Explaining the Cambrian "Explosion" of Animals

Charles R. Marshall
Department of Earth and Planetary Sciences, Department of Organismic and Evolutionary Biology, and Department of Invertebrate Paleontology, Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts 02138; email: cmarshall@oeb.harvard.edu

Key Words
bilaterian phyla, evolution, paleontology, development, ecology

Abstract
The Cambrian “explosion” is a unique episode in Earth history, when essentially all the animal phyla first appear in the fossil record. A variety of environmental, developmental (genetic), and ecological explanations for this complex and somewhat protracted event are reviewed, with a focus on how well each explains the observed increases in disparity and diversity, the time of onset of the radiation, its duration, and its uniqueness. The increase in disparity (the origin of the phyla) and diversity are best understood as being the result of the interplay of the combinatorial bilaterian developmental system and the increase in the number of needs the first bilaterians had to meet as complex ecological interactions developed. The time of onset is constrained by the evolution of the environment, whereas its duration appears to be controlled primarily by rates of developmental innovation. The uniqueness of the event is either due to ensuing developmental limitation, to ecological saturation, or simply to the exhaustion of ecologically viable morphologies that could be produced by the nascent bilaterian developmental system.
INTRODUCTION

The Cambrian “explosion,” or radiation, is perhaps the most significant evolutionary transition seen in the fossil record. Essentially all of the readily fossilizable animal body plans first appear in the fossil record during this interval (Valentine 2002). We move from the depths of the Precambrian world, where the sedimentary record is essentially devoid of animal fossils, to the Phanerozoic, where animal life leaves pervasive evidence of its existence, both as body fossils and as disturbers of the sediment.

Numerous explanations for the Cambrian “explosion” have been posited (note here that I am not considering in any detail explanations for the precursor to the Cambrian “explosion,” the Ediacaran radiation). Classification of this rich panoply of explanations is somewhat arbitrary but typically explanations center on one of the following factors: (a) changes in the abiotic environment, (b) changes in the genetic or developmental capacity of the taxa involved, or (c) changes in the biotic environment, i.e., in ecology. All of these factors must have played a role, but how important was each? To what extent did the Cambrian “explosion” flow from an interaction between them? How might we develop a conceptual framework for understanding that interaction?

Developing a coherent explanation for the Cambrian “explosion” faces several challenges. First, most workers interested in the Cambrian “explosion” approach the problem through their primary discipline(s) of activity, whether that be paleontology, geology, geochemistry, ecology, climate modeling, developmental biology, etc. Thus, developing a balanced multidisciplinary explanation is at the outset handicapped; the phenomenology and conceptual apparatus needed to understand the problem are very rich. This difficulty is exacerbated by the fact that different subdisciplines use different approaches for exploring the unknown. For example, explanations that flow from the geological record typically focus on events that occurred just before or at the Precambrian/Cambrian boundary, and thus tend to be at their heart correlative rather than causal in nature; a challenge faced by this type of study is determining how (or whether) proposed triggering events or changes relate causally to the radiation. Ecological or developmental explanations are more obviously evolutionary in nature and so there is typically a plausible causal relation between these explanations and the biological radiation seen in the fossil record. However, these explanations are not usually tied in any significant way to the geologic record, to the history of the planet.

Part of the difficulty in developing an explanation for the Cambrian “explosion” also stems from the fact that it is an unfolding in history; each successive step not only flows from the conditions established in the previous steps but also is shaped by interactions at the current step of unfolding. So in some sense there cannot be a simple explanation for the Cambrian “explosion” that is fully satisfying; understanding the nature of causation of complex historical events is difficult.

Here, I first briefly review our knowledge of the Cambrian “explosion,” outline what needs to be explained, and then summarize some of the more prominent explanations for the event, focusing on how effectively each explains the various components of the “explosion” that need to be accounted for. I then show how recent advances in our understanding of the evolution of fitness landscapes provide a way of understanding the interactions between organisms, their environment (both the
ANATOMY OF THE CAMBRIAN “EXPLOSION”

The Cambrian “explosion” has a complex anatomy, and a great deal has been written about this, including recent and insightful reviews by Budd & Jensen (2000), the various contributions in Zhuravlev & Riding (2001), and Valentine (2002, 2004). The edited volume by Zhuravlev & Riding (2001) provides a particularly fine introduction to the environment, community patterns and dynamics, and radiation of the major groups of organisms through the Cambrian radiation, whereas Valentine (2004) provides a masterly treatment of the evolution of the phyla, with a special emphasis on their origins. Here, I simply outline the key elements of the Cambrian “explosion.” However, before doing so, it is important to remember that we see the Cambrian “explosion” through the windows permitted by the fossil and geological records. So when talking about the Cambrian “explosion,” we are typically referring to the appearance of large-body (can be seen by the naked eye) and preservable (and therefore largely skeletonized) forms. Finally, I place the word “explosion” in quotation marks because, while the Cambrian radiation occurred quickly compared with the time between the Cambrian and the present, it still extended over some 20 million years of the earliest Cambrian, or longer if you add in the last 30 million years of the Ediacaran and the entire 55 million year duration of the Cambrian.

Chronology

Apart from a few older problematic fossils (both body and putative trace), the first relatively continuous fossil record of animal life begins after the Gaskiers glaciation, some 580 mya (Figure 1). This Ediacaran phase is followed by the Cambrian itself, starting some 542–543 mya. In the past few years, there have been significant advances in understanding the absolute chronology of this interval of time, using both new radiometric dates and their association with Carbon-isotopic profiles (e.g., Bowring et al. 1993, Grotzinger et al. 1995, Landing et al. 1998, Condon et al. 2005), but much work still needs to be done, especially in placing key fossil localities into the emerging chronologic scheme.

Trace Fossils from the Ediacaran into the Cambrian

The fossil record of animal tracks and trails is particularly important when trying to understand the Cambrian “explosion.” While diploblasts can leave trace fossils (Collins et al. 2000), most trace fossils are thought to be the product of triploblastic activity (Budd & Jensen 2000, Droser et al. 2002). Horizontal traces appear with regularity some 555 mya and increase in size and complexity throughout the Cambrian (Droser et al. 2002) (Figure 1). The first three-dimensional traces, those that reflect vertical as well as horizontal movement in the substrate, belong to...
Figure 1
Complex anatomy of the Cambrian "explosion." Dates from Grotzinger et al. (1995), Landing et al. (1998), Gradstein et al. (2004), and Condon et al. (2005). Neoproterozoic carbonate carbon isotope curve from Condon et al. (2005), Early Cambrian curve largely from Maloof et al. (2005) but also from Kirschvink & Raub (2003), and Middle and Late Cambrian from Montanez et al. (2000). Note the wide range of values in part of the Early Cambrian; this is partly due to geographic variation, but also to variation measured in Morocco. Disparity from Bowring et al. (1993). Diversity based on tabulation by Foote (2003) derived from Sepkoski’s compendium of marine genera (Sepkoski 1997, 2002); all taxa found in the interval, as well as those that range through the interval, are counted. Short-term idiosyncrasies in the rock record can add noise to diversity curves, so to dampen that effect, taxa found in just one interval can be omitted (singletons omitted). Note that standing diversities were much lower than the values shown; many of the taxa found in a stratigraphic interval did not coexist. The boundary crosser curve (M. Foote, personal communication) gives the number of taxa that must have coexisted at the points shown; however, because traditional stratigraphic boundaries are based on times of unusual taxonomic turnover, these estimates may underestimate typical standing diversities.
DIPLOBLASTIC AND TRIPLOBLASTIC (BILATERIAN) ANIMAL PHYLA

The animal phyla are divided into three groups. First are the sponges (Porifera) that do not have organized cell layers. Second are the diploblasts, which have two primary cell layers: an outer ectodermal layer and an inner endodermal layer. The Cnidaria (corals and jellyfish) and the jellyfish-like group, the Ctenophores, are the only diploblastic phyla. All of the remaining animal phyla (31 depending on the authority) are triploblasts, also collectively referred to as the Bilateria. These grow from three primary cell layers: the outer ectoderm, the intermediate mesoderm (from which our skeleton and most of our muscles are derived), and the inner endoderm, which includes the gut. The idea that there are a large number of extinct phyla (up to 70; see Lewin 1988) is now no longer widely accepted. Over 99% of all living animals are triploblasts.

Treptichnus pedum (Droser et al. 1999). This taxon’s first appearance in Newfoundland lies a short distance below the formally defined base of the Cambrian (Gehling et al. 2001). Critical reviews of the trace fossil record are forthcoming (M.L. Droser, personal communication).

Body Fossils of the Ediacaran

There have been excellent recent reviews of body fossils of the Ediacaran (Valentine 2002, 2004; Narbonne 2005). Most of the taxa are unskeletonized and may be as large as two meters in size. A wide range of morphologies is known, but the taxonomic affinities of most are controversial. Some may well be diploblasts, for example, Charniodiscus (Conway Morris 1992, Williams 1997), whereas some are quite possibly stem-group bilaterians, such as Kimberella (Fedonkin & Waggoner 1997). However, it has also been suggested that most of the soft-bodied taxa belong to a completely different and now extinct clade of animals, the Vendobionta (Seilacher 1989, 1992, Buss & Seilacher 1994, Runnegar 1995), or might not be animals at all (e.g., see Retallack 1994, Peterson et al. 2003). The Ediacaran biota first appears shortly after the Gaskiers glacial and all but become extinct at, or shortly after (Jensen et al. 1998), the Precambrian/Cambrian boundary (Figure 1).

In the last five million years of the Ediacaran, the first lightly skeletonized forms appear (Figure 1), including the relatively simple multicone-shaped Cloudina (Grant 1990, Hua et al. 2005) and the rather more complex, but still small, Namacalathus and its relatives (Grotzinger et al. 2000). The taxonomic affinities of these forms are not known.

Body Fossils of the Cambrian

For the vast majority of the Phanerozoic, most fossils are easily classified, especially at the phylum level. However, the earliest Cambrian faunas were dominated by small...
shelly fossils, many of which are hard to diagnose, even at the phylum level. Thus, discussions of the body fossil record of the Cambrian must start with these largely problematic fossils.

**The small shellies.** The earliest Cambrian, the Manykaian (Nemakit-Daldynian), which until relatively recently was informally considered part of the Precambrian, is relatively devoid of fossils. However, by the Tommotian the small shelly fossils, many of which are only hundreds of microns in size, became much more numerous (see review by Kouchinsky 2001) (Figure 1). While many represent individual animals, others represent individual components of the armor of much larger animals (Conway Morris & Peel 1995). Some of the described genera belong to known phyla such as Brachiopoda and Mollusca. However, many are problematic, including the cambroclaves, coeloscleritopherans, cribricyatheans, machaeridians, tommotiids, as well as a diverse array of incertae sedis.

**Larger, morphologically diverse taxa.** It is not until the Atdabanian Series (Figure 1) that we find the first abundant body fossils that most would immediately recognize as belonging to typical Phanerozoic taxa, such as the first trilobites and echinoderms (Zhuravlev & Riding 2001). The first window of exceptional preservation in the Cambrian is the Chengjiang biota in Yunnan, China (Hou et al. 1991, 2004). The famous Burgess Shale biota is Middle Cambrian (Whittington 1985, Briggs et al. 1994), by which time the Cambrian explosion is all but spent. However, while the increases in disparity and diversity that characterize the Cambrian “explosion” are largely over by the end of the Early Cambrian, many crown group higher taxa are first found in the Middle and Late Cambrian and beyond (Budd & Jensen 2000).

**How Reliable Is the Fossil Record?**

The anatomy of the Cambrian “explosion” described above is based on a literal reading of the rock and fossil records. However, it is likely that evolutionary lineages have their origins in rocks older than their first observed occurrences in the fossil record (Marshall 1998). Attempts to quantify the incompleteness of the fossil record typically lead to range extensions that are small compared with observed stratigraphic ranges. For example, Valentine and colleagues (1991), in the only quantitative treatment of the suddenness of the Cambrian “explosion,” conclude that the suddenness of the adaptive radiation is real, even when the incompleteness of the fossil and rock records is taken into account. In contrast, attempts to use molecular clocks to estimate the time of origin of the animal phyla have led to much larger estimates of the incompleteness of the fossil record.

**Completeness of the fossil record and molecular (DNA) clocks.** Numerous papers using molecular (DNA) clocks have suggested a very deep pre-Phanerozoic history for the animal phyla (see Runnegar 1982 and Wray et al. 1996 for the first papers), with the possibility that the phyla may have diverged over a billion years.
ago. However, extracting accurate divergence times from DNA sequences is a subtle and difficult art, made even more complicated by the fact that the divergence times of lineages (which molecular clocks estimate) may significantly predate the time of emergence of diagnosable morphologies (which the fossil record estimates) (Marshall 1998). However, some studies (albeit a minority), including those that recognize that different clades evolve at different rates, the importance of concatenating collected sequence data, and the use of multiple and better refined fossil calibration points, place the divergence of the bilaterian phyla within the Ediacaran, and most likely after the last Snowball Earth (Aris-Brosou & Yang 2003, Peterson et al. 2004). So, while certainly still contentious, the fossil record and molecular estimates of the times of origin of the animal phyla may well be in general accord, although most agree that most phyla have at least some Precambrian history.

Quality of the fossil record. Although the fossil record typically misses the early history of major clades, it nonetheless provides much more accurate information about the time at which (fossilizable) groups became ecologically important. Thus, for example, although molecular clocks suggest that the major living groups of mammals diverged in the Cretaceous while the fossil record is largely silent about these origins (but see Tavaré et al. 2002), it is clear that these groups did not become ecologically important until after the end-Cretaceous mass extinction, where the fossil record is unequivocal about their rise in importance (Alroy 1999). Similarly, it is seems likely that the first appearance of skeletonized animals, as well as the first invasion of the substrate by animals, near the Ediacaran/Cambrian boundary represents the true first appearance of ecologically important skeletonized and burrowing forms, even though their unskeletonized forebears almost certainly had an Ediacaran, and perhaps more ancient, existence.

THE CAMBRIAN “EXPLOSION”: WHAT NEEDS EXPLAINING?

There are five major components of the Cambrian “explosion” that need to be explained: (a) the spectacular increase in animal disparity, (b) the rise in animal diversity, (c) why the time of onset of the explosion was some 543–542 mya, (d) why the duration of the explosion was some tens of million years long, and (e) why the event appears unique.

There are also two problems that emerge once we begin to examine the fossil record in some detail: (f) Where are the (largely) missing fossils of the bilaterian stem-groups, and (g) What are the phylogenetic affinities of the Ediacaran biota?

Disparity

Paleontologists distinguish between diversity, the number of species, and disparity, the morphological differences, or distinctness, between species (thus the disparity of a sea anemone, a whale, and a sea urchin is much greater than the disparity of a hundred beetles). Perhaps the most important aspect of the Cambrian “explosion” is the first
appearance of essentially all the readily fossilizable bilaterian animal phyla (Valentine 2002). Each phylum represents (for the most part) a fundamentally different mode of organization; comparing their morphologies is akin to comparing proverbial apples and oranges. Whether disparity is measured by the number of phyla, or by their constituent classes (Figure 1), the rate of accumulation of disparity during the (Early) Cambrian “explosion” is unprecedented and unique.

**Diversity**

Along with the dramatic increase in disparity comes a dramatic increase in the number of species (Figure 1), although the increase in diversity persists longer than the increase in disparity.

**Time of Onset**

Broadly speaking, the onset of the Cambrian “explosion” is in the Ediacaran when there is evidence of stem group bilaterian fossils, when the trace fossil record suggests bilaterians large enough to burrow appeared, and when large animals, the enigmatic Ediacaran biota, occur with some abundance. Nonetheless, the first three-dimensional trace fossils and the first appearance of numerous heavily skeletonized taxa occur in the Cambrian, so designating the base of the Cambrian as the time of onset of the radiation is appropriate (Figure 1). The question as to why it happened then, and not say 10 million, or 100 million, or even a 1000 million years later or earlier is both a particularly interesting and difficult question.

**Duration**

Depending on when exactly one thinks the Cambrian “explosion” began, it is clear that there is a considerable temporal anatomy to the radiation (Figure 1). From the first appearance of heavily skeletonized animals to the first body fossils of trilobites, the radiation took some 20 million years. If one starts with the first abundant trace fossils through to the end of the Cambrian, then the radiation ran for some 65 million years. Why did it not happen much faster, say in just a few million years? Why not much more slowly?

**Uniqueness**

Never before or since has there been such a dramatic emergence of animal disparity as seen in the Cambrian “explosion.” Even after the largest mass extinction of the Phanerozoic, the end-Permian mass extinction, no new phyla appeared. Why?

**Stem-Group Bilaterians**

While the fossil record of the well-skeletonized animal phyla is pretty good, we have virtually no fossils that are unambiguously assignable to the most basal stem groups...
of these phyla, those first branches that lie between the last common ancestor of all bilaterians and the last common ancestors of the living representatives of each of the phyla. Part of the reason for this is that these stem groups, by definition, lack easily diagnosable characters (Budd & Jensen 2000), but nonetheless their absence is striking. Where are they?

Phylogenetic Status of Many of the Ediacarans

Finally, as noted above, the phylogenetic status of many of the Ediacaran taxa is uncertain. These uncertainties make unraveling this time of prelude to the Cambrian “explosion” difficult (e.g., see Valentine 2002). What were their biological affinities?

What the Cambrian “Explosion” Is Not

By the end of the Cambrian “explosion” the basic body plans of the animal phyla were in place. However, compared with the rest of the Phanerozoic the Cambrian was still an unusual time. Cambrian rates of per genus origination and extinction were on average much higher and more volatile (Bowring et al. 1993, Foote 2003) than the rest of the Phanerozoic. Further, Cambrian diversity was considerably lower than most of the rest of the Phanerozoic (Sepkoski 1997). It was not until the subsequent Ordovician radiation that typical Phanerozoic diversities were reached. It was also not until the Ordovician that typical Phanerozoic ecosystems emerged. So, the Ordovician radiation was also a time of great significance (Webby et al. 2004). And animals did not invade the terrestrial realm until later in the Paleozoic, and they did not master the aerial realm until the later Paleozoic (arthropods) and in the Mesozoic (pterosaurs and birds) and Cenozoic (bats); there are many other major events in the history of animal life other than the Cambrian “explosion.”

PROPOSED CAUSES (AND THEIR EXPLANATORY POWER)

A great number of causes, or at least necessary conditions, for the Cambrian “explosion” have been suggested. Here, I have selected a subset of these, with the goal of giving a sense of the range of ideas that have been proposed and as a means of dissecting out the explanatory power different kinds of explanations provide with regard to each of the components of the explosion that need to be explained.

Environmental Explanations

There has been a resurgence in this class of explanation, partly owing to increased interest in the Cambrian “explosion” within the wider geoscience community. There are two classes of environmental explanation. First, there are those that point to the fact that the environment had to have been conducive to complex animal life before there could have been complex animal life. Perhaps the most dominant explanation of this kind is associated with the accumulation of enough oxygen to sustain large animals. The second class of environmental explanation is often motivated by the
discovery of some unusual, large-scale, environmental change that occurred at about
the time of the initiation of the radiation. Three such changes include (a) the ge-
ological evidence that points to the Snowball Earths, (b) the large negative carbon
isotopic anomaly at the Precambrian/Cambrian boundary, followed by a period of
unusual volatility, and (c) paleomagnetic evidence for true polar wander (whereby, on
geological timescales, the continents rapidly shifted their positions).

**Increase in atmospheric oxygen levels.** Budd & Jensen (2000) provide an in-depth
analysis of the relationship between atmospheric oxygen levels and the capacity to
support large animals. It has long been argued that low levels of oxygen would have
prevented the evolution of large animals (Knoll & Carroll 1999), and thus that the
Cambrian “explosion” might have been triggered by the accumulation of enough
oxygen to support large animals. Testing this hypothesis is made difficult by the fact
that we do not know the minimum oxygen requirements of the first animals, nor do
we have precise enough values of Proterozoic oxygen levels to put firm constraints
on just how much oxygen there was through this interval [Canfield (2005) provides
an informative review of the complex controls and ways of measuring oxygen levels
in the Precambrian]. However, many feel that there was sufficient oxygen to support
large animals well before the Cambrian “explosion” (Knoll & Carroll 1999).

Many of the Ediacarans had large body sizes, so there was clearly enough oxygen
to support large animals by the latest Neoproterozoic (assuming at least some were
animals!). The question still remains as to whether the appreciable lag between the
emergence of the Ediacaran fauna and the skeletonization event that heralds the
beginning of the Cambrian (some 30 million years later) is due to the waiting time
for some other key factor (such as the development of the bilaterian developmental
system or the first morphologies of predation), or whether oxygen levels had to
increase beyond mid-Ediacaran levels to support the Cambrian radiation, perhaps
because higher levels were required to support the physiologies and generation of
tissues [e.g., collagen (Towe 1970)] needed for predation, etc. That is, it is still an
open question as to whether there was a single oxygen threshold that needed to be
crossed for the emergence of the bilaterian phyla, or whether there was a series of
thresholds that had to be crossed.

**Snowball Earths.** With the growing evidence of significant global late Neoprotero-
ozoic deep freezes, the Snowball Earths (Kirschvink 1992, Hoffman et al. 1998, see
discussion in Peterson et al. 2005), it has been argued that these may have played
a critical role in the Cambrian “explosion,” given that the first unambiguous and
continuous fossil record of animals (the first of the Ediacaran biota) appears shortly
after the last of these freezes, the Marinoan, which occurred some 635 mya, and
the Gaskiers (**Figure 1**), dated at 580 mya (but which was probably only a regional
glaciation). In terms of the Cambrian “explosion,” this explanation suffers the same
problem that the explanation based on increased oxygen levels suffers: The Snow-
ball Earths seem to be too early to alone account for the radiation (and, ignoring
the Gaskiers glacial, even too early to have played a major role in the emergence
of large animals in the Ediacaran). It is also hard to see how a major environmental
catastrophe could have lead to fundamentally new levels of developmental and morphological organization. However, the fact that the animal fossil record begins after the last of these great Neoproterozoic glaciations is certainly suggestive, and perhaps the link to animal evolution (especially the divergence of the phyla which may have predated their first appearance in the fossil record by a substantial degree) is through the effects the Snowball Earths had on global oxygen levels, or some other key, as yet unidentified, environmental factor.

**Carbon isotopic anomaly at the Precambrian/Cambrian boundary.** The presence of an unusually large negative carbon isotopic anomaly, of between 7 to 9 per mil, at the Neoproterozoic/Cambrian boundary (Figure 1) points to an environmental disturbance at a critical time in Earth history, so it is perhaps natural to assume it played a role in the Cambrian “explosion” (Knoll & Carroll 1999, Amthor et al. 2003). While there is no generally accepted explanation for the cause of the anomaly (but see Rothman et al. 2003), it is has been assumed that it represents a sufficiently large environmental disturbance that it caused a mass extinction of the Ediacaran biota, and in doing so laid the foundation for Cambrian radiation (Knoll & Carroll 1999, Amthor et al. 2003). By analogy with the end-Cretaceous extinction, where it is generally believed that the mass extinction of the incumbent dinosaurs paved the way for the radiation of mammals, it is argued that by removing the incumbent Ediacaran biota, the stage was now set for the Cambrian “explosion.”

However, using the recovery from the end-Cretaceous mass extinction as an analogy for the Cambrian “explosion” does not seem appropriate; in the case of dinosaurs and mammals, both are members of the same clade, the Amniota; there is no increase in disparity, nor a significant change in diversity as a consequence of the end-Cretaceous extinction; all we see is a replacement of various diapsid clades (dinosaurs, pterosaurs, ichthyosaurs, plesiosaurs, etc.) with an ecologically and morphologically equivalent sister clade [the synapsids (mammals)]. This is in sharp contrast to the Cambrian explosion, where there is a dramatic increase in both disparity and diversity.

As with the Snowball Earths, it is hard to see how a simple (even if large) environmental disturbance can lead to an increase in disparity. In fact, one wonders whether the carbon isotopic anomaly might be independent of the unfolding Cambrian “explosion,” or even a consequence of the radiation. Interestingly, when the diversity of animals was still relatively low in the Early Cambrian, the carbon isotopic values show an unusual volatility (Figure 1), similar to the volatility seen in the Early Triassic (Payne et al. 2004) when diversity was also low after the end-Permian mass extinction. We have much to understand about the cause of these isotopic excursions and the complex relationships between global carbon budgets, the biota (and where that biota lives in the water column), global oxygen levels, etc.

**True polar wander.** Starting with the controversial hypothesis of Cambrian true polar wander (Kirschvink et al. 1997), Kirschvink & Raub (2003) have linked this hypothesis to the unusual carbon isotopic record of the Cambrian, proposing that the rapid movement of the continents (on geological, but not ecological timescales) caused huge methane “burps” (Kirschvink & Raub 2003), which in turn triggered the
Cambrian “explosion.” They argue that the increased temperatures induced by the methane releases drove the observed increase in diversity. No explanation is offered as to why an increase in diversity, per se, should have led to new levels of disparity. This remains an intriguing hypothesis.

Summary. While it is clear that the environment needs to have been conducive to the evolution of large animals for the Cambrian “explosion” to proceed, none of the environmental explanations address why an environment permissive of complex animal life should necessarily lead to the evolution of complex animal life, and especially why we should see a shift from diploblastic-grade organisms to complex triploblasts. In the case of the perturbation explanations, it is unclear how the time of the perturbation (Snowball Earths or onset of the unsettled carbon cycle in the Cambrian) relates to the increased disparity and diversity, and how it might explain either the duration or uniqueness of the event (although this is not saying that a relationship might not be found).

Developmental Explanations
Animals cannot evolve if the genes for making them are not yet in place. So clearly, developmental/genetic innovation must have played a central role in the radiation. However, until recently, we have known so little about the genetic basis of animal form that developmental explanations have been necessarily rudimentary. But this is now rapidly changing.

Origin of the bilaterian developmental system. There is an ongoing revolution in our understanding of the genetic basis of morphological form (Carroll et al. 2001, Davidson 2001), and several workers point to the origin of the bilaterian developmental system, including the origin of the Hox genes, etc., as the primary cause of the “explosion” [e.g., Peterson & Davidson 2000 (but see Peterson et al. 2005), Erwin & Davidson 2002], and there can be little doubt that the origin of the bilaterian developmental system was critical to the Cambrian “explosion.” We still understand little of how and when the system originated, but clearly gene duplication was one major component (Carroll et al. 2001), for example, with the origin of the Hox cluster from the paraHox cluster of genes.

Developmental data have also dramatically changed the way we look at the origin of disparity. While we can’t determine the genetic composition of extinct taxa directly, we can use the phylogenetic distribution of developmental genes in living species to make inferences about the genetic capacities of their last common ancestors (Carroll et al. 2001). While we only have detailed genetic data from a very few species (the so-called model systems), the last common ancestor of these species also happens to be the last common ancestor of all the bilaterian phyla! So, in fact, we can make quite powerful inferences about the genetic capacities of animals that lived at the Precambrian/Cambrian transition (Carroll et al. 2001).

Spectacularly, many key developmental genes and gene families are shared between all animals. Comparative development genetics has shown us that while
morphologically the animal phyla might be “apples and oranges,” genetically they are fundamentally comparable. Figure 2 summarizes some of the most important developmental genes that are inferred to have been ancestral to all bilaterian animals, along with the morphologies they initiate. As the genomes of diploblasts and sponges are explored, it is also becoming clear that many of these genes predate the first bilaterians (e.g., Erwin & Davidson 2002, Finnerty et al. 2004).

However, the significance of the presence of these shared genes is still an open question (Erwin & Davidson 2002). Does the presence of the *tinman/NK2.5* gene in the last common ancestor of the bilaterians indicate the presence of a heart and circulatory system in that ancestor, or does the gene simply mark a special type of muscle that was later and independently co-opted to initiate the development of fully developed circulatory systems in different lineages (Erwin & Davidson 2002)? If the latter view is correct then there must have been considerable developmental sophistication en route from the last common ancestor of the bilaterians to the living phyla.

**The combinatorial genome.** What is clear is that the animal development system is combinatorial in nature. Changes in the way the genes are wired is a major, if not the major, source of morphological innovation (e.g., see Averof & Patel 1997, Carroll et al. 2001, Davidson 2001). The combinatorial nature of the developmental system is very important. For example, Kauffman and colleagues (Sole et al. 2003), with the Cambrian “explosion” in mind, show how small increases in the complexity of a combinatorial developmental system (including developing relatively simple rules for changing gene regulation and cell-cell interaction) can lead to an extraordinary range of stable spatial patterns of gene expression. Thus, they argue that a relatively small increase in underlying genomic complexity can lead to rich morphogenetic potential. This conceptualization is also reflected in Wolfram’s (2002) *A New Kind of Science*, where he also advances the idea that much of biological complexity probably arises from the operation of relatively simple rules within large numbers of cellular automata (cells).

**Canalization/developmental entrenchment.** Developmental explanations for the Cambrian “explosion,” particularly its uniqueness, predate the spectacular recent advances in our understanding of the genetic basis of development (see above). For example, Valentine (1986), among many others, argued that early in the Phanerozoic, development was less constrained, or canalized, than it is now. A similar notion flows from Wimsatt’s (1986) discussion of the developmental lock. The common thread in this literature [see Raff (1996) for a deeply informed discussion] is the idea that as evolution proceeds, the developmental stages responsible for laying down the body plans are overlain with down-stream genetic pathways, making it progressively harder to modify those earlier pathways; the developmental stages when the phylum-level body plans are laid down become entrenched and the body plans become frozen in place. This is a very appealing idea, but we still have a lot to learn about how encumbered the gene networks responsible for morphogenesis really are, and it is quite possible that they are not as encumbered as one might think (Valentine 1995).

Paleontological data also have some bearing on this issue. It has long been noted that intraspecific variation is often largest in the earliest members of a clade.
Figure 2
A few of the key developmental genes, and the morphologies they may have conferred, inferred to have been present in the last common ancestor of all the bilaterian phyla (the ur-bilaterian), based on the phylogenetic distribution of developmental genes in mouse and fly. Top: The anterior/posterior (A/P) axis may have been subdivided by nested, overlapping domains of Hox gene expression. The dorsal/ventral (D/V) axis may have been controlled by ancestral genes of the short gastrulation (sog)/chordin and TGF-β families. Middle: Different tissue layers were regionally patterned along the A/P axis, including the gut (paraHox gene cluster) and nervous system (orthodenticle (otd), empty spiracles (ems), Hox genes). Segmentation (seriation) may have been present through the action of the genes ancestral to engrailed and hairy. Bottom: Ancestral photoreceptor organs (Pax6), circulatory pump (tinman/NK2.5) and outgrowths/ingrowths of the body wall (Distal-less (Dll)) are also inferred to have been part of the morphogenetic potential of the ur-bilaterian. From Carroll et al. (2001), published with permission.
For example, some Cambrian trilobite species exhibit varying numbers of thoracic segments, whereas post-Cambrian trilobites are much more stereotyped in their segment number. Observations of this kind have been used to argue in favor of canalization. However, Hughes et al. (1999) document a Silurian trilobite species with Cambrian levels of segment variability. One possibility is that this lineage had broken the pattern of entrenchment that occurred after the Cambrian, but a more plausible hypothesis, advocated by Hughes et al. (1999), is that segment number did not become more entrenched with time, and that the observed decrease in the variation in segment number was controlled by increased ecological/functional constraints, not entrenchment (for example, segment number might be expected to be more constrained once trilobites developed enrollment as a form of defense).

**Summary.** The developmental class of explanation, per se, does not address the question of why the origin of such a system should, ipso facto, lead to increased diversity or disparity. In fact, if at least one Ediacaran is a bilaterian (Kimberella, Spriggina, Dickinsonia, or Arkarua, for example), then the bilaterian developmental system existed at least a few tens of millions of years prior to the Cambrian “explosion,” suggesting something more than just developmental innovation might be needed to account for the “explosion.” I will return to this issue below.

**Ecological Explanations**

Ecological explanations (e.g., Zhuravlev 2001) have been overshadowed in the recent literature by the developmental and environmental explanations. Typically, these explanations center on the invention of new trophic capacities, whether predation (Evans 1912, Hutchinson 1961, Vermeij 1990, Bengtson 2002) or cropping (e.g., Stanley 1973, 1976), as a major cause of the Cambrian “explosion.” Many center on predation, specifically in an effort to explain the massive skeletonization event that characterizes the fossil record of the “explosion” (Vermeij 1990). All of these have a common thread of coevolution, escalation, or arms races (Vermeij 1987, 2004). Some workers have specifically focused on the ecological ramifications of the invasion of the water column by small animals to avoid predation on the sea floor (Signor & Vermeij 1994, Peterson et al. 2005), an idea that finds support from the fossil record of pelagic algae, which evolved antipredator defenses in the Early Cambrian (Butterfield 1997, 2001).

It is hard to see how these ecological factors could not have been a major component of Cambrian radiation, and any satisfying explanation for the radiation must take stock of these factors. Open questions include why the radiation should have been unique, and why it should have taken as long as it did.

Valentine (1980) and Valentine & Walker (1986) have approached the uniqueness problem by invoking a theoretical niche space. They argue that if large morphological changes are less likely to occur than small changes, and if the probability of a new major innovation being successful depends on there being limited competition at the time the large morphological change occurred (so that the nascent higher taxon has a chance to adjust to its new ecological niche), then at the initiation of the radiation many higher taxa should appear, but as the ecosystem’s niches fill up, it will
be progressively harder for new body plans to become established; the world’s niche space becomes saturated. The reason, then, why the end-Permian mass extinction did not lead to new phyla is that the relatively few Triassic survivors still had most of the niche space occupied despite the decimation, so that new body plans could not get an ecological foothold. The Valentine model implies that there is a wide range of other potential phyla out there, and that it is the lack of ecological opportunity that keeps them from becoming manifest.

More recently, some have focused on some very specific morphologies that are clearly important in biotic interaction, such as the evolution of macroscopic eyes (and color perception) as being key (Parker 1998). However, while it is clear that the evolution of eyesight is crucial to the specific morphologies that have evolved (e.g., camouflage, etc.), there are more ways of locating organisms than through eyesight (e.g., see Marcotte 1999), so there almost certainly would have been some sort of radiation even if large compound eyes had not evolved in the Cambrian.

**Theoretical Explanations**

Others have taken a more theoretical approach to understanding the Cambrian “explosion,” especially Stewart Kauffman (1993, chapter 3). Specifically, he has been interested in the exploration of fitness landscapes. He notes that for rough landscapes (those with many peaks), the rate of evolution dramatically slows as the landscape is explored. By analogy with the Cambrian “explosion,” his ideas translate into an argument that one should expect to see a rapid burst of evolution followed by a steady decline in rate, simply as a consequence of the time it takes to find progressively more optimal solutions.

So unlike Valentine’s (1980) niche saturation model where selection (ecological opportunity and competition) limits the rate of origination of higher taxa as the ecospace fills, in Kauffman’s model the total disparity available is preset (as defined by the fitness landscape), and it is the inevitable decrease in the rate at which fitter genotypes (morphologies) are found as the landscape is explored that produces the burst in higher taxon origination that characterizes the Cambrian “explosion.” That is, in Kauffman’s model the Cambrian “explosion” corresponds to the initial steps in the exploration of the fitness landscape.

The challenge for this class of explanation is understanding how the theoretical constructs related to the real world. In the case of Kauffman’s NK models, the roughness of the landscape is controlled by $K$, the number of interactions between the $N$ genes. However, it is difficult to meaningfully assign a value of $K$ to a set of genes, and it is even more difficult to interpret these landscapes in morphological terms; i.e., the NK model does not explicitly incorporate the phenotype into the calculation of the fitnesses.

**Summary: Explanatory Power of Posited Causes of the Cambrian “Explosion”**

It is clear that environmental, developmental, and ecological factors must have played a role in the Cambrian radiation; however, the questions still stand as to which factors
are most important, which variant(s) of each class of explanation is most likely correct, and how the various factors interact.

Table 1 presents a summary of the explanations discussed above, showing the extent to which each is able to account for the various core aspects of the Cambrian “explosion” that needs to be explained. As with all simple classifications, some subtlety has been lost in constructing the table; nonetheless, a few generalities emerge.

First, environmental explanations are not (currently) sufficient to explain the emergence of animal disparity, the animal phyla. Second, both developmental and ecological explanations seem important in explaining the origin of animal body plans, so what is the relationship between the two? Third, most explanations do not explicitly explain why there should be an increase in diversity through the radiation; usually, an increase in diversity seems to simply follow from the arguments for the origin of disparity. Fourth, only environmental explanations have any direct bearing on the time of onset of the explosion. But these only speak to the emergence of the conditions conducive to the onset of the radiation, but are unable to address how long after those conditions are met whether or when the radiation might proceed. Fifth, the only hypotheses that address the issue of the duration of the radiation do so in terms of how long it takes to explore the combinatorial genetic potential of the early genome, although we have no a priori understanding of why this should be on the timescales of tens of millions of years, rather than an order of magnitude (or more) faster or slower. And finally, we have two classes of explanation that bear on the uniqueness of the event, the developmental (entrenchment), as well as the ecological (niche saturation).

It is clear that the environment must be permissive of animals before they could have evolved. It is also clear that the genetic machinery for making animals must have been in place, at least in a rudimentary way, before they could have evolved. And finally, organisms must be able to leave viable offspring to survive and evolve, so ecology had to be important too.

A FRAMEWORK FOR INTEGRATING ENVIRONMENTAL, ECOLOGICAL, AND DEVELOPMENTAL DATA

Fitness Landscapes (of the Morphogenetic Kind)

Following the rich tradition begun by Sewell Wright (1931, 1932), fitness landscapes provide a fruitful way of thinking about the interaction between developmental potential and evolutionary success, the ability to pass one’s genes on to the next generation. The coordinate system in most fitness landscapes is based on genes and their alleles. However, the Cambrian “explosion” finds its expression in the fossil record morphologically, so it is more appropriate to use a morphogenetic rather than a genic coordinate system. Hence, theoretical morphospaces (McGhee 1999), where each axis of the landscape represents a distinct morphogenetic rule and where the position along each axis corresponds to a particular variant of the rule, is appropriate here. Every point in the space corresponds to a unique morphology that arises from the morphogenetic rules.
Table 1 Evaluation of the power of individual factors hypothesized to be relevant to an understanding of the Cambrian “explosion” to account for the major aspects of the “explosion” that need to be explained. See text for discussion

<table>
<thead>
<tr>
<th>Key event or process</th>
<th>Aspect of the Cambrian &quot;explosion&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Burst of disparity</td>
</tr>
<tr>
<td>Environmental</td>
<td></td>
</tr>
<tr>
<td>Increased oxygen levels</td>
<td>Yellow</td>
</tr>
<tr>
<td>Snowball Earth</td>
<td>Gray</td>
</tr>
<tr>
<td>Extinction of incumbents</td>
<td>Red</td>
</tr>
<tr>
<td>True polar wander and methane “burps”</td>
<td>Orange</td>
</tr>
<tr>
<td>Developmental</td>
<td></td>
</tr>
<tr>
<td>Origin of bilaterian developmental system</td>
<td>Green</td>
</tr>
<tr>
<td>Canalization/developmental lock</td>
<td>Yellow</td>
</tr>
<tr>
<td>Ecological</td>
<td></td>
</tr>
<tr>
<td>Origin of predation/arms race</td>
<td>Orange</td>
</tr>
<tr>
<td>Niche saturation</td>
<td>Red</td>
</tr>
<tr>
<td>Theoretical</td>
<td></td>
</tr>
<tr>
<td>Exploration of fitness</td>
<td>Red</td>
</tr>
<tr>
<td>Evolution of fitness landscapes (this review)</td>
<td>Red</td>
</tr>
</tbody>
</table>

*Green: Proposed factor is both necessary and sufficient to explain designated aspect of the “explosion.” Yellow: Proposed factor is necessary, but not sufficient, to explain designated aspect of the “explosion.” Orange: Proposed factor is consistent with designated aspect of the “explosion,” but importance/relevance of factor is not yet established. Red: Proposed factor is not in accord with the fossil record. Gray: Proposed factor either (a) implicitly or explicitly assumes aspect of the “explosion” will follow from the trigger or generative event, without a detailed analysis of why this should be or (b) naturally follows from proposed factor (e.g., an increase in diversity will naturally follow an increase in disparity).

**1. Event too early to be immediately causal, but may have permitted evolution of Ediacaran biota, a necessary prelude to the Cambrian “explosion.” 2. If the stem lineages were both small and unskeletonized, then we would not expect to see them in the fossil record. These animals will have been small if there was limited oxygen, or if the controversial developmental set-aside cell theory (Davidson et al. 1995; but see Peterson et al. 2003 for alternate view) holds, or if large size was a consequence of predation. 3. If roughening (see text) is correct (i.e., the invention of complex food webs) then origin of bilaterian developmental system alone is not enough to account for disparity increase. 4. In the absence of a combinatorial developmental system we might not expect a burst in disparity. 5. Explanation certainly explains uniqueness of skeletonization event, but it is not as obvious why body plans should only appear during the initial escalatory radiation. 6. Duration of radiation represents rapid initial steps in exploration of animal fitness landscape. 7. Duration of radiation controlled by the time it took to explore roughened fitness landscape. Post-Cambrian innovation was due to further evolution of landscape. 8. Uniqueness follows from the fact that the bilaterian developmental system only has so many locally optimal designs given the needs that must be met, which were made manifest as evolution explored the roughened landscape.
To convert a theoretical morphospace into a fitness landscape, each morphology must be assigned a fitness, a measure of how well it is able to perform all the functions necessary to pass offspring successfully into the next generation. By contouring the fitness values for all the morphologies, the fitness landscape is established. There is a vast literature on fitness landscapes (almost exclusively of the genic kind) and how evolution explores them (e.g., see Gavrilets 2004), but relatively little on how the landscapes themselves might evolve.

It is important to keep in mind some limitations of the metaphor of the fitness landscape. First, movement is measured in terms of change in the morphogenetic rules, several steps removed from the genetic changes that are responsible for those rules. That is, we are still profoundly ignorant of how changes in the genome translate into changes in morphology, despite the spectacular advances we have made in understanding the genetic basis of morphogenesis.

**Startling Results from Computer Simulations of Evolution on Fitness Landscapes**

Important to explaining the Cambrian “explosion” is understanding what controls the number of peaks on the landscape (the expected diversity), and the morphological disparity of those local optima. Before the recognition of the combinatorial nature of the genetic machinery, it was generally assumed that one would need new genes, new morphogenetic potentials, to evolve new morphologies; i.e., one would need to increase either or both the size and dimensionality of the landscape to account for the Cambrian “explosion.” However, some very elegant computer simulations by Karl Niklas show how increases in diversity and disparity may also be achieved through the roughening of fitness landscapes, i.e., without the need for new genes and morphogenetic potentials.

**What determines the roughness of a fitness landscape?** Niklas (1994, 1997, 2004), using a six-dimensional morphogenetic fitness landscape for land plants, provides deep insight into what controls the roughness, the density of peaks, on a landscape. He demonstrates that the roughness is controlled by the number of needs the organism must satisfy (rather than the degree of interaction between the genes, as is the case of Kauffman’s genic fitness landscapes). The insight is made possible by the fact that Niklas can evaluate the fitness of each morphology of his computer-simulated plants based on quantitative measures of each plant’s ability to perform realistic ecological tasks, including the ability to produce and disperse seeds, harvest light, avoid mechanical breakage of its branches, and minimize the risk of desiccation through minimizing its surface area.

The roughness of the landscape stems from trade-offs associated with having to maximize the performance of all required tasks simultaneously. The central importance in trade-offs is key here, and so I have elevated its importance by labeling it a principle (Marshall 1995, 2003), the Principle of Frustration.

**The principle of frustration and its role in roughening fitness landscapes.** Usually trade-offs come to the fore in discussions of specific functions, often when trying
THE PRINCIPLE OF FRUSTRATION

This principle captures the notion that different needs will often have (partially) conflicting solutions, so that the overall optimal design for an organism will rarely be optimal for any of the specific tasks it needs to perform (i.e., there are trade-offs).

to explain why a specific functional system of an organism does not have a fully optimal design, or performance, from an engineering point of view.

However, Niklas’ computer simulations show the central importance of the Principle of Frustration in shaping the overall morphology of organisms [the name is inspired by Kauffman’s work, e.g., Kauffman (1993)]. When Niklas selects for just one task, the number of local optima in his plant morphogenetic fitness landscape are few, typically between one and three. However, as he increases the number of needs that must be met, frustration sets in, and both the number of local optima increases (with just 4 needs there are some 20 locally optimal solutions), as does the disparity of the locally optimal solutions (Figure 3). Critically important is that the increase in diversity and disparity does not involve any change in the developmental system. As the number of frustrated needs increases, the fitness landscape roughens, and the increased disparity flows from the rich combinatorial potential embodied in the six simple morphogenetic rules.

It is easy to see why this perhaps counter-intuitive result occurs once one recognizes that an optimal solution for one need will typically compromise the ability of the organism to perform some other need, a principle well understood by engineers of complex systems (Csete & Doyle 2002). In the case of Niklas’ plants, maximizing reproductive success requires many branches (actually branch tips, where the seeds form) as high as possible to maximize dispersal. Maximizing light harvesting, while also requiring many branches, requires the branches to be shorter as they get higher to reduce overshadowing, and the branches need to spread out horizontally to maximize light interception, rather than vertically. Minimizing the risk of breaking branches means limiting horizontal branch length, which flies in the face of the needs of light interception. And finally, minimizing surface area means limiting branch length, regardless of orientation, which compromises both reproductive success and light interception. Some of the trade-offs are in branch length, some are spatial (the need for horizontal versus vertical branches), others a combination of both. The key point is that when all tasks need to be performed, the trade-offs combine to produce a wide range of local optima, given the rules for making the plants.

Thus, it is frustration that leads to an increase in the roughness of a fitness landscape as the number of needs increase (Figure 3). While the number of local optima in a fitness landscape will clearly depend on the specific morphogenetic system (e.g., whether we are dealing with plants or animals, etc.) and on the range of environments that system finds itself in (e.g., terrestrial, aquatic, polar, tropical, etc.), the roughness of the landscape will also usually depend on the number of needs that must be met, or tasks that need to be performed.
Figure 3

Roughening of a fitness landscape. (a) Cartoons of two-dimensional slices (the x-y planes) through hypothetical, high-dimensional morphospaces (the third dimension, the z-axis, is used to designate the fitness of each morphology). The number of peaks is a function of the number of needs the morphologies must satisfy. In the case of Niklas' computer simulation of plants, based on a six-dimensional morphogenetic space, the roughness rises from 1–3 peaks to 20 as the number of needs rise from 1 to 4. (b) Morphologies produced by Niklas' computer simulations (from Niklas 2004). Upper row: The three locally optimal morphologies when only one need must be met, the minimizing of mechanical failure (breakage) of the branches. Lower row: The 20 local optima when fitness is based on four needs: reproductive success, light interception, resistance to mechanical failure, and minimizing desiccation (see text). Each morphology is arbitrarily assigned a peak in the corresponding fitness landscape in (a). (c) Using Niklas' analysis as an analogy for the Cambrian explosion, three representative Ediacaran morphologies and 20 representative Cambrian morphologies of triploblastic animals from the Chengjiang and Burgess Shale biotas are shown. Drawings by Samar Bush, modified from Briggs et al. (1994) and Hou et al. (2004).

Realized versus unrealized morphologies in the real world. When the morphologies produced in Niklas’ simulation are compared to those seen in the fossil record, there is a remarkable correspondence (Niklas 1997). To a first order, Niklas is able to replicate in a computer what is seen in the fossil record by the end of the Devonian, the period of greatest gross morphological innovation in the terrestrial invasion by plants. The startling possibility is that evolution has found essentially all
the locally optimal ways of being a terrestrial plant (ignoring the fine morphology associated with leaves, reproductive organs, roots, etc., as well as major modifications in the way living plants grow and reproduce compared with these early plants), and that it explored the morphogenetic space in just about one geological period.

The Niklas study opens up the possibility that evolution is able to find essentially all the locally optimal morphologies consistent with a given underlying developmental system on geological timescales. That is, all the processes associated with variation (point mutation, recombination, hybridization, gene conversion, insertion and deletion, post-transcriptional changes in mRNA processing, etc.) are able to effectively explore fitness landscapes on geological timescales; evolution is able solve the \( np \)-hard problem of exploring the rich combinatorial potential embedded in the genome in the order of 10–20 million years.

Viewing the Cambrian “explosion” in the context of the evolution of fitness landscapes opens up the possibility that uniqueness of the Cambrian “explosion” may simply represent the exhaustion of ecologically viable alternatives that can be generated by the bilaterian developmental system (Table 1), rather than being due to ensuing developmental limitation (developmental entrenchment) or ecological saturation, which simply prevented new body plans for gaining a foothold.

ROUGHENING AND THE CAMBRIAN “EXPLOSION”

For most clades, for most of their history, I suspect their fitness landscapes neither roughen nor smooth appreciably. The fundamental needs of organisms are set, the ecological roles that they play are in place, and so the panoply of needs that must be met, while perhaps bewildering in detail, are relatively stable for long periods of geologic time. Roughening is not a common mode of evolution.

However, I suspect that roughening played a major role in driving both diversity and disparity in the Cambrian “explosion” of animals (Marshall 2003). As discussed above (see also Figure 2), the phylogenetic distribution of key developmental genes in living species suggests that the basic developmental toolkit, a combinatorial toolkit, for bilaterian animals was already in place prior to the radiation (Carroll et al. 2001), certainly before the end of the Ediacaran. So, what roughened the basal bilaterian landscape?

The Arms Race Roughens the Garden of Ediacara

Compared with Phanerozoic animals, the Ediacaran biota are remarkable by their lack of macroscopic sensory organs (compound eyes, antennae, setae, etc.), the virtual absence of macroscopic organs for interacting with other organisms or the environment (there are no legs, swimming appendages, claws, etc.). They show no signs of predation, except for bore holes in some of the earliest skeletonized fossils, *Cloudina*, from the latest Ediacaran (Hua et al. 2003). It appears, indeed, to have been the Garden of Ediacara (McMenamin 1998); macroscopic adult body/adult body interactions were minimal. As paleoecologists recognize, fully developed ecosystems are far from being realized in the Ediacaran (e.g., see chapters in part II of Zhuravlev & Riding 2001).
With the advent of ecological interactions between macroscopic adults (especially interactions associated with predation) (see section on Ecological Explanations for the Cambrian “explosion,” above), the number of needs each organism had to meet must have increased markedly: Now there were myriad predators to contend with, and a myriad number of ways to avoid them, which in turn led to more specialized ways of predation as different species developed different avoidance strategies, etc. Even with no changes in the bilaterian developmental system as it existed by the end of the Ediacaran, the diversity and disparity of animals should have risen sharply, as the fitness landscape roughened in response to dramatic increase in the level of frustration (Figure 3). The combinatoric richness already present in the Ediacaran genome was extracted through the richness of biotic interaction as the Cambrian “explosion” unfolded (Marshall 2003).

I offer this as a null hypothesis explanation for the Cambrian explosion. However, it is quite likely that the size of the fitness landscape (for example, through gene duplications and through increases in the number and complexity of the cis-regulatory elements; see Valentine 2000, Shubin & Marshall 2000) and the dimension of the landscape (as new developmental capacities evolved) increased as well. But if we could go back and replay the history of life, limiting mutational events to simply rewiring the existing genetic networks, we still would have seen a spectacular explosion of new complex animal forms. That is, I suspect that roughening was the primary driver of the Cambrian “explosion,” given that the environment was conducive and the bilaterian developmental system was in place.

This extension of Niklas’ work on the evolution of plant fitness landscapes to the evolution of animals unites the ecological notion of an arms race, with the discovery of the bilaterian developmental system and the recognition that this system is combinatorial in nature. It suggests that if we replayed the tape of life, with the same developmental programs, we would see similar morphologies each time: In the debate that sprung up across the Atlantic between the importance of contingency (Gould 1989) and inevitability (Conway Morris 1998, 2003) in the history of life (to oversimplify the issue somewhat!), perhaps the truth of the matter lies more to the East than most would be willing to grant.

On the other hand, if the environmental conditions necessary for the evolution of large metazoans had occurred earlier, when the animal developmental system was perhaps in a simpler state, then perhaps the body plans seen would have been different; while the environment, per se, probably exerts very little control on patterns of observed disparity and diversity, perhaps the time the environment became conducive to the evolution of animals played a major role in the morphologies that evolved.

**Phylogenetic Status of the Ediacarans and Missing Bilaterian Stem Groups**

If the ideas of Niklas can be appropriately transferred to the Cambrian “explosion,” they offer two possibilities with regard to the phylogenetic status of many of the Ediacaran taxa as well as the missing bilaterian stem groups. First, if one sketches the basic morphology one might expect from the first bilaterians, based on the inferred
morphogenetic potential derived from the ancestral set of developmental genes, one obtains a morphology that is at least by crude measure, similar to many of the Ediacarans (compare the ur-bilaterian of Figure 2 with the representative Ediacarans shown in Figure 3). Perhaps some of the Ediacarans are in fact the missing bilaterians stem groups. Indeed, a number of taxa have been variously assigned as stem groups of existing phyla, for example, *Arkarua* as basal echinoderm (Gehling 1987); *Kimberella* as basal mollusk (Fedonkin and Waggoner 1997); and perhaps taxa like *Dickinsonia* and *Spriggina* are basal ecdysozoans, the relatively recently recognized grouping of phyla initially recognized on the basis of 18S rRNA data (for review see Halanych 2004) that include Annelida and Arthropoda, among others.

**TOWARD A FULLY INTEGRATED EXPLANATION FOR THE CAMBRIAN “EXPLOSION”**

There are many issues that remain. In terms of understanding the origin of the animal phyla, to what extent is roughening alone responsible for the morphologies seen? While the basic developmental systems used by bilaterian organisms, and most likely diploblasts as well (e.g., see Finnerty et al. 2004), was in place in the late Neoproterozoic, since that time there have been duplications of many key genes and gene families, as well as their regulatory elements (Valentine 2000, Carroll et al. 2001). To what extent was this increase in genomic complexity important in the Cambrian “explosion”? We also need to learn how to compare the animal phyla morphogenetically, not just developmentally. Recent progress in this area includes Jacobs and colleagues’ (2005) characterization of the ways in which the process of terminal addition in the basal bilaterian has been modified in different phyla. Through the process of comparative morphogenesis, we might yet be able to develop a morphogenetic space for animals and begin to test the ideas initiated by Niklas’ computer simulations.

In terms of diversity, if roughening was so important in the Cambrian radiation, why is there an even larger increase in diversity in the Ordovician radiation?

In terms of the duration of the radiation, what is the primary control on the rate of the radiation? Is it the time it takes evolution to explore the landscape (which presumably is controlled by the rate at which *cis*-regulatory changes that lead to new morphologies occur, among other genomic processes), or does the roughening occur piecemeal, where a new species helps roughen the landscape of another clade, which in turn, roughens the landscape of another clade, etc. That is, is there a steady dance as the fitness landscapes of each of the major clades coevolve? This dance will be controlled by the construction of Phanerozoic-style food webs from their Ediacaran precursors (e.g., see Butterfield 2001, Bengtson 2002, Peterson et al. 2005), so paleoecological analysis will be central to understanding this coevolution. And if this is true, why does phylum-level innovation die away as the Cambrian unfolds? Is it because with each coevolutionary step, development becomes more entrenched, or is it because ecological escalation makes the morphogenetic space harder to explore? Unlike the plant morphospace, perhaps evolution has not been able to explore the bilaterian morphospace to the same degree.
In terms of the time onset of the radiation, did events like the Snowball Earths play a crucial role or not? Why did the radiation not occur 100 million years earlier, or 100 million years later?

In terms of environmental change, what is the waiting time between the first appearance of an environment conducive to the evolution of large animals and the roughening of the animal fitness landscape? Is it substantial (e.g., controlled by the waiting time for the right morphologies to evolve to begin Phanerozoic-style ecological interactions, including predation)? Or is it that there is some key environmental control that once removed leads quite quickly to the radiation.

In summation, explaining the Cambrian “explosion” of bilaterian animals will remain a rich field of enquiry for quite some time to come!

ACKNOWLEDGEMENTS

I thank Sam Bowring, Nigel Hughes, Andy Knoll, Adam Maloof, and an anonymous reviewer for very helpful comments on the manuscript, Samar Bush for help with Figure 3, as well as Jennifer Jongsma and Eliza Jewitt for help with the final manuscript.

LITERATURE CITED


www.annualreviews.org • Explaining the Cambrian “Explosion” 379


Beautifully illustrated and comprehensive survey of the Chengjiang taxa.


A key book by a leader working at the interface of evolution and development.
Impressive multidisciplinary volume, with a paleontological emphasis.

Wright S. 1931. Evolution in Mendelian populations. *Genetics* 16:97–159

Authoritative compilation on diverse aspects of the Cambrian “explosion.”
Contents

Threads: A Life in Geochemistry
Karl K. Turekian ................................................................. 1

Reflections on the Conception, Birth, and Childhood of Numerical Weather Prediction
Edward N. Lorenz ............................................................... 37

Binary Minor Planets
Derek C. Richardson and Kevin J. Walsh .................................. 47

Mössbauer Spectroscopy of Earth and Planetary Materials
M. Darby Dyar, David G. Agresti, Martha W. Schaefer, Christopher A. Grant, and Elizabeth C. Sklute .......................................................... 83

Phanerozoic Biodiversity Mass Extinctions
Richard K. Bambach ............................................................. 127

The Yarkovsky and YORP Effects: Implications for Asteroid Dynamics
William F. Bottke, Jr., David Vokrouhlický, David P. Rubincam, and David Nesvorný ........................................................... 157

Planetesimals to Brown Dwarfs: What is a Planet?
Gibor Basri and Michael E. Brown ......................................... 193

History and Applications of Mass-Independent Isotope Effects
Mark H. Thiemens ............................................................... 217

Seismic Triggering of Eruptions in the Far Field: Volcanoes and Geysers
Michael Manga and Emily Brodsky ......................................... 263

Dynamics of Lake Eruptions and Possible Ocean Eruptions
Youxue Zhang and George W. Kling ........................................ 293

Bed Material Transport and the Morphology of Alluvial River Channels
Michael Church ................................................................. 325

Explaining the Cambrian “Explosion” of Animals
Charles R. Marshall ............................................................ 355
Cosmic Dust Collection in Aerogel
Mark J. Burchell, Giles Graham, and Anton Kearsley .................................................. 385

Using Thermochronology to Understand Orogenic Erosion
Peter W. Reiners and Mark T. Brandon ................................................................. 419

High-Mg Andesites in the Setouchi Volcanic Belt, Southwestern Japan:
Analogy to Archean Magmatism and Continental Crust Formation?
Yoshiyuki Tatsumi .................................................................................................. 467

Hydrogen Isotopic (D/H) Composition of Organic Matter During
Diagenesis and Thermal Maturation
Arndt Schimmelmann, Alex L. Sessions, and Maria Mastalerz .......................... 501

The Importance of Secondary Cratering to Age Constraints on
Planetary Surfaces
Alfred S. McEwen and Edward B. Bierhaus ......................................................... 535

Dates and Rates: Temporal Resolution in the Deep Time Stratigraphic
Record
Douglas H. Erwin .................................................................................................. 569

Evidence for Aseismic Deformation Rate Changes Prior to
Earthquakes
Evelyn A. Roeloffs .................................................................................................. 591

Water, Melting, and the Deep Earth H₂O Cycle
Marc M. Hirschmann .............................................................................................. 629

The General Circulation of the Atmosphere
Tapio Schneider .................................................................................................. 655

INDEXES

Subject Index ........................................................................................................... 689
Cumulative Index of Contributing Authors, Volumes 24–34 ......................... 707
Cumulative Index of Chapter Titles, Volumes 24–34 ........................................ 710

ERRATA

An online log of corrections to Annual Review of Earth and Planetary Sciences
chapters may be found at http://earth.annualreviews.org