several hundred genes during sporulation⁶⁸. In the yeast *S. cerevisiae*, a small set of transcription factors orchestrates the regulation of genes involved in cell-type differences⁶⁹. In metazoans, cell-type differences (for example, muscle compared with neural) and body-region identity are regulated typically by a few proteins, whereas pattern formation within tissues is regulated by a larger set of proteins⁵⁵. Only about 3–5% of the proteins encoded by animals' genomes are transcriptional regulators^{70,71}. Therefore, the total number of genes is not a driver of cell type or other indices of morphological complexity. For instance, greater gene number (for example, in *C. elegans* compared to *D. melanogaster*, or in zebrafish compared to humans) does not dictate greater cell-type number or any other index of morphological complexity (indeed, *C. elegans* is a highly derived ecdysozoan that has lost certain morphological features (eyes) and *Hox* genes that were present in the common ancestor it shares with *Drosophila*).

It seems that most expansion in the genetic toolkit for bilaterian development occurred in two intervals of bilaterian evolution. First, in the stem lineage leading to bilaterians, and later, in the evolution of vertebrates from a chordate ancestor⁷². These two intervals of *Hox* gene expansion do correlate with the evolution of increased grades of

complexity. Interestingly, protostomes and more basal deuterostomes possess similar complements of regulators that control cell-type, tissue and regional identities.

Expansions in the number of regulatory proteins offer the potential, but are not necessary for, the evolution of increased complexity and the expansion of diversity. The evolutionary trends in arthropod and vertebrate axial complexity and diversity are due to the evolution of genomic complexity at a different level than gene number, that is, at the level of the evolution of the regulatory elements that act in *cis* to control gene expression. Within these phyla, no substantial increase in the number of genes involved in regional specification occurred. Rather, the evolution of advanced forms with greater numbers of differentiated serial structures has occurred through an expansion in the number of regulatory elements that control region-specific expression of genes. The expansion of regulatory elements constitutes increases in genomic complexity in all four senses described earlier — in the number of different parts (regulatory elements) in a regulatory system, in the number of different interactions of these parts, in the number of levels in developmental hierarchies, and in the number of parts and interactions at a given spatial scale. Regulatory evolution creates new combinations of gene expression and therefore enables increases in the information content of genomes and the generative potential of development without expansion of gene number. The role of regulatory evolution is therefore key to understanding how morphological complexity and diversity evolve in macroscopic forms.

Chance and necessity

Perhaps the most surprising conclusion one might draw from the consideration of the complexity of genes and genomes is that the generative potential of genomes is far greater than is realized in evolution. Kauffman⁷³ has pointed out that there is a vast difference between the potential number of combinations of possible gene expression states and those that actually exist in any organism. Given just two inputs into each gene, a system of 100,000 genes has 2^{100,000} different possible states. Yet, if we use cell-type number as an indicator of gene expression states, only 200–300 states are realized (more states are realized in modular organisms through differential expression in different modules). Furthermore, given that multicellularity seems to be readily evolved and that very few regulatory proteins can orchestrate markedly different cell physiologies, it is curious that more multicellular forms have not evolved.

We do not understand why the actual complexity realized in evolution is far less than what seems to be possible genetically. The observed limits of form seem to be due to a combination of both chance and necessity, a product of historical contingency and imposed by external agents (for example, selection) and internal rules (for example, constraints). The demands of natural selection may exclude or favour certain forms, but it is widely agreed that selection cannot be the whole story. Internally imposed constraints also shape the world of possible morphologies and are themselves factors that can evolve⁷⁴.

Are there universal rules to the shapes of life?

The parochial question nested within the mystery of the existence of life on other bodies is that of the existence of forms like the ones that have occurred on Earth. A few extrapolations seem to be reasonably grounded in the overall trends of life's history reviewed here. Assuming a cellular basis of life elsewhere, the passive trends towards increases in organismal size, complexity and diversity from some initial minima are certain to prevail in any system. It must be kept in mind, however, that few macroscopic forms evolved in the first 3 billion years of life on Earth. Therefore, the time required for any quantum change in morphology is entirely contingent upon the particular history of any system. As for the shapes of life, macroscopic forms are most likely to be multicellular and there is a finite set of simple geometries — such as those that dominated the early history of life on Earth (linear and branched filaments, cylinders and

spheres) — that are likely to satisfy the constraints imposed by diffusion and biomechanics and that are therefore likely to be universal^{2,75}.

But the evolution of motile, modular mega-organisms may be a different story. Only after 3 billion years of physiological and anatomical evolution, vast changes in the environment and ecology (that were partly biogenic in nature), and extensive genetic and developmental innovations did such beasts emerge on Earth. And, although some symmetrical body organization is likely of macro-forms⁷⁵, there is no basis to assert that bilateral, radial or spiral forms were or would be inevitable. Nor, sadly, is their continued evolution assured as the ecological dice are now in the hands of a single species that is on a path to extinguishing a substantial fraction of all diversity before the question of life elsewhere may be answered. □

1. Monod, J. *Chance and Necessity* (Vintage Books, New York, 1971).
2. Bonner, J. T. *The Evolution of Complexity by Means of Natural Selection* (Princeton Univ. Press, Princeton, 1988).
3. McShea, D. W. Metazoan complexity and evolution: is there a trend? *Evolution* **50**, 477–492 (1996).
4. Knoll, A. H. & Bambach, R. K. Directionality in the history of life: diffusion from the left wall or repeated scaling of the right? *Paleobiology* **26** (Suppl.) 1–14 (2000).
5. Gould, S. J. *Full House* (Harmony Books, New York, 1996).
6. Foote, M. Cambrian and recent morphological disparity. *Science* **258**, 1816–1818 (1992).
7. McShea, D. W. Mechanisms of large-scale evolutionary trends. *Evolution* **48**, 1747–1763 (1994).
8. Stanley, S. M. An explanation for Cope's rule. *Evolution* **27**, 1–26 (1973).
9. Maynard Smith, J. Time in evolutionary process. *Studia Generale* **23**, 266–272 (1970).
10. Gould, S. J. Presidential Address. Trends as changes in variance: a new slant on progress and directionality in evolution. *J. Paleont.* **62**, 319–329 (1988).
11. Wagner, P. J. Contrasting the underlying patterns of active trends in morphological evolution. *Evolution* **50**, 990–1007 (1996).
12. Hawking, S. W. *A Brief History of Time: From the Big Bang to Black Holes* (Bantam Books, Toronto, 1988).
13. Bonner, J. T. The origins of multicellularity. *Integ. Biol.* **1**, 28–36 (1998).
14. Schopf, J. W. Microfossils of the Early Archean Apex chert: new evidence of the antiquity of life. *Science* **260**, 640–645 (1993).
15. Knoll, A. H. The early evolution of eukaryotes: a geological perspective. *Science* **256**, 622–627 (1992).
16. Butterfield, N. J. *Bangiomorpha pubescens* n. gen., n. sp.: implications for the evolution of sex, multicellularity, and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes. *Paleobiology* **26**, 386–404 (2000).
17. Martin, M. W. et al. Age of Neoproterozoic bilaterian body of trace fossils, White Sea, Russia: implications for metazoan evolution. *Science* **288**, 841–845 (2000).
18. Knoll, A. H. & Carroll, S. B. Early animal evolution: emerging views from comparative biology and geology. *Science* **284**, 2129–2137 (1999).
19. Briggs, D. E. G., Erwin, D. H. & Collier, F. J. *The Fossils of the Burgess Shale* (Smithsonian Institution Press, Washington DC, 1994).
20. Kenrick, P. & Crane, P. R. *The Origin and Early Diversification of Land Plants* (Smithsonian Institution Press, Washington DC, 1997).
21. Graham, L. E. *Origin of Land Plants* (Wiley, New York, 1993).
22. Kirk, D. L. *Volvox: Molecular Genetic Origins of Multicellularity and Cellular Differentiation* (Cambridge Univ. Press, Cambridge, 1998).
23. Qui, Y.-L., Cho, Y., Cox, J. C. & Palmer, J. D. The gain of three mitochondrial introns identifies liverworts as the earliest land plants. *Nature* **394**, 671–674 (1998).
24. Valentine, J. W., Collins, A. G. & Meyer, C. P. Morphological complexity increase in metazoans. *Paleobiology* **20**, 131–142 (1994).
25. Shimeld, S. M. & Holland, P. W. H. Vertebrate innovations. *Proc. Natl Acad. Sci. USA* **97**, 4434–4437 (2000).
26. de Rosa, R. et al. *Hox* genes in brachiopods and priapulids and protostome evolution. *Nature* **399**, 772–776 (1999).
27. Knoll, A. H. Proterozoic and early Cambrian protists: evidence for accelerating evolutionary tempo. *Proc. Natl Acad. Sci. USA* **91**, 6743–6750 (1994).
28. Conway Morris, S. *The Crucible of Creation: The Burgess Shale and the Rise of Animals* (Oxford Univ. Press, Oxford, 1998).
29. Budd, G. E. & Jensen, S. A critical reappraisal of the fossil record of the bilaterian phyla. *Biol. Rev.* **75**, 253–295 (2000).
30. Benton, M. J. Models for the diversification of life. *Trends Ecol. Evol.* **12**, 490–495.
31. Cope, E. D. The method of creation of organic forms. *Proc. Am. Phil. Soc.* **12**, 229–263 (1871).
32. Alroy, J. Cope's rule and the dynamics of body mass evolution in North American fossil mammals. *Science* **280**, 731–734 (1998).
33. MacFadden, B. J. Fossil horses from "Eohippus" (Hyracotherium) to Equus: scaling, Cope's law, and the evolution of body size. *Paleobiology* **12**, 355–369 (1986).
34. Jablonski, D. Body-size evolution in Cretaceous molluscs and the status of Cope's rule. *Nature* **385**, 250–252 (1997).
35. De Robertis, E. M. & Sasai, Y. A common plan for dorsoventral patterning in Bilateria. *Nature* **380**, 37–40 (1996).
36. De Robertis, E. M. The ancestry of segmentation. *Nature* **387**, 25–26 (1997).
37. Valentine, J. Late Precambrian bilaterians: grades and clades. *Proc. Natl Acad. Sci. USA* **91**, 6751–6757 (1994).
38. Carlson, S. J. Evolutionary trends in the articulate brachiopod hinge mechanism. *Paleobiology* **18**, 344–366 (1992).
39. Saunders, W. B., Work, D. M. & Nikolaeva, S. V. Evolution of complexity in Paleozoic ammonoid structures. *Science* **286**, 760–763 (1999).
40. Cisne, J. L. Evolution of the world fauna of aquatic free-living arthropods. *Evolution* **28**, 337–366 (1974).
41. McShea, D. W. Evolutionary change in the morphological complexity of the mammalian vertebral column. *Evolution* **47**, 730–740 (1993).
42. Budd, G. The morphology of *Opabinia regalis* and the reconstruction of the arthropod stem-group. *Lethaia* **29**, 1–14 (1996).
43. Foote, M. Discordance and concordance between morphological and taxonomic diversity. *Paleobiology* **19**, 185–204 (1993).
44. Thomas, R. D. K. The skeleton space: a finite set of organic designs. *Evolution* **47**, 341–360 (1993).
45. Thomas, R. D. K., Shearman, R. M. & Stewart, G. W. Evolutionary exploitation of design options by the first animals with hard skeletons. *Science* **288**, 1239–1242 (2000).
46. Stebbins, G. L. *The Basis of Progressive Evolution* (Univ. North Carolina Press, Chapel Hill, 1969).
47. Raup, D. M. On the early origins of major biologic groups. *Paleobiology* **9**, 107–115 (1983).
48. Smith, A. B. *Systematics and the Fossil Record* (Blackwell Scientific Publications, Oxford, 1994).
49. Bateson, W. *Materials for the Study of Variation* (Macmillan, London, 1894).
50. Gregory, W. K. Reduplication in evolution. *Quart. Rev. Biol.* **10**, 272–290 (1935).
51. Rensch, B. *Evolution above the Species Level* (Columbia Univ. Press, New York, 1960).
52. McShea, D. W. Complexity and evolution: what everybody knows. *Biol. Philosophy* **6**, 303–324 (1991).
53. Niklas, K. J. *Plant Allometry: The Scaling of Form and Process* (Univ. Chicago Press, Chicago, 1994).
54. Davidson, E., Peterson, K. & Cameron, R. Origin of bilaterian body plans: evolution of developmental regulatory mechanisms. *Science* **270**, 1319–1325 (1995).
55. Carroll, S. B., Grenier, J. & Weatherbee, S. D. *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design* (Blackwell, Cambridge, 2001).
56. Ma, H. & DePamphilis, C. The ABCs of floral evolution. *Cell* **101**, 5–8 (2000).
57. Ambrose, B. A. et al. Molecular and genetic analyses of the *silky1* gene reveal conservation in floral organ specification between eudicots and monocots. *Mol. Cell* **5**, 569–579 (2000).
58. Carroll, S. Homeotic genes and the evolution of arthropods and chordates. *Nature* **376**, 479–485 (1995).
59. Weatherbee, S. D. & Carroll, S. B. Selector genes and limb identity in arthropods and vertebrates. *Cell* **97**, 283–286 (1999).
60. Akam, M. Arthropods: developmental diversity within a (super) phylum. *Proc. Natl Acad. Sci. USA* **97**, 4438–4441 (2000).
61. Grenier, J., Garber, T., Warren, R., Whittington, P. & Carroll, S. Evolution of the entire arthropod *Hox* gene set predates the origin and radiation of the onychophoran/arthropod clade. *Curr. Biol.* **7**, 547–553 (1997).
62. Lewis, E. B. A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565–570 (1978).
63. Akam, M. The molecular basis for metameric pattern in the *Drosophila* embryo. *Development* **101**, 1–22 (1987).
64. Raff, R. A. *The Shape of Life* (Univ. Chicago Press, Chicago, 1996).
65. Raff, R. & Sly, B. J. Modularity and dissociation in the evolution of gene expression territories in development. *Evol. Dev.* **2**, 102–113 (2000).
66. Gerhart, J. & Kirschner, M. *Cells, Embryos, and Evolution* (Blackwell Science, Malden, 1997).
67. Kirschner, M. & Gerhart, J. Evolvability. *Proc. Natl Acad. Sci. USA* **95**, 8420–8427 (1998).
68. Fawcett, P., Eichenberger, P., Losick, R. & Youngman, P. The transcriptional profile of early to middle sporulation in *Bacillus subtilis*. *Proc. Natl Acad. Sci. USA* **97**, 8063–8068 (2000).
69. Johnson, A. D. Molecular mechanisms of cell-type determination in yeast. *Curr. Opin. Genet. Dev.* **5**, 552–558 (1998).
70. Adams, M. D. et al. The genome sequence of *Drosophila melanogaster*. *Science* **287**, 2185–2195 (2000).
71. Ruvkin, G. & Hobart, O. The taxonomy of developmental control in *C. elegans*. *Science* **282**, 2033–2041 (1998).
72. Garcia-Fernandez, J. & Holland, P. W. Archetypal organization of the amphioxus *Hox* gene cluster. *Nature* **370**, 563–566 (1994).
73. Kauffman, S. *At Home in the Universe* (Oxford University Press, New York, 1995).
74. Budd, G. E. Does evolution in body patterning genes drive morphological change—or vice versa? *BioEssays* **21**, 326–332 (1999).
75. Stewart, I. *Life's Other Secret* (Wiley, New York, 1998).
76. Fraser, C. M. et al. The minimal gene complement of *Mycoplasma genome*. *Science* **270**, 397–404 (1995).
77. Andersson, S. G. E. et al. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature* **396**, 133–143 (1998).
78. Fleischmann, R. D. et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**, 496–412 (1995).
79. Blattner, F. R. et al. The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1462 (1997).
80. Parkhill, J. et al. The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals highly variable sequences. *Nature* **403**, 665–668 (2000).
81. Deckert, G. et al. The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. *Nature* **392**, 353–358 (1998).
82. Parkhill, J. et al. Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491. *Nature* **404**, 502–506 (2000).
83. Klenk, H.-P. et al. The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. *Nature* **390**, 364–370 (1997).
84. Bult, C. J. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* **273**, 1058–1073 (1996).
85. Kaneko, T. et al. Sequence analysis of the genome of the unicellular Cyanobacterium *Synechocystis* sp. strain. *DNA Res.* **3**, 109–136 (1996).
86. Kunst, F. et al. The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. *Nature* **390**, 249–256 (1997).
87. Goffeau, A. et al. The yeast genome directory. *Nature* **387**, 5–105 (1997).
88. European Union Arabidopsis Genome Sequencing Consortium & Cold Spring Harbor, Washington University in St Louis and PE Biosystems Arabidopsis Sequencing Consortium. Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*. *Nature* **402**, 769–777 (1999).
89. Lin, X. et al. Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. *Nature* **402**, 761–768 (1999).
90. Rubin, G. M. et al. Comparative genomics of the eukaryotes. *Science* **287**, 2204–2215 (2000).
91. Brocks, J. J., Logan, G. A., Buick, R. & Summons, R. E. Archean molecular fossils and the early rise of eukaryotes. *Science* **285**, 1033–1036 (1999).

Acknowledgements

I thank A. Knoll for discussions and helpful pointers; R. Losick and A. Johnson for information on microbial gene regulation; G. Budd, J. Crow, N. King and J. True for suggestions on the text; J. Carroll for preparation of the manuscript; and L. Olds for the artwork. S.B.C. is an Investigator of the Howard Hughes Medical Institute.