The origin of biological information and the higher taxonomic categories

Stephen C. Meyer

Palm Beach Atlantic University, 901 S. Flagler Dr., West Palm Beach, Florida 33401
e-mail: stevemeyer@discovery.org

Introduction

In a recent volume of the Vienna Series in Theoretical Biology (2003), Gerd B. Müller and Stuart Newman argue that what they call the "origination of organismal form" remains an unsolved problem. In making this claim, Müller and Newman (2003:3–10) distinguish two distinct issues, namely, (1) the causes of form generation in the individual organism during embryological development and (2) the causes responsible for the production of novel organismal forms in the first place during the history of life. To distinguish the latter case (phylogeny) from the former (ontogeny), Müller and Newman use the term "origination" to designate the causal processes by which biological form first arose during the evolution of life. They insist that "the molecular mechanisms that bring about biological form in modern day embryos should not be confused" with the causes responsible for the origin (or "origination") of novel biological forms during the history of life (p. 3). They further argue that we know more about the causes of ontogenesis, due to advances in molecular biology, molecular genetics and developmental biology, than we do about the causes of phylogenesis—the ultimate origination of new biological forms during the remote past.

In making this claim, Müller and Newman are careful to affirm that evolutionary biology has succeeded in explaining how pre-existing forms diversify under the twin influences of natural selection and variation of genetic traits. Sophisticated mathematically-based models of population genetics have proven adequate for mapping and understanding quantitative variability and populational changes in organisms. Yet Müller and Newman insist that population genetics, and thus evolutionary biology, has not identified a specifically causal explanation for the origin of true morphological novelty during the history of life. Central to their concern is what they see as the inadequacy of the variation of genetic traits as a source of new form and structure. They note, following Darwin himself, that the sources of new form and structure must precede the action of natural selection (2003: 3)—that selection must act on what already exists. Yet, in their view, the "genocentricity" and "incrementalism" of the neo-Darwinian mechanism has meant that an adequate source of new form and structure has yet to be identified by theoretical biologists. Instead, Müller and Newman see the need to identify epigenetic sources of morphological innovation during the evolution of life. In the meantime, however, they insist neo-Darwinism lacks any "theory of the generative" (p. 7).

As it happens, Müller and Newman are not alone in this judgment. In the last decade or so a host of scientific essays and books have questioned the efficacy of selection and mutation as a mechanism for generating morphological novelty, as even a brief literature survey will establish. Thomson (1992:107) expressed doubt that large-scale morphological changes could accumulate via minor phenotypic changes at the population genetic level. Miklos (1993:29) argued that neo-Darwinism fails to provide a mechanism that can produce

large-scale innovations in form and complexity. Gilbert et al. (1996) attempted to develop a new theory of evolutionary mechanisms to supplement classical neo-Darwinism, which, they argued, could not adequately explain macroevolution. As they put it in a memorable summary of the situation: "starting in the 1970s, many biologists began questioning its [neo-Darwinism's] adequacy in explaining evolution. Genetics might be adequate for explaining microevolution, but microevolutionary changes in gene frequency were not seen as able to turn a reptile into a mammal or to convert a fish into an amphibian. Microevolution looks at adaptations that concern the survival of the fittest, not the arrival of the fittest. As Goodwin (1995) points out, 'the origin of species—Darwin's problem—remains unsolved' " (p. 361). Though Gilbert et al. (1996) attempted to solve the problem of the origin of form by proposing a greater role for developmental genetics within an otherwise neo-Darwinian framework,[1] numerous recent authors have continued to raise questions about the adequacy of that framework itself or about the problem of the origination of form generally (Webster & Goodwin 1996; Shubin & Marshall 2000; Erwin 2000; Conway Morris 2000, 2003b; Carrol 2000; Wagner 2001; Becker & Lönnig 2001; Stadler et al. 2001; Lönnig & Saedler 2002; Wagner & Stadler 2003; Valentine 2004:189–194).

What lies behind this skepticism? Is it warranted? Is a new and specifically causal theory needed to explain the origination of biological form?

This review will address these questions. It will do so by analyzing the problem of the origination of organismal form (and the

[1] Specifically, Gilbert et al. (1996) argued that changes in morphogenetic fields might produce large-scale changes in the developmental programs and, ultimately, body plans of organisms. Yet they offered no evidence that such fields—if indeed they exist—can be altered to produce advantageous variations in body plan, though this is a necessary condition of any successful causal theory of macroevolution.

corresponding emergence of higher taxa) from a particular theoretical standpoint. Specifically, it will treat the problem of the origination of the higher taxonomic groups as a manifestation of a deeper problem, namely, the problem of the origin of the information (whether genetic or epigenetic) that, as it will be argued, is necessary to generate morphological novelty.

In order to perform this analysis, and to make it relevant and tractable to systematists and paleontologists, this paper will examine a paradigmatic example of the origin of biological form and information during the history of life: the Cambrian explosion. During the Cambrian, many novel animal forms and body plans (representing new phyla, sub-phyla and classes) arose in a geologically brief period of time. The following information-based analysis of the Cambrian explosion will support the claim of recent authors such as Müller and Newman that the mechanism of selection and genetic mutation does not constitute an adequate causal explanation of the origination of biological form in the higher taxonomic groups. It will also suggest the need to explore other possible causal factors for the origin of form and information during the evolution of life and will examine some other possibilities that have been proposed.

The Cambrian Explosion

The "Cambrian explosion" refers to the geologically sudden appearance of many new animal body plans about 530 million years ago. At this time, at least nineteen, and perhaps as many as thirty-five phyla of forty total (Meyer et al. 2003), made their first appearance on Earth within a narrow five- to ten-million-year window of geologic time (Bowring et al. 1993, 1998a:1, 1998b:40; Kerr 1993; Monastersky 1993; Aris-Brosou & Yang 2003). Many new sub-phyla, between 32 and 48 of 56 total (Meyer et al. 2003), and classes of animals also arose at this time with representatives of these new higher taxa manifesting signifi-

cant morphological innovations. The Cambrian explosion thus marked a major episode of morphogenesis in which many new and disparate organismal forms arose in a geologically brief period of time.

To say that the fauna of the Cambrian period appeared in a geologically sudden manner also implies the absence of clear transitional intermediate forms connecting Cambrian animals with simpler pre-Cambrian forms. And, indeed, in almost all cases, the Cambrian animals have no clear morphological antecedents in earlier Vendian or Precambrian fauna (Miklos 1993, Erwin et al. 1997:132, Steiner & Reitner 2001, Conway Morris 2003b:510, Valentine et al. 2003:519–520). Further, several recent discoveries and analyses suggest that these morphological gaps may not be merely an artifact of incomplete sampling of the fossil record (Foote 1997, Foote et al. 1999, Benton & Ayala 2003, Meyer et al. 2003), suggesting that the fossil record is at least approximately reliable (Conway Morris 2003b:505).

As a result, debate now exists about the extent to which this pattern of evidence comports with a strictly monophyletic view of evolution (Conway Morris 1998a, 2003a, 2003b:510; Willmer 1990, 2003). Further, among those who accept a monophyletic view of the history of life, debate exists about whether to privilege fossil or molecular data and analyses. Those who think the fossil data provide a more reliable picture of the origin of the Metazoan tend to think these animals arose relatively quickly—that the Cambrian explosion had a "short fuse." (Conway Morris 2003b:505–506, Valentine & Jablonski 2003). Some (Wray et al. 1996), but not all (Ayala et al. 1998), who think that molecular phylogenies establish reliable divergence times from pre-Cambrian ancestors think that the Cambrian animals evolved over a very long period of time—that the Cambrian explosion had a "long fuse." This review will not address these questions of historical pattern. Instead, it will analyze whether the neo-Darwinian process of mu-

tation and selection, or other processes of evolutionary change, can generate the form and information necessary to produce the animals that arise in the Cambrian. This analysis will, for the most part,[2] therefore, not depend upon assumptions of either a long or short fuse for the Cambrian explosion, or upon a monophyletic or polyphyletic view of the early history of life.

Defining Biological Form and Information

Form, like life itself, is easy to recognize but often hard to define precisely. Yet, a reasonable working definition of form will suffice for our present purposes. Form can be defined as the four-dimensional topological relations of anatomical parts. This means that one can understand form as a unified arrangement of body parts or material components in a distinct shape or pattern (topology)—one that exists in three spatial dimensions and which arises in time during ontogeny.

Insofar as any particular biological form constitutes something like a distinct arrangement of constituent body parts, form can be seen as arising from constraints that limit the possible arrangements of matter. Specifically, organismal form arises (both in phylogeny and ontogeny) as possible arrangements of material parts are constrained to establish a specific or particular arrangement with an identifiable three dimensional topography—one that we would recognize as a particular protein, cell type, organ, body plan or organism. A particular

---

[2] If one takes the fossil record at face value and assumes that the Cambrian explosion took place within a relatively narrow 5–10 million year window, explaining the origin of the information necessary to produce new proteins, for example, becomes more acute in part because mutation rates would not have been sufficient to generate the number of changes in the genome necessary to build the new proteins for more complex Cambrian animals (Ohno 1996:8475–8478). This review will argue that, even if one allows several hundred million years for the origin of the metazoan, significant probabilistic and other difficulties remain with the neo-Darwinian explanation of the origin of form and information.

"form," therefore, represents a highly specific and constrained arrangement of material components (among a much larger set of possible arrangements).

Understanding form in this way suggests a connection to the notion of information in its most theoretically general sense. When Shannon (1948) first developed a mathematical theory of information he equated the amount of information transmitted with the amount of uncertainty reduced or eliminated in a series of symbols or characters. Information, in Shannon's theory, is thus imparted as some options are excluded and others are actualized. The greater the number of options excluded, the greater the amount of information conveyed. Further, constraining a set of possible material arrangements by whatever process or means involves excluding some options and actualizing others. Thus, to constrain a set of possible material states is to generate information in Shannon's sense. It follows that the constraints that produce biological form also impart information. Or conversely, one might say that producing organismal form by definition requires the generation of information.

In classical Shannon information theory, the amount of information in a system is also inversely related to the probability of the arrangement of constituents in a system or the characters along a communication channel (Shannon 1948). The more improbable (or complex) the arrangement, the more Shannon information, or information-carrying capacity, a string or system possesses.

Since the 1960s, mathematical biologists have realized that Shannon's theory could be applied to the analysis of DNA and proteins to measure the information-carrying capacity of these macromolecules. Since DNA contains the assembly instructions for building proteins, the information-processing system in the cell represents a kind of communication channel (Yockey 1992: 110). Further, DNA conveys information via specifically arranged sequences of nu-

cleotide bases. Since each of the four bases has a roughly equal chance of occurring at each site along the spine of the DNA molecule, biologists can calculate the probability, and thus the information-carrying capacity, of any particular sequence **n** bases long.

The ease with which information theory applies to molecular biology has created confusion about the type of information that DNA and proteins possess. Sequences of nucleotide bases in DNA, or amino acids in a protein, are highly improbable and thus have large information-carrying capacities. But, like meaningful sentences or lines of computer code, genes and proteins are also *specified* with respect to function. Just as the meaning of a sentence depends upon the specific arrangement of the letters in a sentence, so too does the function of a gene sequence depend upon the specific arrangement of the nucleotide bases in a gene. Thus, molecular biologists beginning with Crick equated *information* not only with complexity but also with "specificity," where "specificity" or "specified" has meant "necessary to function" (Crick 1958:144, 153; Sarkar, 1996:191).[3] Molecular biologists such as Monod and Crick understood biological information—the information stored in DNA and proteins—as something more than mere complexity (or improbability). Their notion of information associated both biochemical contingency and combinatorial complexity with DNA sequences (allowing DNA's carrying capacity to be calculated), but it also affirmed that sequences of nucleotides and amino acids in functioning macromolecules possessed a high degree of *specificity* relative to the maintenance of cellular function.

The ease with which information theory applies to molecular biology has also created confusion about the location of infor-

---

[3] As Crick put it, "information means here the *precise* determination of sequence, either of bases in the nucleic acid or on amino acid residues in the protein" (Crick 1958:144, 153).

mation in organisms. Perhaps because the information carrying capacity of the gene could be so easily measured, it has been easy to treat DNA, RNA and proteins as the sole repositories of biological information. Neo-Darwinists in particular have assumed that the origination of biological form could be explained by recourse to processes of genetic variation and mutation alone (Levinton 1988:485). Yet if one understands organismal form as resulting from constraints on the possible arrangements of matter at many levels in the biological hierarchy— from genes and proteins to cell types and tissues to organs and body plans—then clearly biological organisms exhibit many levels of information-rich structure.

Thus, we can pose a question, not only about the origin of genetic information, but also about the origin of the information necessary to generate form and structure at levels higher than that present in individual proteins. We must also ask about the origin of the "specified complexity," as opposed to mere complexity, that characterizes the new genes, proteins, cell types and body plans that arose in the Cambrian explosion. Dembski (2002) has used the term "complex specified information" (CSI) as a synonym for "specified complexity" to help distinguish functional biological information from mere Shannon information—that is, specified complexity from mere complexity. This review will use this term as well.

The Cambrian Information Explosion

The Cambrian explosion represents a remarkable jump in the specified complexity or "complex specified information" (CSI) of the biological world. For over three billion years, the biological realm included little more than bacteria and algae (Brocks et al. 1999). Then, beginning about 570–565 million years ago (mya), the first complex multicellular organisms appeared in the rock strata, including sponges, cnidarians, and the peculiar Ediacaran biota (Grotzin-

ger et al. 1995). Forty million years later, the Cambrian explosion occurred (Bowring et al. 1993). The emergence of the Ediacaran biota (570 mya), and then to a much greater extent the Cambrian explosion (530 mya), represented steep climbs up the biological complexity gradient.

One way to estimate the amount of new CSI that appeared with the Cambrian animals is to count the number of new cell types that emerged with them (Valentine 1995:91–93). Studies of modern animals suggest that the sponges that appeared in the late Precambrian, for example, would have required five cell types, whereas the more complex animals that appeared in the Cambrian (e.g., arthropods) would have required fifty or more cell types. Functionally more complex animals require more cell types to perform their more diverse functions. New cell types require many new and specialized proteins. New proteins, in turn, require new genetic information. Thus an increase in the number of cell types implies (at a minimum) a considerable increase in the amount of specified genetic information. Molecular biologists have recently estimated that a minimally complex single-celled organism would require between 318 and 562 kilobase pairs of DNA to produce the proteins necessary to maintain life (Koonin 2000). More complex single cells might require upward of a million base pairs. Yet to build the proteins necessary to sustain a complex arthropod such as a trilobite would require orders of magnitude more coding instructions. The genome size of a modern arthropod, the fruitfly *Drosophila melanogaster,* is approximately 180 million base pairs (Gerhart & Kirschner 1997:121, Adams et al. 2000). Transitions from a single cell to colonies of cells to complex animals represent significant (and, in principle, measurable) increases in CSI.

Building a new animal from a single-celled organism requires a vast amount of new genetic information. It also requires a way of arranging gene products—proteins—into higher levels of organization.

New proteins are required to service new cell types. But new proteins must be organized into new systems within the cell; new cell types must be organized into new tissues, organs, and body parts. These, in turn, must be organized to form body plans. New animals, therefore, embody hierarchically organized systems of lower-level parts within a functional whole. Such hierarchical organization itself represents a type of information, since body plans comprise both highly improbable and functionally specified arrangements of lower-level parts. The specified complexity of new body plans requires explanation in any account of the Cambrian explosion.

Can neo-Darwinism explain the discontinuous increase in CSI that appears in the Cambrian explosion—either in the form of new genetic information or in the form of hierarchically organized systems of parts? We will now examine the two parts of this question.

Novel Genes and Proteins

Many scientists and mathematicians have questioned the ability of mutation and selection to generate information in the form of novel genes and proteins. Such skepticism often derives from consideration of the extreme improbability (and specificity) of functional genes and proteins.

A typical gene contains over one thousand precisely arranged bases. For any specific arrangement of four nucleotide bases of length n, there is a corresponding number of possible arrangements of bases, 4^n. For any protein, there are 20^n possible arrangements of protein-forming amino acids. A gene 999 bases in length represents one of 4^{999} possible nucleotide sequences; a protein of 333 amino acids is one of 20^{333} possibilities.

Since the 1960s, some biologists have thought functional proteins to be rare among the set of possible amino acid sequences. Some have used an analogy with human language to illustrate why this should be the case. Denton (1986, 309–311), for example,

has shown that meaningful words and sentences are extremely rare among the set of possible combinations of English letters, especially as sequence length grows. (The ratio of meaningful 12-letter words to 12-letter sequences is $1/10^{14}$; the ratio of 100-letter sentences to possible 100-letter strings is $1/10^{100}$.) Further, Denton shows that most meaningful sentences are *highly isolated* from one another in the space of possible combinations, so that random substitutions of letters will, after a very few changes, inevitably degrade meaning. Apart from a few closely clustered sentences accessible by random substitution, the overwhelming majority of meaningful sentences lie, probabilistically speaking, beyond the reach of random search.

Denton (1986:301–324) and others have argued that similar constraints apply to genes and proteins. They have questioned whether an undirected search via mutation and selection would have a reasonable chance of locating new islands of function—representing fundamentally new genes or proteins—within the time available (Eden 1967, Shützenberger 1967, Løvtrup 1979). Some have also argued that alterations in sequencing would likely result in loss of protein function before fundamentally new function could arise (Eden 1967, Denton 1986). Nevertheless, neither the extent to which genes and proteins are sensitive to functional loss as a result of sequence change, nor the extent to which functional proteins are isolated within sequence space, has been fully known.

Recently, experiments in molecular biology have shed light on these questions. A variety of mutagenesis techniques have shown that proteins (and thus the genes that produce them) are indeed highly specified relative to biological function (Bowie & Sauer 1989, Reidhaar-Olson & Sauer 1990, Taylor et al. 2001). Mutagenesis research tests the sensitivity of proteins (and, by implication, DNA) to functional loss as a result of alterations in sequencing. Studies of proteins have long shown that amino acid res-

idues at many active positions cannot vary without functional loss (Perutz & Lehmann 1968). More recent protein studies (often using mutagenesis experiments) have shown that functional requirements place significant constraints on sequencing even at non-active site positions (Bowie & Sauer 1989, Reidhaar-Olson & Sauer 1990, Chothia et al. 1998, Axe 2000, Taylor et al. 2001). In particular, Axe (2000) has shown that multiple as opposed to single position amino acid substitutions inevitably result in loss of protein function, even when these changes occur at sites that allow variation when altered in isolation. Cumulatively, these constraints imply that proteins are highly sensitive to functional loss as a result of alterations in sequencing, and that functional proteins represent highly isolated and improbable arrangements of amino acids—arrangements that are far more improbable, in fact, than would be likely to arise by chance alone in the time available (Reidhaar-Olson & Sauer 1990; Behe 1992; Kauffman 1995:44; Dembski 1998:175–223; Axe 2000, 2004). (See below the discussion of the neutral theory of evolution for a precise quantitative assessment.)

Of course, neo-Darwinists do not envision a completely random search through the set of all possible nucleotide sequences—so-called "sequence space." They envision natural selection acting to preserve small advantageous variations in genetic sequences and their corresponding protein products. Dawkins (1996), for example, likens an organism to a high mountain peak. He compares climbing the sheer precipice up the front side of the mountain to building a new organism by chance. He acknowledges that this approach up "Mount Improbable" will not succeed. Nevertheless, he suggests that there is a gradual slope up the backside of the mountain that could be climbed in small incremental steps. In his analogy, the backside climb up "Mount Improbable" corresponds to the process of natural selection acting on random changes in the genetic text. What

chance alone cannot accomplish blindly or in one leap, selection (acting on mutations) can accomplish through the cumulative effect of many slight successive steps.

Yet the extreme specificity and complexity of proteins presents a difficulty, not only for the chance origin of specified biological information (i.e., for random mutations acting alone), but also for selection and mutation acting in concert. Indeed, mutagenesis experiments cast doubt on each of the two scenarios by which neo-Darwinists envision new information arising from the mutation/selection mechanism (for review, see Lönnig 2001). For neo-Darwinism, new functional genes either arise from non-coding sections in the genome or from preexisting genes. Both scenarios are problematic.

In the first scenario, neo-Darwinists envision new genetic information arising from those sections of the genetic text that can presumably vary freely without consequence to the organism. According to this scenario, non-coding sections of the genome, or duplicated sections of coding regions, can experience a protracted period of "neutral evolution" (Kimura 1983) during which alterations in nucleotide sequences have no discernible effect on the function of the organism. Eventually, however, a new gene sequence will arise that can code for a novel protein. At that point, natural selection can favor the new gene and its functional protein product, thus securing the preservation and heritability of both.

This scenario has the advantage of allowing the genome to vary through many generations, as mutations "search" the space of possible base sequences. The scenario has an overriding problem, however: the size of the combinatorial space (i.e., the number of possible amino acid sequences) and the extreme rarity and isolation of the functional sequences within that space of possibilities. Since natural selection can do nothing to help *generate* new functional sequences, but rather can only preserve such sequences once they have arisen, chance alone—random variation—must do the

work of information generation—that is, of finding the exceedingly rare functional sequences within the set of combinatorial possibilities. Yet the probability of randomly assembling (or "finding," in the previous sense) a functional sequence is extremely small.

Cassette mutagenesis experiments performed during the early 1990s suggest that the probability of attaining (at random) the correct sequencing for a short protein 100 amino acids long is about 1 in 10^{65} (Reidhaar-Olson & Sauer 1990, Behe 1992:65–69). This result agreed closely with earlier calculations that Yockey (1978) had performed based upon the known sequence variability of cytochrome c in different species and other theoretical considerations. More recent mutagenesis research has provided additional support for the conclusion that functional proteins are exceedingly rare among possible amino acid sequences (Axe 2000, 2004). Axe (2004) has performed site directed mutagenesis experiments on a 150-residue protein-folding domain within a β-lactamase enzyme. His experimental method improves upon earlier mutagenesis techniques and corrects for several sources of possible estimation error inherent in them. On the basis of these experiments, Axe has estimated the ratio of (a) proteins of typical size (150 residues) that perform a specified function via any folded structure to (b) the whole set of possible amino acids sequences of that size. Based on his experiments, Axe has estimated this ratio to be 1 to 10^{77}. Thus, the probability of finding a functional protein among the possible amino acid sequences corresponding to a 150-residue protein is similarly 1 in 10^{77}.

Other considerations imply additional improbabilities. First, new Cambrian animals would require proteins much longer than 100 residues to perform many necessary specialized functions. Ohno (1996) has noted that Cambrian animals would have required complex proteins such as lysyl oxidase in order to support their stout body structures. Lysyl oxidase molecules in ex-

tant organisms comprise over 400 amino acids. These molecules are both highly complex (non-repetitive) and functionally specified. Reasonable extrapolation from mutagenesis experiments done on shorter protein molecules suggests that the probability of producing functionally sequenced proteins of this length at random is so small as to make appeals to chance absurd, even granting the duration of the entire universe. (See Dembski 1998:175–223 for a rigorous calculation of this "Universal Probability Bound"; See also Axe 2004.) Yet, second, fossil data (Bowring et al. 1993, 1998a:1, 1998b:40; Kerr 1993; Monastersky 1993), and even molecular analyses supporting deep divergence (Wray et al. 1996), suggest that the duration of the Cambrian explosion (between $5–10 \times 10^6$ and, at most, 7×10^7 years) is far smaller than that of the entire universe ($1.3–2 \times 10^{10}$ years). Third, DNA mutation rates are far too low to generate the novel genes and proteins necessary to building the Cambrian animals, given the most probable duration of the explosion as determined by fossil studies (Conway Morris 1998b). As Ohno (1996:8475) notes, even a mutation rate of 10^{-9} per base pair per year results in only a 1% change in the sequence of a given section of DNA in 10 million years. Thus, he argues that mutational divergence of pre-existing genes cannot explain the origin of the Cambrian forms in that time.[4]

---

[4] To solve this problem Ohno himself proposes the existence of a hypothetical ancestral form that possessed virtually all the genetic information necessary to produce the new body plans of the Cambrian animals. He asserts that this ancestor and its "pananimalian genome" might have arisen several hundred million years before the Cambrian explosion. On this view, each of the different Cambrian animals would have possessed virtually identical genomes, albeit with considerable latent and unexpressed capacity in the case of each individual form (Ohno 1996:8475–8478). While this proposal might help explain the origin of the Cambrian animal forms by reference to pre-existing genetic information, it does not solve, but instead merely displaces, the problem of the origin of the genetic information necessary to produce these new forms.

The selection/mutation mechanism faces another probabilistic obstacle. The animals that arise in the Cambrian exhibit structures that would have required many new *types* of cells, each of which would have required many novel proteins to perform their specialized functions. Further, new cell types require *systems* of proteins that must, as a condition of functioning, act in close coordination with one another. The unit of selection in such systems ascends to the system as a whole. Natural selection selects for functional advantage. But new cell types require whole systems of proteins to perform their distinctive functions. In such cases, natural selection cannot contribute to the process of information generation until *after* the information necessary to build the requisite *system* of proteins has arisen. Thus random variations must, again, do the work of information generation—and now not simply for one protein, but for many proteins arising at nearly the same time. Yet the odds of this occurring by chance alone are, of course, far smaller than the odds of the chance origin of a single gene or protein—so small in fact as to render the chance origin of the genetic information necessary to build a new cell type (a necessary but not sufficient condition of building a new body plan) problematic given even the most optimistic estimates for the duration of the Cambrian explosion.

Dawkins (1986:139) has noted that scientific theories can rely on only so much "luck" before they cease to be credible. The neutral theory of evolution, which, by its own logic, prevents natural selection from playing a role in generating genetic information until after the fact, relies on entirely too much luck. The sensitivity of proteins to functional loss, the need for long proteins to build new cell types and animals, the need for whole new *systems* of proteins to service new cell types, the probable brevity of the Cambrian explosion relative to mutation rates—all suggest the immense improbability (and implausi-

bility) of any scenario for the origination of Cambrian genetic information that relies upon random variation alone unassisted by natural selection.

Yet the neutral theory requires novel genes and proteins to arise—essentially— by random mutation alone. Adaptive advantage accrues *after* the generation of new functional genes and proteins. Thus, natural selection cannot play a role *until* new information-bearing molecules have independently arisen. Thus neutral theorists envision the need to scale the steep face of a Dawkins-style precipice of which there is *no* gradually sloping backside—a situation that, by Dawkins' own logic, is probabilistically untenable.

In the second scenario, neo-Darwinists envision novel genes and proteins arising by numerous successive mutations in the preexisting genetic text that codes for proteins. To adapt Dawkins's metaphor, this scenario envisions gradually climbing down one functional peak and then ascending another. Yet mutagenesis experiments again suggest a difficulty. Recent experiments show that, even when exploring a region of sequence space populated by proteins of a single fold and function, most multiple-position changes quickly lead to loss of function (Axe 2000). Yet to turn one protein into another with a completely novel structure and function requires specified changes at many sites. Indeed, the number of changes necessary to produce a new protein greatly exceeds the number of changes that will typically produce functional losses. Given this, the probability of escaping total functional loss during a random search for the changes needed to produce a new function is extremely small—and this probability diminishes exponentially with each additional requisite change (Axe 2000). Thus, Axe's results imply that, in all probability, random searches for novel proteins (through sequence space) will result in functional loss long before any novel functional protein will emerge.

Blanco et al. have come to a similar conclusion. Using directed mutagenesis, they have determined that residues in both the hydrophobic core and on the surface of the protein play essential roles in determining protein structure. By sampling intermediate sequences between two naturally occurring sequences that adopt different folds, they found that the intermediate sequences "lack a well defined three-dimensional structure." Thus, they conclude that it is unlikely that a new protein fold would evolve from a pre-existing fold via a series of folded intermediates sequences (Blanco et al. 1999: 741).

Thus, although this second neo-Darwinian scenario has the advantage of starting with functional genes and proteins, it also has a lethal disadvantage: any process of random mutation or rearrangement in the genome would in all probability generate nonfunctional intermediate sequences before fundamentally new functional genes or proteins would arise. Clearly, nonfunctional intermediate sequences confer no survival advantage on their host organisms. Natural selection favors *only* functional advantage. It cannot select or favor nucleotide sequences or polypeptide chains that do not yet perform biological functions, and still less will it favor sequences that efface or destroy preexisting function.

Evolving genes and proteins will range through a series of nonfunctional intermediate sequences that natural selection will not favor or preserve but will, in all probability, eliminate (Blanco et al. 1999, Axe 2000). When this happens, selection-driven evolution will cease. At this point, neutral evolution of the genome (unhinged from selective pressure) may ensue, but, as we have seen, such a process must overcome immense probabilistic hurdles, even granting cosmic time.

Thus, whether one envisions the evolutionary process beginning with a noncoding region of the genome or a preexisting functional gene, the functional specificity and complexity of proteins impose very stringent limitations on the efficacy of mutation and selection. In the first case, function must arise first, before natural selection can act to favor a novel variation. In the second case, function must be continuously maintained in order to prevent deleterious (or lethal) consequences to the organism and to allow further evolution. Yet the complexity and functional specificity of proteins implies that both these conditions will be extremely difficult to meet. Therefore, the neo-Darwinian mechanism appears to be inadequate to generate the new information present in the novel genes and proteins that arise with the Cambrian animals.

Novel Body Plans

The problems with the neo-Darwinian mechanism run deeper still. In order to explain the origin of the Cambrian animals, one must account not only for new proteins and cell types, but also for the origin of new body plans. Within the past decade, developmental biology has dramatically advanced our understanding of how body plans are built during ontogeny. In the process, it has also uncovered a profound difficulty for neo-Darwinism.

Significant morphological change in organisms requires attention to timing. Mutations in genes that are expressed late in the development of an organism will not affect the body plan. Mutations expressed early in development, however, could conceivably produce significant morphological change (Arthur 1997:21). Thus, events expressed early in the development of organisms have the only realistic chance of producing large-scale macroevolutionary change (Thomson 1992). As John and Miklos (1988:309) explain, macroevolutionary change requires alterations in the very early stages of ontogenesis.

Yet recent studies in developmental biology make clear that mutations expressed early in development typically have deleterious effects (Arthur 1997:21). For example, when early-acting body plan mol-

ecules, or morphogens such as *bicoid* (which helps to set up the anterior-posterior head-to-tail axis in *Drosophila*), are perturbed, development shuts down (Nüsslein-Volhard & Wieschaus 1980, Lawrence & Struhl 1996, Müller & Newman 2003).[5] The resulting embryos die. Moreover, there is a good reason for this. If an engineer modifies the length of the piston rods in an internal combustion engine without modifying the crankshaft accordingly, the engine won't start. Similarly, processes of development are tightly integrated spatially and temporally such that changes early in development will require a host of other coordinated changes in separate but functionally interrelated developmental processes downstream. For this reason, mutations will be much more likely to be deadly if they disrupt a functionally deeply-embedded structure such as a spinal column than if they affect more isolated anatomical features such as fingers (Kauffman 1995:200).

This problem has led to what McDonald (1983) has called "a great Darwinian paradox" (p. 93). McDonald notes that genes that are observed to vary within natural populations do not lead to major adaptive changes, while genes that could cause major changes—the very stuff of macroevolution—apparently do not vary. In other words, mutations of the kind that macroevolution doesn't need (namely, viable genetic mutations in DNA expressed late in development) do occur, but those that it does need (namely, beneficial body plan mutations expressed early in development)

apparently don't occur.[6] According to Darwin (1859:108) natural selection cannot act until favorable variations arise in a population. Yet there is no evidence from developmental genetics that the kind of variations required by neo-Darwinism—namely, favorable body plan mutations—ever occur.

Developmental biology has raised another formidable problem for the mutation/selection mechanism. Embryological evidence has long shown that DNA does not wholly determine morphological form (Goodwin 1985, Nijhout 1990, Sapp 1987, Müller & Newman 2003), suggesting that mutations in DNA alone cannot account for the morphological changes required to build a new body plan.

DNA helps directs protein synthesis.[7] It also helps to regulate the timing and expression of the synthesis of various proteins within cells. Yet, DNA alone does not determine how individual proteins assemble themselves into larger systems of proteins; still less does it solely determine how cell types, tissue types, and organs arrange themselves into body plans (Harold 1995:

---

[5] Some have suggested that mutations in "master regulator" Hox genes might provide the raw material for body plan morphogenesis. Yet there are two problems with this proposal. First, Hox gene expression begins only after the foundation of the body plan has been established in early embryogenesis (Davidson 2001:66). Second, Hox genes are highly conserved across many disparate phyla and so cannot account for the morphological differences that exist between the phyla (Valentine 2004:88).

[6] Notable differences in the developmental pathways of similar organisms have been observed. For example, congeneric species of sea urchins (from genus *Heliocidaris*) exhibit striking differences in their developmental pathways (Raff 1999:110–121). Thus, it might be argued that such differences show that early developmental programs can in fact be mutated to produce new forms. Nevertheless, there are two problems with this claim. First, there is no direct evidence that existing differences in sea urchin development arose by mutation. Second, the observed differences in the developmental programs of different species of sea urchins do not result in new body plans, but instead in highly conserved structures. Despite differences in developmental patterns, the endpoints are the same. Thus, even if it can be assumed that mutations produced the differences in developmental pathways, it must be acknowledged that such changes did not result in novel form.

[7] Of course, many post-translation processes of modification also play a role in producing a functional protein. Such processes make it impossible to predict a protein's final sequencing from its corresponding gene sequence alone (Sarkar 1996:199–202).

2774, Moss 2004). Instead, other factors—such as the three-dimensional structure and organization of the cell membrane and cytoskeleton and the spatial architecture of the fertilized egg—play important roles in determining body plan formation during embryogenesis.

For example, the structure and location of the cytoskeleton influence the patterning of embryos. Arrays of microtubules help to distribute the essential proteins used during development to their correct locations in the cell. Of course, microtubules themselves are made of many protein subunits. Nevertheless, like bricks that can be used to assemble many different structures, the tubulin subunits in the cell's microtubules are identical to one another. Thus, neither the tubulin subunits nor the genes that produce them account for the different shape of microtubule arrays that distinguish different kinds of embryos and developmental pathways. Instead, the structure of the microtubule array itself is determined by the location and arrangement of its subunits, not the properties of the subunits themselves. For this reason, it is not possible to predict the structure of the cytoskeleton of the cell from the characteristics of the protein constituents that form that structure (Harold 2001:125).

Two analogies may help further clarify the point. At a building site, builders will make use of many materials: lumber, wires, nails, drywall, piping, and windows. Yet building materials do not determine the floor plan of the house, or the arrangement of houses in a neighborhood. Similarly, electronic circuits are composed of many components, such as resistors, capacitors, and transistors. But such lower-level components do not determine their own arrangement in an integrated circuit. Biological systems also depend on hierarchical arrangements of parts. Genes and proteins are made from simple building blocks—nucleotide bases and amino acids—arranged in specific ways. Cell types are made of, among other things, systems of specialized proteins. Organs are made of specialized arrangements of cell types and tissues. And body plans comprise specific arrangements of specialized organs. Yet, clearly, the properties of individual proteins (or, indeed, the lower-level parts in the hierarchy generally) do not fully determine the organization of the higher-level structures and organizational patterns (Harold 2001:125). It follows that the genetic information that codes for proteins does not determine these higher-level structures either.

These considerations pose another challenge to the sufficiency of the neo-Darwinian mechanism. Neo-Darwinism seeks to explain the origin of new information, form, and structure as a result of selection acting on randomly arising variation at a very low level within the biological hierarchy, namely, within the genetic text. Yet major morphological innovations depend on a specificity of arrangement at a much higher level of the organizational hierarchy, a level that DNA alone does not determine. Yet if DNA is not wholly responsible for body plan morphogenesis, then DNA sequences can mutate indefinitely, without regard to realistic probabilistic limits, and still not produce a new body plan. Thus, the mechanism of natural selection acting on random mutations in DNA cannot *in principle* generate novel body plans, including those that first arose in the Cambrian explosion.

Of course, it could be argued that, while many single proteins do not by themselves determine cellular structures and/or body plans, proteins acting in concert with other proteins or suites of proteins could determine such higher-level form. For example, it might be pointed out that the tubulin subunits (cited above) are assembled by other helper proteins—gene products—called Microtubule Associated Proteins (MAPS). This might seem to suggest that genes and gene products alone do suffice to determine the development of the three-dimensional structure of the cytoskeleton.

Yet, MAPS, and indeed many other nec-

essary proteins, are only part of the story. The location of specified target sites on the interior of the cell membrane also helps to determine the shape of the cytoskeleton. Similarly, so does the position and structure of the centrosome which nucleates the microtubules that form the cytoskeleton. While both the membrane targets and the centrosomes are made of proteins, the location and form of these structures is not wholly determined by the proteins that form them. Indeed, centrosome structure and membrane patterns *as a whole* convey three-dimensional structural information that helps determine the structure of the cytoskeleton and the location of its subunits (McNiven & Porter 1992:313–329). Moreover, the centrioles that compose the centrosomes replicate independently of DNA replication (Lange et al. 2000:235–249, Marshall & Rosenbaum 2000:187–205). The daughter centriole receives its form from the overall structure of the mother centriole, not from the individual gene products that constitute it (Lange et al. 2000). In ciliates, microsurgery on cell membranes can produce heritable changes in membrane patterns, even though the DNA of the ciliates has not been altered (Sonneborn 1970:1–13, Frankel 1980:607–623; Nanney 1983:163–170). This suggests that membrane patterns (as opposed to membrane constituents) are impressed directly on daughter cells. In both cases, form is transmitted from parent three-dimensional structures to daughter three-dimensional structures directly and is not wholly contained in constituent proteins or genetic information (Moss 2004).

Thus, in each new generation, the form and structure of the cell arises as the result of *both* gene products and pre-existing three-dimensional structure and organization. Cellular structures are built from proteins, but proteins find their way to correct locations in part because of pre-existing three-dimensional patterns and organization inherent in cellular structures. Pre-existing three-dimensional form present in the preceding generation (whether inherent in the cell membrane, the centrosomes, the cytoskeleton or other features of the fertilized egg) contributes to the production of form in the next generation. Neither structural proteins alone, nor the genes that code for them, are sufficient to determine the three-dimensional shape and structure of the entities they form. Gene products provide necessary, but not sufficient conditions, for the development of three-dimensional structure within cells, organs and body plans (Harold 1995:2767). But if this is so, then natural selection acting on genetic variation alone cannot produce the new forms that arise in history of life.

Self-Organizational Models

Of course, neo-Darwinism is not the only evolutionary theory for explaining the origin of novel biological form. Kauffman (1995) doubts the efficacy of the mutation/selection mechanism. Nevertheless, he has advanced a self-organizational theory to account for the emergence of new form, and presumably the information necessary to generate it. Whereas neo-Darwinism attempts to explain new form as the consequence of selection acting on random mutation, Kauffman suggests that selection acts, not mainly on random variations, but on emergent patterns of order that self-organize via the laws of nature.

Kauffman (1995:47–92) illustrates how this might work with various model systems in a computer environment. In one, he conceives a system of buttons connected by strings. Buttons represent novel genes or gene products; strings represent the law-like forces of interaction that obtain between gene products—i.e., proteins. Kauffman suggests that when the complexity of the system (as represented by the number of buttons and strings) reaches a critical threshold, new modes of organization can arise in the system "for free"—that is, naturally and spontaneously—after the manner of a phase transition in chemistry.

Another model that Kauffman develops is a system of interconnected lights. Each light can flash in a variety of states—on, off, twinkling, etc. Since there is more than one possible state for each light, and many lights, there are a vast number of possible states that the system can adopt. Further, in his system, rules determine how past states will influence future states. Kauffman asserts that, as a result of these rules, the system will, if properly tuned, eventually produce a kind of order in which a few basic patterns of light activity recur with greater-than-random frequency. Since these actual patterns of light activity represent a small portion of the total number of possible states in which the system can reside, Kaufman seems to imply that self-organizational laws might similarly result in highly improbable biological outcomes—perhaps even sequences (of bases or amino acids) within a much larger sequence space of possibilities.

Do these simulations of self-organizational processes accurately model the origin of novel genetic information? It is hard to think so.

First, in both examples, Kaufmann presupposes but does not explain significant sources of preexisting information. In his buttons-and-strings system, the buttons represent proteins, themselves packets of CSI, and the result of pre-existing genetic information. Where does this information come from? Kauffman (1995) doesn't say, but the origin of such information is an essential part of what needs to be explained in the history of life. Similarly, in his light system, the order that allegedly arises for "for free" actually arises only if the programmer of the model system "tunes" it in such a way as to keep it from either (a) generating an excessively rigid order or (b) devolving into chaos (pp. 86–88). Yet this necessary tuning involves an intelligent programmer selecting certain parameters and excluding others—that is, inputting information.

Second, Kauffman's model systems are not constrained by functional consider-

ations and thus are not analogous to biological systems. A system of interconnected lights governed by pre-programmed rules may well settle into a small number of patterns within a much larger space of possibilities. But because these patterns have no function, and need not meet any functional requirements, they have no specificity analogous to that present in actual organisms. Instead, examination of Kauffman's (1995) model systems shows that they do not produce sequences or systems characterized by *specified* complexity, but instead by large amounts of symmetrical order or internal redundancy interspersed with aperiodicity or (mere) complexity (pp. 53, 89, 102). Getting a law-governed system to generate repetitive patterns of flashing lights, even with a certain amount of variation, is clearly interesting, but not biologically relevant. On the other hand, a system of lights flashing the title of a Broadway play would model a biologically relevant self-organizational process, at least if such a meaningful or functionally specified sequence arose without intelligent agents previously programming the system with equivalent amounts of CSI. In any case, Kauffman's systems do not produce *specified* complexity, and thus do not offer promising models for explaining the new genes and proteins that arose in the Cambrian.

Even so, Kauffman suggests that his self-organizational models can specifically elucidate aspects of the Cambrian explosion. According to Kauffman (1995:199–201), new Cambrian animals emerged as the result of "long jump" mutations that established new body plans in a discrete rather than gradual fashion. He also recognizes that mutations affecting early development are almost inevitably harmful. Thus, he concludes that body plans, once established, will not change, and that any subsequent evolution must occur within an established body plan (Kauffman 1995:201). And indeed, the fossil record does show a curious (from a neo-Darwinian point of view) top-down pattern of appearance, in which high-

er taxa (and the body plans they represent) appear first, only later to be followed by the multiplication of lower taxa representing variations within those original body designs (Erwin et al. 1987, Lewin 1988, Valentine & Jablonski 2003:518). Further, as Kauffman expects, body plans appear suddenly and persist without significant modification over time.

But here, again, Kauffman begs the most important question, which is: what produces the new Cambrian body plans in the first place? Granted, he invokes "long jump mutations" to explain this, but he identifies no specific self-organizational process that can produce such mutations. Moreover, he concedes a principle that undermines the plausibility of his own proposal. Kauffman acknowledges that mutations that occur early in development are almost inevitably deleterious. Yet developmental biologists know that these are the only kind of mutations that have a realistic chance of producing large-scale evolutionary change—i.e., the big jumps that Kauffman invokes. Though Kauffman repudiates the neo-Darwinian reliance upon random mutations in favor of self-organizing order, in the end, he must invoke the most implausible kind of random mutation in order to provide a self-organizational account of the new Cambrian body plans. Clearly, his model is not sufficient.

Punctuated Equilibrium

Of course, still other causal explanations have been proposed. During the 1970s, the paleontologists Eldredge and Gould (1972) proposed the theory of evolution by punctuated equilibrium in order to account for a pervasive pattern of "sudden appearance" and "stasis" in the fossil record. Though advocates of punctuated equilibrium were mainly seeking to describe the fossil record more accurately than earlier gradualist neo-Darwinian models had done, they did also propose a mechanism—known as species selection—by which the large morphologi-

cal jumps evident in fossil record might have been produced. According to punctuationalists, natural selection functions more as a mechanism for selecting the fittest species rather than the most-fit individual among a species. Accordingly, on this model, morphological change should occur in larger, more discrete intervals than it would given a traditional neo-Darwinian understanding.

Despite its virtues as a descriptive model of the history of life, punctuated equilibrium has been widely criticized for failing to provide a mechanism sufficient to produce the novel form characteristic of higher taxonomic groups. For one thing, critics have noted that the proposed mechanism of punctuated evolutionary change simply lacked the raw material upon which to work. As Valentine and Erwin (1987) note, the fossil record fails to document a large pool of species prior to the Cambrian. Yet the proposed mechanism of species selection requires just such a pool of species upon which to act. Thus, they conclude that the mechanism of species selection probably does not resolve the problem of the origin of the higher taxonomic groups (p. 96).[8] Further, punctuated equilibrium has not addressed the more specific and fundamental problem of explaining the origin of the new biological information (whether genetic or epigenetic) necessary to produce novel biological form. Advocates of punctuated equilibrium might assume that the new species (upon which natural selection acts) arise by known micro-evolutionary processes of speciation (such as founder ef-

---

[8] Erwin (2004:21), although friendly to the possibility of species selection, argues that Gould provides little evidence for its existence. "The difficulty" writes Erwin of species selection, ". . . is that we must rely on Gould's arguments for theoretical plausibility and sufficient relative frequency. Rarely is a mass of data presented to justify and support Gould's conclusion." Indeed, Gould (2002) himself admitted that species selection remains largely a hypothetical construct: "I freely admit that well-documented cases of species selection do not permeate the literature" (p. 710).

fect, genetic drift or bottleneck effect) that do not necessarily depend upon mutations to produce adaptive changes. But, in that case, the theory lacks an account of how the specifically *higher* taxa arise. Species selection will only produce more fit species. On the other hand, if punctuationalists assume that processes of genetic mutation can produce more fundamental morphological changes and variations, then their model becomes subject to the same problems as neo-Darwinism (see above). This dilemma is evident in Gould (2002:710) insofar as his attempts to explain adaptive complexity inevitably employ classical neo-Darwinian modes of explanation.[9]

Structuralism

Another attempt to explain the origin of form has been proposed by the structuralists such as Gerry Webster and Brian Goodwin (1984, 1996). These biologists, drawing on the earlier work of D'Arcy Thompson (1942), view biological form as the result of structural constraints imposed upon matter by morphogenetic rules or laws. For reasons similar to those discussed above, the structuralists have insisted that these generative or morphogenetic rules do not reside in the lower level building materials of or-

---

[9] "I do not deny either the wonder, or the powerful importance, of organized adaptive complexity. I recognize that we know no mechanism for the origin of such organismal features other than conventional natural selection at the organismic level—for the sheer intricacy and elaboration of good biomechanical design surely precludes either random production, or incidental origin as a side consequence of active processes at other levels" (Gould 2002:710). "Thus, we do not challenge the efficacy or the cardinal importance of organismal selection. As previously discussed, I fully agree with Dawkins (1986) and others that one cannot invoke a higher-level force like species selection to explain 'things that organisms do'—in particular, the stunning panoply of organismic adaptations that has always motivated our sense of wonder about the natural world, and that Darwin (1859) described, in one of his most famous lines (3), as 'that perfection of structure and coadaptation which most justly excites our admiration'" (Gould 2002:886).

ganisms, whether in genes or proteins. Webster and Goodwin (1984:510–511) further envision morphogenetic rules or laws operating ahistorically, similar to the way in which gravitational or electro-magnetic laws operate. For this reason, structuralists see phylogeny as of secondary importance in understanding the origin of the higher taxa, though they think that transformations of form can occur. For structuralists, constraints on the arrangement of matter arise not mainly as the result of historical contingencies—such as environmental changes or genetic mutations—but instead because of the continuous ahistorical operation of fundamental laws of form—laws that organize or inform matter.

While this approach avoids many of the difficulties currently afflicting neo-Darwinism (in particular those associated with its "genocentricity"), critics (such as Maynard Smith 1986) of structuralism have argued that the structuralist explanation of form lacks specificity. They note that structuralists have been unable to say just where laws of form reside—whether in the universe, or in every possible world, or in organisms as a whole, or in just some part of organisms. Further, according to structuralists, morphogenetic laws are mathematical in character. Yet, structuralists have yet to specify the mathematical formulae that determine biological forms.

Others (Yockey 1992; Polanyi 1967, 1968; Meyer 2003) have questioned whether physical laws could in principle generate the kind of complexity that characterizes biological systems. Structuralists envision the existence of biological laws that produce form in much the same way that physical laws produce form. Yet the forms that physicists regard as manifestations of underlying laws are characterized by large amounts of symmetric or redundant order, by relatively simple patterns such as vortices or gravitational fields or magnetic lines of force. Indeed, physical laws are typically expressed as differential equations (or algorithms) that almost by definition describe recurring phe-

nomena—patterns of compressible "order" not "complexity" as defined by algorithmic information theory (Yockey 1992:77–83). Biological forms, by contrast, manifest greater complexity and derive in ontogeny from highly complex initial conditions—i.e., non-redundant sequences of nucleotide bases in the genome and other forms of information expressed in the complex and irregular three-dimensional topography of the organism or the fertilized egg. Thus, the kind of form that physical laws produce is not analogous to biological form—at least not when compared from the standpoint of (algorithmic) complexity. Further, physical laws lack the information content to specify biology systems. As Polanyi (1967, 1968) and Yockey (1992:290) have shown, the laws of physics and chemistry allow, but do not determine, distinctively biological modes of organization. In other words, living systems are consistent with, but not deducible, from physical-chemical laws (1992:290).

Of course, biological systems do manifest some reoccurring patterns, processes and behaviors. The same type of organism develops repeatedly from similar ontogenetic processes in the same species. Similar processes of cell division re-occur in many organisms. Thus, one might describe certain biological processes as law-governed. Even so, the existence of such biological regularities does not solve the problem of the origin of form and information, since the recurring processes described by such biological laws (if there be such laws) only occur as the result of pre-existing stores of (genetic and/or epigenetic) information and these information-rich initial conditions impose the constraints that produce the recurring behavior in biological systems. (For example, processes of cell division recur with great frequency in organisms, but depend upon information-rich DNA and proteins molecules.) In other words, distinctively biological regularities depend upon pre-existing biological information. Thus, appeals to higher-level biological laws presuppose, but do not explain, the

origination of the information necessary to morphogenesis.

Thus, structuralism faces a difficult in principle dilemma. On the one hand, physical laws produce very simple redundant patterns that lack the complexity characteristic of biological systems. On the other hand, distinctively biological laws—if there are such laws—depend upon pre-existing information-rich structures. In either case, laws are not good candidates for explaining the origination of biological form or the information necessary to produce it.

Cladism: An Artifact of Classification?

Some cladists have advanced another approach to the problem of the origin of form, specifically as it arises in the Cambrian. They have argued that the problem of the origin of the phyla is an artifact of the classification system, and therefore, does not require explanation. Budd and Jensen (2000), for example, argue that the problem of the Cambrian explosion resolves itself if one keeps in mind the cladistic distinction between "stem" and "crown" groups. Since crown groups arise whenever new characters are added to simpler more ancestral stem groups during the evolutionary process, new phyla will inevitably arise once a new stem group has arisen. Thus, for Budd and Jensen what requires explanation is not the crown groups corresponding to the new Cambrian phyla, but the earlier more primitive stem groups that presumably arose deep in the Proterozoic. Yet since these earlier stem groups are by definition less derived, explaining them will be considerably easier than explaining the origin of the Cambrian animals *de novo*. In any case, for Budd and Jensen the explosion of new phyla in the Cambrian does not require explanation. As they put it, "given that the early branching points of major clades is an inevitable result of clade diversification, the alleged phenomenon of the phyla appearing early and remaining morphologically static is not seen to require

particular explanation" (Budd & Jensen 2000:253).

While superficially plausible, perhaps, Budd and Jensen's attempt to explain away the Cambrian explosion begs crucial questions. Granted, as new characters are added to existing forms, novel morphology and greater morphological disparity will likely result. But what causes new characters to arise? And how does the information necessary to produce new characters originate? Budd and Jensen do not specify. Nor can they say how derived the ancestral forms are likely to have been, and what processes, might have been sufficient to produce them. Instead, they simply assume the sufficiency of known neo-Darwinian mechanisms (Budd & Jensen 2000:288). Yet, as shown above, this assumption is now problematic. In any case, Budd and Jensen do not explain what causes the origination of biological form and information.

Convergence and Teleological Evolution

More recently, Conway Morris (2000, 2003c) has suggested another possible explanation based on the tendency for evolution to converge on the same structural forms during the history of life. Conway Morris cites numerous examples of organisms that possess very similar forms and structures, even though such structures are often built from different material substrates and arise (in ontogeny) by the expression of very different genes. Given the extreme improbability of the same structures arising by random mutation and selection in disparate phylogenies, Conway Morris argues that the pervasiveness of convergent structures suggests that evolution may be in some way "channeled" toward similar functional and/or structural endpoints. Such an end-directed understanding of evolution, he admits, raises the controversial prospect of a teleological or purposive element in the history of life. For this reason, he argues that the phenomenon of convergence has received less attention

than it might have otherwise. Nevertheless, he argues that just as physicists have re-opened the question of design in their discussions of anthropic fine-tuning, the ubiquity of convergent structures in the history of life has led some biologists (Denton 1998) to consider extending teleological thinking to biology. And, indeed, Conway Morris himself intimates that the evolutionary process might be "underpinned by a purpose" (2000:8, 2003b:511).

Conway Morris, of course, considers this possibility in relation to a very specific aspect of the problem of organismal form, namely, the problem of explaining why the same forms arise repeatedly in so many disparate lines of decent. But this raises a question. Could a similar approach shed explanatory light on the more general causal question that has been addressed in this review? Could the notion of purposive design help provide a more adequate explanation for the origin of organismal form generally? Are there reasons to consider design as an explanation for the origin of the biological information necessary to produce the higher taxa and their corresponding morphological novelty?

The remainder of this review will suggest that there are such reasons. In so doing, it may also help explain why the issue of teleology or design has re-emerged within the scientific discussion of biological origins (Denton 1986, 1998; Thaxton et al. 1992; Kenyon & Mills 1996; Behe 1996, 2004; Dembski 1998, 2002, 2004; Conway Morris 2000, 2003a, 2003b; Lönnig 2001; Lönnig & Saedler 2002; Nelson & Wells 2003; Meyer 2003, 2004; Bradley 2004) and why some scientists and philosophers of science have considered teleological explanations for the origin of form and information despite strong methodological prohibitions against design as a scientific hypothesis (Gillespie 1979, Lenior 1982:4).

First, the possibility of design as an explanation follows logically from a consideration of the deficiencies of neo-Darwinism and other current theories as explana-

tions for some of the more striking "appearances of design" in biological systems. Neo-Darwinists such as Ayala (1994:5), Dawkins (1986:1), Mayr (1982:xi–xii) and Lewontin (1978) have long acknowledged that organisms appear to have been designed. Of course, neo-Darwinists assert that what Ayala (1994:5) calls the "obvious design" of living things is only apparent since the selection/mutation mechanism can explain the origin of complex form and organization in living systems without an appeal to a designing agent. Indeed, neo-Darwinists affirm that mutation and selection— and perhaps other similarly undirected mechanisms—are fully sufficient to explain the appearance of design in biology. Self-organizational theorists and punctuationalists modify this claim, but affirm its essential tenet. Self-organization theorists argue that natural selection acting on self-organizing order can explain the complexity of living things—again, without any appeal to design. Punctuationalists similarly envision natural selection acting on newly arising species with no actual design involved.

And clearly, the neo-Darwinian mechanism does explain many appearances of design, such as the adaptation of organisms to specialized environments that attracted the interest of 19th century biologists. More specifically, known micro-evolutionary processes appear quite sufficient to account for changes in the size of Galapagos finch beaks that have occurred in response to variations in annual rainfall and available food supplies (Weiner 1994, Grant 1999).

But does neo-Darwinism, or any other fully materialistic model, explain all appearances of design in biology, including the body plans and information that characterize living systems? Arguably, biological forms—such as the structure of a chambered nautilus, the organization of a trilobite, the functional integration of parts in an eye or molecular machine—attract our attention in part because the organized complexity of such systems seems reminiscent of our own designs. Yet, this review

has argued that neo-Darwinism does not adequately account for the origin of all appearances of design, especially if one considers animal body plans, and the information necessary to construct them, as especially striking examples of the appearance of design in living systems. Indeed, Dawkins (1995:11) and Gates (1996:228) have noted that genetic information bears an uncanny resemblance to computer software or machine code. For this reason, the presence of CSI in living organisms, and the discontinuous increases of CSI that occurred during events such as the Cambrian explosion, appears at least suggestive of design.

Does neo-Darwinism or any other purely materialistic model of morphogenesis account for the origin of the genetic and other forms of CSI necessary to produce novel organismal form? If not, as this review has argued, could the emergence of novel information-rich genes, proteins, cell types and body plans have resulted from actual design, rather than a purposeless process that merely mimics the powers of a designing intelligence? The logic of neo-Darwinism, with its specific claim to have accounted for the appearance of design, would itself seem to open the door to this possibility. Indeed, the historical formulation of Darwinism in dialectical opposition to the design hypothesis (Gillespie 1979), coupled with neo-Darwinism's inability to account for many salient appearances of design including the emergence of form and information, would seem logically to re-open the possibility of actual (as opposed to apparent) design in the history of life.

A second reason for considering design as an explanation for these phenomena follows from the importance of explanatory power to scientific theory evaluation and from a consideration of the potential explanatory power of the design hypothesis. Studies in the methodology and philosophy of science have shown that many scientific theories, particularly in the historical sciences, are formulated and justified as inferences to the best explanation (Lipton 1991:

32–88, Brush 1989:1124–1129, Sober 2000:44). Historical scientists, in particular, assess or test competing hypotheses by evaluating which hypothesis would, if true, provide the best explanation for some set of relevant data (Meyer 1991, 2002; Cleland 2001:987–989, 2002:474–496).[10] Those

---

[10] Theories in the historical sciences typically make claims about what happened in the past, or what happened in the past to cause particular events to occur (Meyer 1991:57–72). For this reason, historical scientific theories are rarely tested by making predictions about what will occur under controlled laboratory conditions (Cleland 2001:987, 2002:474–496). Instead, such theories are usually tested by comparing their explanatory power against that of their competitors with respect to already known facts. Even in the case in which historical theories make claims about past causes they usually do so on the basis of pre-existing knowledge of cause and effect relationships. Nevertheless, prediction may play a limited role in testing historical scientific theories since such theories may have implications as to what kind of evidence is likely to emerge in the future. For example, neo-Darwinism affirms that new functional sections of the genome arise by trial and error process of mutation and subsequent selection. For this reason, historically many neo-Darwinists expected or predicted that the large non-coding regions of the genome—so-called "junk DNA—would lack function altogether (Orgel & Crick 1980). On this line of thinking, the non-functional sections of the genome represent nature's failed experiments that remain in the genome as a kind of artifact of the past activity of the mutation and selection process. Advocates of the design hypotheses on the other hand, would have predicted that non-coding regions of the genome might well reveal hidden functions, not only because design theorists do not think that new genetic information arises by a trial and error process of mutation and selection, but also because designed systems are often functionally polyvalent. Even so, as new studies reveal more about the functions performed by the non-coding regions of the genome (Gibbs 2003), the design hypothesis can no longer be said to make this claim in the form of a specifically future-oriented prediction. Instead, the design hypothesis might be said to gain confirmation or support from its ability to explain this now known evidence, albeit after the fact. Of course, neo-Darwinists might also amend their original prediction using various auxiliary hypotheses to explain away the presence of newly discovered functions in the non-coding regions of DNA. In both cases, considerations of *ex post facto* explanatory power re-emerge as central to assessing and testing competing historical theories.

with greater explanatory power are typically judged to be better, more probably true, theories. Darwin (1896:437) used this method of reasoning in defending his theory of universal common descent. Moreover, contemporary studies on the method of "inference to the best explanation" have shown that determining which among a set of competing possible explanations constitutes the best depends upon judgments about the causal adequacy, or "causal powers," of competing explanatory entities (Lipton 1991:32–88). In the historical sciences, uniformitarian and/or actualistic (Gould 1965, Simpson 1970, Rutten 1971, Hooykaas 1975) canons of method suggest that judgments about causal adequacy should derive from our present knowledge of cause and effect relationships. For historical scientists, "the present is the key to the past" means that present experience-based knowledge of cause and effect relationships typically guides the assessment of the plausibility of proposed causes of past events.

Yet it is precisely for this reason that current advocates of the design hypothesis want to reconsider design as an explanation for the origin of biological form and information. This review, and much of the literature it has surveyed, suggests that four of the most prominent models for explaining the origin of biological form fail to provide adequate causal explanations for the discontinuous increases of CSI that are required to produce novel morphologies. Yet, we have repeated experience of rational and conscious agents—in particular ourselves—generating or causing increases in complex specified information, both in the form of sequence-specific lines of code and in the form of hierarchically arranged systems of parts.

In the first place, intelligent human agents—in virtue of their rationality and consciousness—have demonstrated the power to produce information in the form of linear sequence-specific arrangements of characters. Indeed, experience affirms that

information of this type routinely arises from the activity of intelligent agents. A computer user who traces the information on a screen back to its source invariably comes to a *mind*—that of a software engineer or programmer. The information in a book or inscription ultimately derives from a writer or scribe—from a mental, rather than a strictly material, cause. Our experience-based knowledge of information-flow confirms that systems with large amounts of specified complexity (especially codes and languages) invariably originate from an intelligent source—from a mind or personal agent. As Quastler (1964) put it, the "creation of new information is habitually associated with conscious activity" (p. 16). Experience teaches this obvious truth.

Further, the highly specified hierarchical arrangements of parts in animal body plans also suggest *design,* again because of our experience of the kinds of features and systems that designers can and do produce. At every level of the biological hierarchy, organisms require specified and highly improbable arrangements of lower-level constituents in order to maintain their form and function. Genes require specified arrangements of nucleotide bases; proteins require specified arrangements of amino acids; new cell types require specified arrangements of systems of proteins; body plans require specialized arrangements of cell types and organs. Organisms not only contain information-rich components (such as proteins and genes), but they comprise information-rich arrangements of those components and the systems that comprise them. Yet we know, based on our present experience of cause and effect relationships, that design engineers—possessing purposive intelligence and rationality—have the ability to produce information-rich hierarchies in which both individual modules and the arrangements of those modules exhibit complexity and specificity—information so defined. Individual transistors, resistors, and capacitors exhibit considerable complexity and specificity of design; at a higher level of organization,

their specific arrangement within an integrated circuit represents additional information and reflects further design. Conscious and rational agents have, as part of their powers of purposive intelligence, the capacity to design information-rich parts and to organize those parts into functional information-rich systems and hierarchies. Further, we know of no other causal entity or process that has this capacity. Clearly, we have good reason to doubt that mutation and selection, self-organizational processes or laws of nature, can produce the information-rich components, systems, and body plans necessary to explain the origination of morphological novelty such as that which arises in the Cambrian period.

There is a third reason to consider purpose or design as an explanation for the origin of biological form and information: purposive agents have just those necessary powers that natural selection lacks as a condition of its causal adequacy. At several points in the previous analysis, we saw that natural selection lacked the ability to generate novel information precisely because it can only act *after* new functional CSI has arisen. Natural selection can favor new proteins, and genes, but only after they perform some function. The job of generating new functional genes, proteins and systems of proteins therefore falls entirely to random mutations. Yet without functional criteria to guide a search through the space of possible sequences, random variation is probabilistically doomed. What is needed is not just a source of variation (i.e., the freedom to search a space of possibilities) or a mode of selection that can operate after the fact of a successful search, but instead a means of selection that (a) operates during a search—before success—and that (b) is guided by information about, or knowledge of, a functional target.

Demonstration of this requirement has come from an unlikely quarter: genetic algorithms. Genetic algorithms are programs that allegedly simulate the creative power of mutation and selection. Dawkins and

Küppers, for example, have developed computer programs that putatively simulate the production of genetic information by mutation and natural selection (Dawkins 1986:47–49, Küppers 1987:355–369). Nevertheless, as shown elsewhere (Meyer 1998: 127–128, 2003:247–248), these programs only succeed by the illicit expedient of providing the computer with a "target sequence" and then treating relatively greater proximity to *future* function (i.e., the target sequence), not actual present function, as a selection criterion. As Berlinski (2000) has argued, genetic algorithms need something akin to a "forward looking memory" in order to succeed. Yet such foresighted selection has no analogue in nature. In biology, where differential survival depends upon maintaining function, selection cannot occur before new functional sequences arise. Natural selection lacks foresight.

What natural selection lacks, intelligent selection—purposive or goal-directed design—provides. Rational agents can arrange both matter and symbols with distant goals in mind. In using language, the human mind routinely "finds" or generates highly improbable linguistic sequences to convey an intended or *pre*conceived idea. In the process of thought, functional objectives precede and constrain the selection of words, sounds and symbols to generate functional (and indeed meaningful) sequences from among a vast ensemble of meaningless alternative combinations of sound or symbol (Denton 1986:309–311). Similarly, the construction of complex technological objects and products, such as bridges, circuit boards, engines and software, result from the application of goal-directed constraints (Polanyi 1967, 1968). Indeed, in all functionally integrated complex systems where the cause is known by experience or observation, design engineers or other intelligent agents applied boundary constraints to limit possibilities in order to produce improbable forms, sequences or structures. Rational agents have repeatedly demonstrated the capacity to constrain the

possible to actualize improbable but initially unrealized future functions. Repeated experience affirms that intelligent agents (minds) uniquely possess such causal powers.

Analysis of the problem of the origin of biological information, therefore, exposes a deficiency in the causal powers of natural selection that corresponds precisely to powers that agents are uniquely known to possess. Intelligent agents have foresight. Such agents can select functional goals *before* they exist. They can devise or select material means to accomplish those ends from among an array of possibilities and then actualize those goals in accord with a *pre*conceived design plan or set of functional requirements. Rational agents can constrain combinatorial space with distant outcomes in mind. The causal powers that natural selection lacks—almost by definition—are associated with the attributes of consciousness and rationality—with purposive intelligence. Thus, by invoking design to explain the origin of new biological information, contemporary design theorists are not positing an arbitrary explanatory element unmotivated by a consideration of the evidence. Instead, they are positing an entity possessing precisely the attributes and causal powers that the phenomenon in question requires as a condition of its production and explanation.

Conclusion

An experience-based analysis of the causal powers of various explanatory hypotheses suggests purposive or intelligent design as a causally adequate—and perhaps the most causally adequate—explanation for the origin of the complex specified information required to build the Cambrian animals and the novel forms they represent. For this reason, recent scientific interest in the design hypothesis is unlikely to abate as biologists continue to wrestle with the problem of the origination of biological form and the higher taxa.

Literature Cited

Adams, M. D. et alia. 2000. The genome sequence of *Drosophila melanogaster.*—Science 287:2185–2195.

Aris-Brosou, S., & Z. Yang. 2003. Bayesian models of episodic evolution support a late Precambrian explosive diversification of the Metazoa.—Molecular Biology and Evolution 20:1947–1954.

Arthur, W. 1997. The origin of animal body plans. Cambridge University Press, Cambridge, United Kingdom.

Axe, D. D. 2000. Extreme functional sensitivity to conservative amino acid changes on enzyme exteriors.—Journal of Molecular Biology 301(3): 585–596.

———. 2004. Estimating the prevalence of protein sequences adopting functional enzyme folds.—Journal of Molecular Biology (in press).

Ayala, F. 1994. Darwin's revolution. Pp. 1–17 *in* J. Campbell and J. Schopf, eds., Creative evolution?! Jones and Bartlett Publishers, Boston, Massachusetts.

———, A. Rzhetsky, & F. J. Ayala. 1998. Origin of the metazoan phyla: molecular clocks confirm paleontological estimates.—Proceedings of the National Academy of Sciences USA. 95:606–611.

Becker, H., & W. Lönnig. 2001. Transposons: eukaryotic. Pp. 529–539 *in* Nature encyclopedia of life sciences, vol. 18. Nature Publishing Group, London, United Kingdom.

Behe, M. 1992. Experimental support for regarding functional classes of proteins to be highly isolated from each other. Pp. 60–71 *in* J. Buell and V. Hearn, eds., Darwinism: science or philosophy? Foundation for Thought and Ethics, Richardson, Texas.

———. 1996. Darwin's black box. The Free Press, New York.

———. 2004. Irreducible complexity: obstacle to Darwinian evolution. Pp. 352–370 *in* W. A. Dembski and M. Ruse, eds., Debating design: from Darwin to DNA. Cambridge University Press, Cambridge, United Kingdom.

Benton, M., & F. J. Ayala. 2003. Dating the tree of life.—Science 300:1698–1700.

Berlinski, D. 2000. "On assessing genetic algorithms." Public lecture. Conference: Science and evidence of design in the universe. Yale University, November 4, 2000.

Blanco, F., I. Angrand, & L. Serrano. 1999. Exploring the confirmational properties of the sequence space between two proteins with different folds: an experimental study.—Journal of Molecular Biology 285: 741–753.

Bowie, J., & R. Sauer. 1989. Identifying determinants of folding and activity for a protein of unknown sequences: tolerance to amino acid substitution.—Proceedings of the National Academy of Sciences, U.S.A. 86:2152–2156.

Bowring, S. A., J. P. Grotzinger, C. E. Isachsen, A. H. Knoll, S. M. Pelechaty, & P. Kolosov. 1993. Calibrating rates of early Cambrian evolution.—Science 261:1293–1298.

———. 1998a. A new look at evolutionary rates in deep time: Uniting paleontology and high-precision geochronology.—GSA Today 8:1–8.

———. 1998b. Geochronology comes of age.—Geotimes 43:36–40.

Bradley, W. 2004. Information, entropy and the origin of life. Pp. 331–351 *in* W. A. Dembski and M. Ruse, eds., Debating design: from Darwin to DNA. Cambridge University Press, Cambridge, United Kingdom.

Brocks, J. J., G. A. Logan, R. Buick, & R. E. Summons. 1999. Archean molecular fossils and the early rise of eukaryotes.—Science 285:1033–1036.

Brush, S. G. 1989. Prediction and theory evaluation: the case of light bending.—Science 246:1124–1129.

Budd, G. E., & S. E. Jensen. 2000. A critical reappraisal of the fossil record of the bilaterial phyla.—Biological Reviews of the Cambridge Philosophical Society 75:253–295.

Carroll, R. L. 2000. Towards a new evolutionary synthesis.—Trends in Ecology and Evolution 15: 27–32.

Cleland, C. 2001. Historical science, experimental science, and the scientific method.—Geology 29: 987–990.

———. 2002. Methodological and epistemic differences between historical science and experimental science.—Philosophy of Science 69: 474–496.

Chothia, C., I. Gelfland, & A. Kister. 1998. Structural determinants in the sequences of immunoglobulin variable domain.—Journal of Molecular Biology 278:457–479.

Conway Morris, S. 1998a. The question of metazoan monophyly and the fossil record.—Progress in Molecular and Subcellular Biology 21:1–9.

———. 1998b. Early Metazoan evolution: reconciling paleontology and molecular biology.—American Zoologist 38 (1998):867–877.

———. 2000. Evolution: bringing molecules into the fold.—Cell 100:1–11.

———. 2003a. The Cambrian "explosion" of metazoans. Pp. 13–32 *in* G. B. Müller and S. A. Newman, eds., Origination of organismal form: beyond the gene in developmental and evolutionary biology. The M.I.T. Press, Cambridge, Massachusetts.

———. 2003b. Cambrian "explosion" of metazoans and molecular biology: would Darwin be sat-

isfied?—International Journal of Developmental Biology 47(7–8):505–515.

———. 2003c. Life's solution: inevitable humans in a lonely universe. Cambridge University Press, Cambridge, United Kingdom.

Crick, F. 1958. On protein synthesis.—Symposium for the Society of Experimental Biology. 12(1958): 138–163.

Darwin, C. 1859. On the origin of species. John Murray, London, United Kingdom.

———. 1896. Letter to Asa Gray. P. 437 in F. Darwin, ed., Life and letters of Charles Darwin, vol. 1. D. Appleton, London, United Kingdom.

Davidson, E. 2001. Genomic regulatory systems: development and evolution. Academic Press, New York, New York.

Dawkins, R. 1986. The blind watchmaker. Penguin Books, London, United Kingdom.

———. 1995. River out of Eden. Basic Books, New York.

———. 1996. Climbing Mount Improbable. W. W. Norton & Company, New York.

Dembksi, W. A. 1998. The design inference. Cambridge University Press, Cambridge, United Kingdom.

———. 2002. No free lunch: why specified complexity cannot be purchased without intelligence. Rowman & Littlefield, Lanham, Maryland.

———. 2004. The logical underpinnings of intelligent design. Pp. 311–330 in W. A. Dembski and M. Ruse, eds., Debating design: from Darwin to DNA. Cambridge University Press, Cambridge, United Kingdom.

Denton, M. 1986. Evolution: a theory in crisis. Adler & Adler, London, United Kingdom.

———. 1998. Nature's destiny. The Free Press, New York.

Eden, M. 1967. The inadequacies of neo-Darwinian evolution as a scientific theory. Pp. 5–12 in P. S. Morehead and M. M. Kaplan, eds., Mathematical challenges to the Darwinian interpretation of evolution. Wistar Institute Symposium Monograph. Allen R. Liss, New York.

Eldredge, N., & S. J. Gould. 1972. Punctuated equilibria: an alternative to phyletic gradualism. Pp. 82–115 in T. Schopf, ed., Models in paleobiology. W. H. Freeman, San Francisco.

Erwin, D. H. 1994. Early introduction of major morphological innovations.—Acta Palaeontologica Polonica 38:281–294.

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

———. 2004. One very long argument.—Biology and Philosophy 19:17–28.

———, J. Valentine, & D. Jablonski. 1997. The origin of animal body plans.—American Scientist 85: 126–137.

———, ———, & J. J. Sepkoski. 1987. A comparative study of diversification events: the early Paleozoic versus the Mesozoic.—Evolution 41: 1177–1186.

Foote, M. 1997. Sampling, taxonomic description, and our evolving knowledge of morphological diversity.—Paleobiology 23:181–206.

———, J. P. Hunter, C. M. Janis, & J. J. Sepkoski. 1999. Evolutionary and preservational constraints on origins of biologic groups: Divergence times of eutherian mammals.—Science 283:1310–1314.

Frankel, J. 1980. Propagation of cortical differences in *tetrahymena*.—Genetics 94:607–623.

Gates, B. 1996. The road ahead. Blue Penguin, Boulder, Colorado.

Gerhart, J., & M. Kirschner. 1997. Cells, embryos, and evolution. Blackwell Science, London, United Kingdom.

Gibbs, W. W. 2003. The unseen genome: gems among the junk.—Scientific American. 289:46–53

Gilbert, S. F., J. M. Opitz, & R. A. Raff. 1996. Resynthesizing evolutionary and developmental biology.—Developmental Biology 173:357–372.

Gillespie, N. C. 1979. Charles Darwin and the problem of creation. University of Chicago Press, Chicago.

Goodwin, B. C. 1985. What are the causes of morphogenesis?—BioEssays 3:32–36.

———. 1995. How the leopard changed its spots: the evolution of complexity. Scribner's, New York, New York.

Gould, S. J. 1965. Is uniformitarianism necessary?—American Journal of Science 263:223–228.

Gould, S. J. 2002. The structure of evolutionary theory. Harvard University Press, Cambridge, Massachusetts.

Grant, P. R. 1999. Ecology and evolution of Darwin's finches. Princeton University Press, Princeton, New Jersey.

Grimes, G. W., & K. J. Aufderheide. 1991. Cellular aspects of pattern formation: the problem of assembly. Monographs in Developmental Biology, vol. 22. Karger, Basel, Switzerland.

Grotzinger, J. P., S. A. Bowring, B. Z. Saylor, & A. J. Kaufman. 1995. Biostratigraphic and geochronologic constraints on early animal evolution.—Science 270:598–604.

Harold, F. M. 1995. From morphogenes to morphogenesis.—Microbiology 141:2765–2778.

———. 2001. The way of the cell: molecules, organisms, and the order of life. Oxford University Press, New York.

Hodge, M. J. S. 1977. The structure and strategy of Darwin's long argument.—British Journal for the History of Science 10:237–245.

Hooykaas, R. 1975. Catastrophism in geology, its scientific character in relation to actualism and

uniformitarianism. Pp. 270–316 *in* C. Albritton, ed., Philosophy of geohistory (1785–1970). Dowden, Hutchinson & Ross, Stroudsburg, Pennsylvania.

John, B., & G. Miklos. 1988. The eukaryote genome in development and evolution. Allen & Unwin, London, United Kingdom.

Kauffman, S. 1995. At home in the universe. Oxford University Press, Oxford, United Kingdom.

Kenyon, D., & G. Mills. 1996. The RNA world: a critique.—Origins & Design 17(1):9–16.

Kerr, R. A. 1993. Evolution's Big Bang gets even more explosive.—Science 261:1274–1275.

Kimura, M. 1983. The neutral theory of molecular evolution. Cambridge University Press, Cambridge, United Kingdom.

Koonin, E. 2000. How many genes can make a cell?: the minimal genome concept.—Annual Review of Genomics and Human Genetics 1:99–116.

Küppers, B. O. 1987. On the prior probability of the existence of life. Pp. 355–369 *in* L. Krüger et al., eds., The probabilistic revolution. M.I.T. Press, Cambridge, Massachusetts.

Lange, B. M. H., A. J. Faragher, P. March, & K. Gull. 2000. Centriole duplication and maturation in animal cells. Pp. 235–249 *in* R. E. Palazzo and G. P. Schatten, eds., The centrosome in cell replication and early development. Current Topics in Developmental Biology, vol. 49. Academic Press, San Diego.

Lawrence, P. A., & G. Struhl. 1996. Morphogens, compartments and pattern: lessons from Drosophila?—Cell 85:951–961.

Lenior, T. 1982. The strategy of life. University of Chicago Press, Chicago.

Levinton, J. 1988. Genetics, paleontology, and macroevolution. Cambridge University Press, Cambridge, United Kingdom.

———. 1992. The big bang of animal evolution.—Scientific American 267:84–91.

Lewin, R. 1988. A lopsided look at evolution.—Science 241:292.

Lewontin, R. 1978. Adaptation. Pp. 113–125 *in* Evolution: a Scientific American book. W. H. Freeman & Company, San Francisco.

Lipton, P. 1991. Inference to the best explanation. Routledge, New York.

Lönnig, W. E. 2001. Natural selection. Pp. 1008–1016 *in* W. E. Craighead and C. B. Nemeroff, eds., The Corsini encyclopedia of psychology and behavioral sciences, 3rd edition, vol. 3. John Wiley & Sons, New York.

———, & H. Saedler. 2002. Chromosome rearrangements and transposable elements.—Annual Review of Genetics 36:389–410.

Løvtrup, S. 1979. Semantics, logic and vulgate neo-darwinism.–Evolutionary Theory 4: 157–172.

Marshall, W. F., & J. L. Rosenbaum. 2000. Are there nucleic acids in the centrosome? Pp. 187–205 *in* R. E. Palazzo and G. P. Schatten, eds., The centrosome in cell replication and early development. Current Topics in Developmental Biology, vol. 49. San Diego, Academic Press.

Maynard Smith, J. 1986. Structuralism versus selection—is Darwinism enough? Pp. 39–46 *in* S. Rose and L. Appignanesi, eds., Science and beyond. Basil Blackwell, London, United Kingdom.

Mayr, E. 1982. Foreword. Pp. xi–xii *in* M. Ruse, Darwinism defended. Pearson Addison Wesley, Boston, Massachusetts.

McDonald, J. F. 1983. The molecular basis of adaptation: a critical review of relevant ideas and observations.—Annual Review of Ecology and Systematics 14:77–102.

McNiven, M. A., & K. R. Porter. 1992. The centrosome: contributions to cell form. Pp. 313–329 *in* V. I. Kalnins, ed., The centrosome. Academic Press, San Diego.

Meyer, S. C. 1991. Of clues and causes: a methodological interpretation of origin of life studies. Unpublished doctoral dissertation, University of Cambridge, Cambridge, United Kingdom.

———. 1998. DNA by design: an inference to the best explanation for the origin of biological information.—Rhetoric & Public Affairs, 1(4):519–555.

———. The scientific status of intelligent design: The methodological equivalence of naturalistic and non-naturalistic origins theories. Pp. 151–211 *in* Science and evidence for design in the universe. Proceedings of the Wethersfield Institute. Ignatius Press, San Francisco.

———. 2003. DNA and the origin of life: information, specification and explanation. Pp. 223–285 *in* J. A. Campbell and S. C. Meyer, eds., Darwinism, design and public education. Michigan State University Press, Lansing, Michigan.

———. 2004. The Cambrian information explosion: evidence for intelligent design. Pp. 371–391 *in* W. A. Dembski and M. Ruse, eds., Debating design: from Darwin to DNA. Cambridge University Press, Cambridge, United Kingdom.

———, M. Ross, P. Nelson, & P. Chien. 2003. The Cambrian explosion: biology's big bang. Pp. 323–402 *in* J. A. Campbell & S. C. Meyer, eds., Darwinism, design and public education. Michigan State University Press, Lansing. See also Appendix C: Stratigraphic first appearance of phyla body plans, pp. 593–598.

Miklos, G. L. G. 1993. Emergence of organizational complexities during metazoan evolution: perspectives from molecular biology, palaeontology and neo-Darwinism.—Mem. Ass. Australas. Palaeontols, 15:7–41.

Monastersky, R. 1993. Siberian rocks clock biological big bang.—Science News 144:148.

Moss, L. 2004. What genes can't do. The M.I.T. Press, Cambridge, Massachusetts.

Müller, G. B. & S. A. Newman. 2003. Origination of organismal form: the forgotten cause in evolutionary theory. Pp. 3–12 in G. B. Müller and S. A. Newman, eds., Origination of organismal form: beyond the gene in developmental and evolutionary biology. The M.I.T. Press, Cambridge, Massachusetts.

Nanney, D. L. 1983. The ciliates and the cytoplasm.—Journal of Heredity, 74:163–170.

Nelson, P., & J. Wells. 2003. Homology in biology: problem for naturalistic science and prospect for intelligent design. Pp. 303–322 in J. A. Campbell and S. C. Meyer, eds., Darwinism, design and public education. Michigan State University Press, Lansing.

Nijhout, H. F. 1990. Metaphors and the role of genes in development.—BioEssays 12:441–446.

Nüsslein-Volhard, C., & E. Wieschaus. 1980. Mutations affecting segment number and polarity in Drosophila.—Nature 287:795–801.

Ohno, S. 1996. The notion of the Cambrian pananimalia genome.—Proceedings of the National Academy of Sciences, U.S.A. 93:8475–8478.

Orgel, L. E., & F. H. Crick. 1980. Selfish DNA: the ultimate parasite.—Nature 284:604–607.

Perutz, M. F., & H. Lehmann. 1968. Molecular pathology of human hemoglobin.—Nature 219:902–909.

Polanyi, M. 1967. Life transcending physics and chemistry.—Chemical and Engineering News, 45(35):54–66.

———. 1968. Life's irreducible structure.—Science 160:1308–1312, especially p. 1309.

Pourquié, O. 2003. Vertebrate somitogenesis: a novel paradigm for animal segmentation?—International Journal of Developmental Biology 47(7–8):597–603.

Quastler, H. 1964. The emergence of biological organization. Yale University Press, New Haven, Connecticut.

Raff, R. 1999. Larval homologies and radical evolutionary changes in early development, Pp. 110–121 in Homology. Novartis Symposium, vol. 222. John Wiley & Sons, Chichester, United Kingdom.

Reidhaar-Olson, J., & R. Sauer. 1990. Functionally acceptable solutions in two alpha-helical regions of lambda repressor.—Proteins, Structure, Function, and Genetics, 7:306–316.

Rutten, M. G. 1971. The origin of life by natural causes. Elsevier, Amsterdam.

Sapp, J. 1987. Beyond the gene. Oxford University Press, New York.

Sarkar, S. 1996. Biological information: a skeptical look at some central dogmas of molecular biology. Pp. 187–233 in S. Sarkar, ed., The philosophy and history of molecular biology: new perspectives. Kluwer Academic Publishers, Dordrecht.

Schützenberger, M. 1967. Algorithms and the neo-Darwinian theory of evolution. Pp. 73–75 in P. S. Morehead and M. M. Kaplan, eds., Mathematical challenges to the Darwinian interpretation of evolution. Wistar Institute Symposium Monograph. Allen R. Liss, New York.

Shannon, C. 1948. A mathematical theory of communication.—Bell System Technical Journal 27:379–423, 623–656.

Shu, D. G., H. L. Lou, S. Conway Morris, X. L. Zhang, S. X. Hu, L. Chen, J. Han, M. Zhu, Y. Li, & L. Z. Chen. 1999. Lower Cambrian vertebrates from south China.—Nature 402:42–46.

Shubin, N. H., & C. R. Marshall. 2000. Fossils, genes, and the origin of novelty. Pp. 324–340 in Deep time. The Paleontological Society.

Simpson, G. 1970. Uniformitarianism: an inquiry into principle, theory, and method in geohistory and biohistory. Pp. 43–96 in M. K. Hecht and W. C. Steere, eds., Essays in evolution and genetics in honor of Theodosius Dobzhansky. Appleton-Century-Crofts, New York.

Sober, E. 2000. The philosophy of biology, 2nd edition. Westview Press, San Francisco.

Sonneborn, T. M. 1970. Determination, development, and inheritance of the structure of the cell cortex. In Symposia of the International Society for Cell Biology 9:1–13.

Solé, R. V., P. Fernández, & S. A. Kauffman. 2003. Adaptive walks in a gene network model of morphogenesis: insight into the Cambrian explosion.—International Journal of Developmental Biology 47(7–8):685–693.

Stadler, B. M. R., P. F. Stadler, G. P. Wagner, & W. Fontana. 2001. The topology of the possible: formal spaces underlying patterns of evolutionary change.—Journal of Theoretical Biology 213:241–274.

Steiner, M., & R. Reitner. 2001. Evidence of organic structures in Ediacara-type fossils and associated microbial mats.—Geology 29(12):1119–1122.

Taylor, S. V., K. U. Walter, P. Kast, & D. Hilvert. 2001. Searching sequence space for protein catalysts.—Proceedings of the National Academy of Sciences, U.S.A. 98:10596–10601.

Thaxton, C. B., W. L. Bradley, & R. L. Olsen. 1992. The mystery of life's origin: reassessing current theories. Lewis and Stanley, Dallas, Texas.

Thompson, D. W. 1942. On growth and form, 2nd edition. Cambridge University Press, Cambridge, United Kingdom.

Thomson, K. S. 1992. Macroevolution: The morpho-

logical problem.—American Zoologist 32:106–
112.

Valentine, J. W. 1995. Late Precambrian bilaterians:
grades and clades. Pp. 87–107 *in* W. M. Fitch
and F.J. Ayala, eds., Tempo and mode in evo-
lution: genetics and paleontology 50 years after
Simpson. National Academy Press, Washington,
D.C.

———. 2004. On the origin of phyla. University of
Chicago Press, Chicago, Illinois.

———, & D. H. Erwin. 1987. Interpreting great de-
velopmental experiments: the fossil record. Pp.
71–107 *in* R. A. Raff and E. C. Raff, eds., De-
velopment as an evolutionary process. Alan R.
Liss, New York.

———, & D. Jablonski. 2003. Morphological and de-
velopmental macroevolution: a paleontological
perspective.—International Journal of Devel-
opmental Biology 47:517–522.

Wagner, G. P. 2001. What is the promise of develop-
mental evolution? Part II: A causal explanation
of evolutionary innovations may be impossi-
ble.—Journal of Experimental Zoology (Mol.
Dev. Evol.) 291:305–309.

———, & P. F. Stadler. 2003. Quasi-independence, ho-
mology and the unity of type: a topological the-
ory of characters.—Journal of Theoretical Bi-
ology 220:505–527.

Webster, G., & B. Goodwin. 1984. A structuralist ap-

proach to morphology.—Rivista di Biologia 77:
503–10.

———, & ———. 1996. Form and transformation:
generative and relational principles in biology.
Cambridge University Press, Cambridge, Unit-
ed Kingdom.

Weiner, J. 1994. The beak of the finch. Vintage Books,
New York.

Willmer, P. 1990. Invertebrate relationships: patterns in
animal evolution. Cambridge University Press,
Cambridge, United Kingdom.

———. 2003. Convergence and homoplasy in the
evolution of organismal form. Pp. 33–50 *in* G.
B. Müller and S. A. Newman, eds., Origination
of organismal form: beyond the gene in devel-
opmental and evolutionary biology. The M.I.T.
Press, Cambridge, Massachusetts.

Woese, C. 1998. The universal ancestor.—Proceedings
of the National Academy of Sciences, U.S.A.
95:6854–6859.

Wray, G. A., J. S. Levinton, & L. H. Shapiro. 1996.
Molecular evidence for deep Precambrian di-
vergences among metazoan phyla.—Science
274:568–573.

Yockey, H. P. 1978. A calculation of the probability of
spontaneous biogenesis by information theo-
ry.—Journal of Theoretical Biology 67:377–
398.

———. 1992. Information theory and molecular bi-
ology. Cambridge University Press, Cambridge,
United Kingdom.