Stephen Freeland

Article

The Evolutionary Origins of Genetic Information

Stephen Freeland

Any living branch of science achieves progress by testing new ideas. The results of these tests determine whether each new idea is accepted as a change to what we thought we knew, is dismissed as incorrect, or simply stagnates, owing to a lack of clear evidence. For evolutionary theory, one such proposition is that some features of genetic information cannot evolve through natural processes unless we allow a role for an intelligent designer. This proposition claims testability by defining information in a way that is usually reserved for human creations, such as computer programming code. The argument is that since we know that intelligent beings create computer code, then perhaps similar features found within genetic information indicate a similar origin. However, many biologists perceive that they are able to understand exactly where life's genetic information comes from (the local environment) by thinking in terms of more fundamental and well-established definitions of information that do not involve intelligent design.

Current science does not have a detailed, widely accepted description for how a genetic information system evolved in the first place. Intelligent design (ID) proponents suggest that this is a key weakness of existing evolutionary theory, consistent with the need for an intelligent designer. I describe the progress that mainstream science has made toward understanding the origin of genetic information ever since the molecular basis of genetic information was first understood, encouraging readers to reach their own conclusions.

Biological evolution describes the natural process that transfers information from a local environment into the chemical known as DNA. Something similar happens when gravity causes raindrops to form a puddle, and the shape of the ground beneath becomes reflected in the underside of the water.

This unusual definition of evolution seeks to clarify an ambiguity in traditional alternatives, such as "biological evolution is a natural process of change in genetic material over time."[1] The phrase "change in genetic material" describes and limits exactly what scientists measure and test to develop their evolutionary theory; however, any description of this sort omits two aspects of a living science. One is the group of all propositions that have been revealed as incorrect through tests (such as *recapitulation*—the claim that the embryos reenact their evolutionary history as they develop from a single fertilized egg cell).[2] Let us call these incorrect propositions "Category 1" omissions. Knowing about them can help scientists avoid wasted time spent repeating previous errors.

ASA member **Stephen Freeland** *is currently project manager for the NASA Astrobiology Institute research team based at the University of Hawaii. Steve's academic background (a bachelor's degree in zoology from Oxford, a master's in biological computation from York University, and a doctorate in genetics from Cambridge) has led him to spend the past twelve years researching the evolution of genetic coding. Steve's current research explores the evolution of the amino acid "alphabet"—the set of twenty building blocks with which life has been making the proteins of metabolism for more than three billion years.*

Underlying this research is a growing interest in the cosmological question, "To what degree is life on Earth (or elsewhere) a result of chance?" As the son of a biology teacher who retrained as a Methodist minister, Steve has been blessed with an encouraging environment with which to explore the interface of science and faith since childhood.
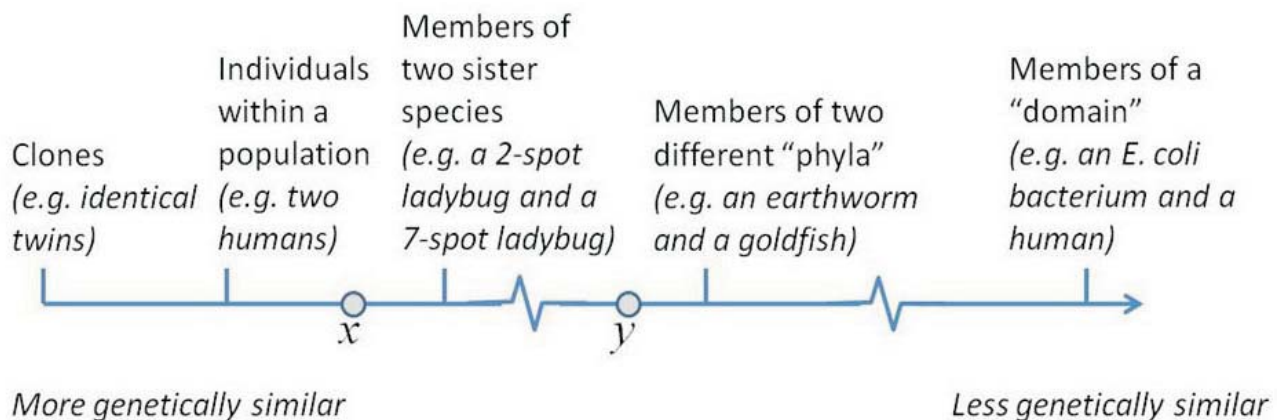
The second element missing from a classic definition comprises all propositions for which science has yet to find clear evidence, for or against. We may refer to these as "Category 2" omissions. Propositions in this second category are especially important to science because all suggestions to change existing scientific understanding start here. In other words, Category 2 propositions can gather supporting evidence until they become accepted as scientific truth, altering what we previously thought we understood, perhaps even requiring a change in definition of that science. (It is both humbling and inspiring to remember that scientific knowledge is presently incomplete in ways that are actively misleading us.) However, many Category 2 propositions follow a different trajectory as careful application of the scientific method reveals them as incorrect, and therefore reclassifies them as Category 1 propositions. A third fate is possible for Category 2 propositions. If they do not generate sufficient evidence to make a clear case, whether it be for or against, then they will stagnate. A proposition often ends in stagnation if it fails to generate clear, testable hypotheses that have the power to transform established theory.

Intelligent design theory (ID) has already started its life in Category 2 by suggesting that current evolutionary theory cannot adequately explain the origin of new genetic information. The unusual definition of evolution written above hints why many scientists, including Christians such as myself, think this is an incorrect (Category 1) proposition. What follows seeks to explain why in greater detail—and to equip you to judge for yourself.

Evaluating Suggestions for Changes to Evolutionary Theory

Start by imagining a line that describes every conceivable degree of genetic difference that could separate any two living organisms (figure 1). In fact, we do not have to rely on imagination—such differences can be measured precisely, due to life's shared biochemistry of DNA and proteins (see box 1). Most criticisms of evolution are, upon careful inspection, claims that evolutionary theory is incomplete. They suggest that evolutionary theory can explain differences only up to a specific point on this line. For example, older versions of creationism claim natural processes cannot change anything more than the frequency (number of copies) of genetic material already present within a species. In effect, this defines a point *X* on the line shown in figure 1. To the right of *X* lie larger differences in genetic material, such as those that separate different species. Under creationism, these differences are considered too large for natural processes to explain, and are therefore explained by divine intervention.

A growing weight of detailed evidence shows that new species form by the accumulation of changing gene frequencies within a population.[3] This evidence has led many contemporary versions of creationism to increase the acceptable limit for evolution, moving point X on the line in figure 1 to point *Y*. An explanation is that God created fundamental *kinds* of animals and plants so that the formation of new species *within these kinds* are legitimate outcomes of natural processes.[4] Accepting this interpretation, it is now the larger degrees of genetic difference lying to the right of *Y* that require supernatural explanation.

Figure 1. Any two or more organisms can be compared for genetic similarity (e.g., in terms of differences in DNA sequence), and thus plotted as a point on a line that runs from "complete genetic similarity" (clones or identical twins) to "very little genetic similarity," such as a human and an *E. coli* bacterium.

For our purposes, what matters is that different versions of creationism all accept some degree of evolution but place a cutoff on the extent of change that evolution can produce, explaining anything above that point by divine intervention. Wherever the cutoff is perceived, the same terminology is used: *microevolution* (anything to the left of the acceptable limit) is attributable to natural processes, but *macroevolution* (anything to the right of this point) requires a new explanation—direct creation by God.

The terms "*microevolution*" and "*macroevolution*" come originally from similar suggestions made within secular science during the early development of evolutionary theory.[5] Biologists working early in the twentieth century were learning how to cause genetic mutations in a laboratory setting. These mutations could, in a single generation, produce large changes in an organism's appearance. Some pioneers of this new science (*genetics*) thought that their discoveries changed evolutionary theory. Darwin had previously described a process of evolution

Box 1. An Introduction to Biological Coding and the Central Dogma of Molecular Biology

A *code* is a system of rules for converting information of one representation into another. For example, Morse Code describes the conversion of information represented by a simple alphabet of dots and dashes to another, more complex alphabet of letters, numbers, and punctuation. The code itself is the system of rules that connects these two representations. Genetic coding involves much the same principles, and it is remarkably uniform throughout life (figure 2): genetic information is stored in the form of *nucleic acid* (DNA and RNA), but organisms are built by (and to a large extent from) interacting networks of *proteins*. Proteins and nucleic acids are utterly different types of molecule; thus it is only by decoding genes into proteins that self-replicating organisms come into being, exposing genetic material to evolution. The decoding process occurs in two distinct stages: during *transcription* local portions of the DNA double-helix are unwound to expose individual genes as templates from which temporary copies are made (*transcribed*) in the chemical sister language RNA. These *messenger RNA* molecules (mRNA's) are then *translated* into protein.

The language-based terminology reflects the fact that both genes and proteins are essentially 1-dimensional arrays of chemical letters. However, the nucleic acid alphabet comprises just four chemical letters (the 4 nucleotides are often abbreviated to A, C, G, and T—but see endnote 28), whereas proteins are built from twenty different amino acids. Clearly, no 1:1 mapping can connect nucleotides to amino acids. Instead nucleotides are translated as nonoverlapping triplets known as *codons*. With four chemical letters grouped into codons of length 3, there are 4x4x4 = 64 possible codons. Each of these 64 codons is assigned to exactly one of 21 meanings (20 amino acids and one "stop translation" signal found at the end of every gene). The genetic code is quite simply the mapping of codons to amino acid meanings (figure 2a). One consequence of this mapping is that most of the amino acids are specified by more than one codon: this is commonly referred to as the *redundancy* of the code.
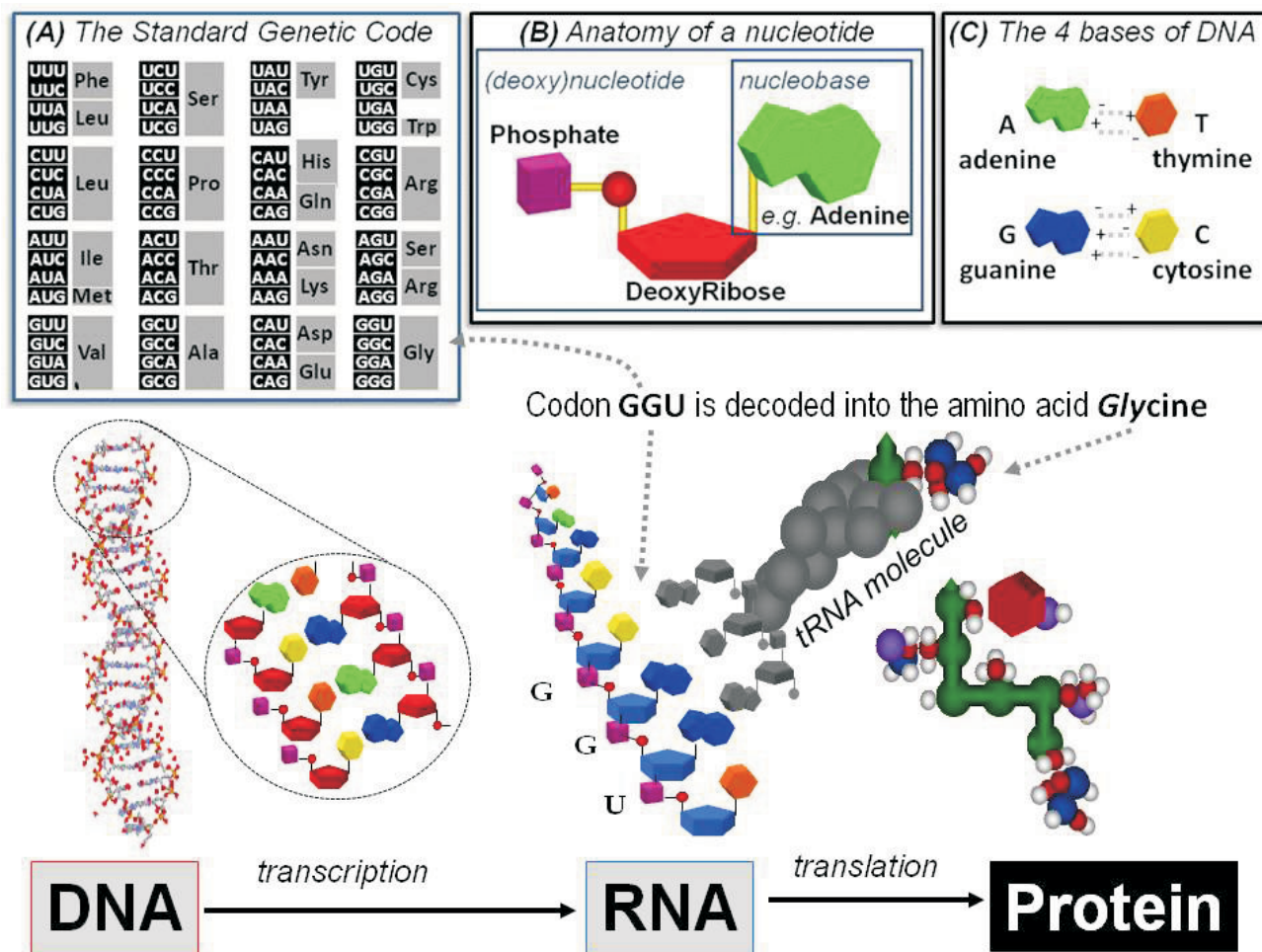
Although the molecular machinery that produces genetic coding is complex (and indeed, less than perfectly understood), the most essential elements for this discussion are the *tRNA*'s and *ribosome*. Each organism uses a set of slightly different tRNA's that each bind a specific amino acid at one end, and recognize a specific codon or subset of codons at the other. As translation of a gene proceeds, appropriate tRNA's bind to successive codons, bringing the desired sequence of amino acids into close, linear proximity where they are chemically linked to form a protein *translation product*. In this sense, tRNA's are adaptors and translators—between them, they represent the molecular basis of genetic coding. The ribosome is a much larger molecule, comprising both RNA and various proteins, which supervises the whole process of translation. It contains a tunnel through which the ribbon of messenger RNA feeds; somewhere near the center of the ribosome, a window exposes just enough genetic material for tRNA's to compete with each other to bind the exposed codons.

by natural selection, and this process could be observed changing the frequencies of genes within populations over one or more generations. However, subtle differences in the genetic makeup of a population seemed too small to connect with the large jumps being witnessed in laboratories, and the latter seemed more relevant to the formation of new species. A typical evolutionary debate from this time also defined a point somewhere near *X* on the line shown in figure 1. Everyone agreed that Darwin's

process could explain changes to the left of this point (microevolution), but some now argued that a fundamentally new phenomenon called *genetic mutation* or *macromutation* was responsible for the larger-scale differences to the right (macroevolution).

At first sight, *macromutationism* and *creationism* seem similar. Both propose a cutoff point for the degree of genetic change that evolutionary theory can explain, and both propose that a new cause must be

Figure 2. The Central Dogma of Molecular Biology. Short stretches of DNA *nucleotides* are copied into the chemical sister language RNA, and these are then decoded into *amino acids*. A triplet of RNA nucleotides (*codon*) corresponds to a single amino acid. Many amino acids linked together form a protein—the basic unit of metabolism. The decoding molecules are *transfer RNA's* (tRNA's), which operate within the context of the ribosome (not shown here). **(A)** The *Standard Genetic Code* is the set of rules defining which codon each amino acid is decoded into. This used to be known as the Universal Genetic Code, but since the 1970s many slight variants have been discovered, all of which diverged from the standard genetic code. **(B)** The building block of DNA is a *nucleotide*. This is not to be confused with a *nucleobase*—nucleobases must be joined to a ribose molecule and a phosphate to become a nucleotide. This distinction is important to the origin of genetic information because although amino acids, bases, and even ribose are widely thought to be produced by non-biological processes throughout the universe, nucleotides are not so (see endnote 48). **(C)** The four bases of DNA are adenine, thymine, guanine, and cytosine. When placed edge to edge, each base finds a matching counterpart in one of the others. This explains why DNA is so stable as a double helix: two strands of nucleotides run in opposite directions, allowing hundreds of bases to match in pairs, like the teeth on a zipper. RNA uses the same four nucleobases, except for Thymine: here RNA uses a chemical variation known as Uracil. Thymine is more chemically stable than Uracil, fitting with a general picture that DNA evolved to carry information with greater stability than RNA once the latter had "subcontracted" the work of chemical catalysis to proteins.

added to explain genetic differences beyond this cut-off. Where the two propositions differ for science is in their potential for tests. Supernatural causes (literally, those that come from beyond nature) cannot be tested directly from within the natural universe. Science can get no nearer than searching for indirect evidence, such as natural phenomena that cannot be explained by any known, natural cause. Evidence of this kind is unlikely to carry creationist propositions from Category 2 suggestions into accepted science. In part, the problem is that specific data used to justify unnatural causes tend to find an equal or better explanation in terms of the natural causes measured by science as new data become available.[6] Mostly, however, the problem is that unnatural phenomena can never be more than *consistent with* a supernatural cause. Even where specific claims for unnatural phenomena have not been refuted, it remains equally possible that science has yet to understand natural causation, and science keeps growing its understanding in ways that support evolution.[7]

In contrast to creationism, the work of the early geneticists referred to strictly natural phenomena (i.e., those occurring within the observable, natural universe). This focus allowed for direct evaluation by science. Through a series of hypotheses and tests, geneticists revealed that early examples of laboratory-induced macromutation were, in fact, large-scale genetic damage caused by powerful doses of radiation and chemicals. Meanwhile, other tests clarified that within nature, genetic mutations of far greater subtlety do indeed account for the minor differences between members of a species (micro-evolution). Further evidence indicated that micro-evolution accumulates over time to account for all larger degrees of evolutionary diversification (macro-evolution). In other words, science not only failed to find supporting evidence for the idea that macro-mutations are responsible for the emergence of new species, but it also undermined the observation that had led to this hypothesis in the first place. Science refuted the claim that macromutations filled a gap within evolutionary theory by discovering that there was no gap to fill. Macromutationist ideas for the origin of new species have therefore moved from Category 2 (ideas for which the evidence is unclear) to Category 1 (ideas that are incorrect), and they are no longer actively researched by evolutionary biolo-

gists.[8] The most noticeable place where remnants of macromutationist ideas are to be found today is within popular culture, in which characters ranging from Spiderman to X-men are stubbornly explained in terms of these outdated views of evolution.

Over the years, secular science has proposed many other novel factors that evolutionary theory should absorb to better explain biological diversity. So far, all have gone the way of macromutationism.[9] However, cutting-edge research is, by definition, constantly probing for evidence to support new insights. For example, one recent claim is that without adding any new causal factors, enough biological evolution will ultimately produce something similar to our own sentient species.[10] Contrary to popular belief, this outcome is not predicted by current evolutionary science.[11] The new claim of *inevitable outcomes* has not been refuted by science, nor has the supporting evidence become overwhelming. In fact, scientists still do not know quite how to weigh the evidence—how to measure inevitability when it comes to evolution. As a result, *inevitable outcomes* remains a Category 2 idea, a topic of active debate and research until scientists gather a clear majority of evidence either to reject it or to accept it into science.[12] If such evidence is not forthcoming, the idea will likely atrophy.

These three propositions, *creationism, macromutationism,* and *inevitable outcomes,* provide context for discussing another idea that has arisen in Category 2: the idea that evolutionary theory would be improved by allowing a role for a guiding intelligence. Nothing is inherently unscientific about this suggestion so long as it can find appropriate evidence (through tests) to help scientists decide, one way or the other. One idea for a test is to ask whether we can identify properties of genetic information that resemble human-created information. The idea is that we are intelligent, so if genetic material looks like the sort of thing we would make, then it might be better explained as the product of intelligent design, especially if science can identify features of genetic information inexplicable by known evolutionary processes.[13] ID names one of these features *specified complexity.* Specified complexity is a measurement that tries to capture the semantic content of information (the amount of *meaning* within a piece of information). The assertion is that natural processes,

lacking a guiding intelligence, can neither produce new genetic information nor can they explain the origin of genetic information because this implies an increase in specified complexity. Each of these claims warrants careful consideration.

Can Natural Processes Generate New Genetic Information?

Unless life began in greater quantity than now exists, evolution requires that natural processes have, over time, increased the total quantity of genetic material (DNA) present on our planet. One way in which science currently believes that genetic information has increased over time is that a natural process has increased the number of copies of DNA molecules without any need for guidance by an intelligent agent. This kind of increase in genetic information is exactly what we see whenever a natural population grows (e.g., bacteria during an infection). Clearly, this type of information increase is not at issue. Indeed, ID refers to this as a *flow of information*, rather than as the *creation of new information*.[14]

Along similar lines, unless life originated containing more DNA than the most genetically complex organism alive today, then some lineages must have increased the quantity of DNA they contain through evolution.[15] Established science knows ways to observe and measure this kind of increase in genetic information. For example, genome-sequencing technology has revealed small variations in the length of genetic material carried by different individuals within every natural population, including our own species.[16] *Indels* (short for "insertion/deletion mutations") form one of the fundamental types of mutation recognized by geneticists. Indels represent microevolution, but why could insertions not accumulate faster than deletions over time, causing genetic material to grow in size? This is exactly what we would expect if microevolution adds up to produce macroevolution. Again, ID agrees with mainstream science that this is entirely within the realm of causation by existing theory and that a focus on *quantities* of DNA is misleading. Genetic differences between a human and an amoeba are only partly attributable to the different quantity of genetic material present in each. For example, the *Amoeba proteus* genome contains 100-fold more DNA than a human genome; other species of Amoebae contain both

much larger and much smaller quantities of genetic information.[17]

More important than the quantity of DNA present in each species is the different order in which nucleotides are linked together to spell out genetic messages. DNA has the unusual property of being aperiodic. This means that the sequence of nucleotides within a DNA molecule is not constrained to any kind of repeating pattern (see Box 1). It is precisely this property that allows DNA or anything with similar properties to carry a large amount of information. For example, written English is an aperiodic sequence built from relatively few symbols. Everything ever written in English can be copied using one simple keyboard. The trick is to arrange these building-block symbols into particular aperiodic sequences. The major difference between this article and *Harry Potter* lies not in the quantity of letters and punctuation used but in the sequence in which these symbols have been assembled. Where current evolutionary science disagrees with ID is in the suggestion that some sequences of genetic information can only be generated by a guiding intelligence.[18] ID asserts that natural processes cannot produce changes in genetic information if these changes correspond to an increase in specified complexity. Specified complexity is measured in a way that tries to capture the difference that separates this article from *Harry Potter*. More accurately, specified complexity is the information that distinguishes any random sequence of symbols from orderings that have meaning.[19]

The idea that some sequences of DNA cannot be produced by natural processes, owing to the information they contain, has no empirical support from modern genetics. In fact, quite the reverse is true. Genetic information is stored in sequences of nucleotides that have been chemically linked together to form a molecule of DNA. Genetics, bioinformatics, biochemistry, and molecular biology all agree that natural processes can cause any nucleotide to become the neighbor of any other within a DNA sequence. Mutations that interconvert each of the four nucleotides have been observed within natural populations and within the laboratory, as have insertions, deletions, and translocations of minisequences from one region of the DNA sequence to another. These elementary components of modern genetics are, in principle, more than sufficient to produce any

DNA sequence from any other. Try this for yourself by listing a series of mutations that convert the word "*evolution*" into "*creation*" with the restriction that each mutation must either change a single letter, insert or delete one or more letters, or move the position of any subgroup of letters. There are many ways to reach the outcome, and this remains true for any two words that you can choose.[20]

The biochemistry that describes how genetic information is stored, replicated, corrected, and translated into proteins is fascinating but requires no novel concepts regarding semantic information. The question is whether the addition of this latter concept can reveal insights, such as limitations too subtle to observe with empirical science. As the companion article by Randy Isaac explains, science recognizes several types of information.[21] One of the most fundamental types is thermodynamic information, a fundamental parameter of physics that reflects all that could be different about the universe. If evolutionary theory implies an increase or loss in thermodynamic information, then it would be in conflict with established ideas belonging to another branch of science. This is not the case. Nothing about biological evolution ever involves an increase (or decrease) in the thermodynamic information present within the universe. Indeed, evolution can be described precisely in terms of thermodynamic processes by which sources of energy bring into being particular states of information within a DNA molecule. The opening definition of this article tries to emphasize this point: "*Biological evolution describes a natural process that transfers information from a local environment into the chemical known as DNA.*" To understand why this causes many biologists to doubt whether additional concepts regarding information are necessary or helpful, one must return to Darwin's original insight.

Within a population of individuals that vary from one another, those that best match their environment will, on average, leave behind the most offspring. Wherever the match is genetically programmed, the version of the genetic program associated with the best match will tend to increase in frequency over time by leaving behind more copies of itself. As these advantageous versions are copied from one generation to the next, they will mix with new variations that either increase or decrease the match. All the while, the environment keeps changing and mutations keep occurring, and thus the matching process

continues. Repeating this process over and over will create a pool of genetic programs that have accumulated variations, maximizing the overall match between organism and environment (quite simply because those that did not match as well left behind fewer copies of themselves).[22]

Through this process, genetic material will evolve to mirror some of the information presented by the environment in which it is copying itself. This information might include patterns in time and space by which ambient temperatures vary, or patterns of chemical resources found in the environment. Things get especially interesting when we realize that some of the most significant information about an organism's environment is specified by other organisms. The color of leaf on which an organism feeds may become reflected in its genetic material, if this type of genetic programming helps the herbivore to hide from predators; conversely, genetic material may evolve to program colorations that contrast with the background of other organisms in an environment where finding and attracting mates is the strategy that leaves behind the most copies. Each reflection originates in physical parameters, but these collide, transfer information, and start new emanations as they become reflected in the genetic material of the organisms. No matter how complex these rebounding reflections of the environment become, they will never create new information (any more than your image in a reflection of a reflection of a reflection contains more information than you do).[23] Viewed in this light, biological evolution is a natural process that distills thermodynamic information from a highly complex environment into molecules of DNA.[24]

Evolution is to DNA what gravity is to a puddle of water. In both cases, it is possible to isolate elements of the whole that carry impressively complex information (species really do contain lots of complex genetic programs written out in DNA, as does the shape produced when a body of H_2O perfectly matches some of the information inherent to the collection of rocks and debris beneath). If we considered only the water, we might be tempted to think that some sort of intelligence had sculpted such a complex and accurate reflection of the environment. We might even measure this information content to demonstrate its improbability of arising by chance. But step back far enough to see the whole

picture, and we realize that evidence consistent with design can be better understood as a result of natural processes (gravity and a preexisting, information-rich environment). In the case of biological evolution, evolution and DNA take the place of gravity and water. Gravity and evolution not only permit the transfer of environmental information into a chemical medium, but inevitably and inexorably lead to this information transfer. Given this understanding, it is hard to see what evolutionary science would gain by accepting other concepts of semantic information that create a problem to be solved by invoking an indeterminate intelligent designer.

Can Natural Processes Account for the Origin of Genetic Information?

The description of evolution given above applies once the world contains a genetic material that can influence its own rate of copying by reflecting the environment. In living systems, these remarkable properties are produced by the *central dogma of molecular biology* (DNA, proteins, and the genetic code that allows the former to specify the latter; see box 1 and figure 2). Perhaps a stronger argument for ID is that no natural process could create such a versatile system in the first place.

It is true that, at present, evolutionary science does not have a clear, detailed, and well-accepted explanation for how the central dogma of molecular biology emerged. But does that mean it is time to embrace ID as a better approach? By analogy, current medical science has not found the cure for cancer. Taken in isolation, this sound bite could lead to the misleading view that existing research directions, developed for decades, are best written off as a failure. This would miss an important context. Many aspects of cancer are now being treated with far greater effectiveness than ever before as a result of ongoing research. However, these cures are not robust (all-encompassing) enough to be summarized in the statement, "we have found the cure for cancer." This status is typical of big questions within science: failure to reach the sound-bite goal should not be mistaken for evidence that the research program has failed. Scientific progress is measured by the insights that research produces, and their implications for where we might usefully look next. These

insights may even open up new awareness of just how much we do not understand, but characterizing the past few decades of cancer research as an exhaustive search that has ended in failure would be more than premature: it would be actively misleading. This final section of the article offers context to help the reader judge whether a similar situation holds for current research into natural processes that explain the origin of genetic information.

Let us start by making entirely clear what scientists are looking for. As the previous section explains, the challenge is not to find a natural process that can create enough information for a simple genetic system. The universe is replete with information capacity and syntax—from the positions of stars within our galaxy (and billions of others) to the arrangement of atoms in a single grain of sand. Within living systems, most of this information is ignored—so the question is not, "where did the information come from?" (unless we wish to talk cosmology—a very different subject) but rather, "how does nature create systems that focus on some of this natural information?" Put another way, the challenge for understanding the origin of genetic systems is to find how natural processes can simplify a large amount of thermodynamic information into a syntax that displays only the disciplined chemical semantics of a self-replicator.

The exact details of life's genetic information system came into focus during the middle of the twentieth century.[25] In 1953, Watson and Crick published the structure of DNA,[26] revealing the innate capacity of this molecule to replicate and evolve indefinitely. Thirteen years later, a consortium of scientists published the details of the genetic code by which the information carried by DNA is translated into specific protein sequences.[27] The system was so fundamental to understanding life, yet so simple and easy to explain that it has become known as the *central dogma of molecular biology* (box 1 and figure 2). However, it was puzzling from an evolutionary perspective. Protein catalysts supervise the construction of individual nucleotides (the building blocks for making DNA and RNA). Other proteins link these nucleotides into DNA or RNA sequences, depending on their type (deoxyribonucleotides into DNA, and ribonucleotides into RNA). Proteins can perform these roles because each one has just the right chemical properties to catalyze a specific chemical reaction

(such as linking a molecule of the nucleotide "A" to T, G, or C to start building a genetic message).[28] Each protein is a long chain of amino acids (typically several hundred) that have been chemically linked together. The function and shape of a protein emerge spontaneously according to the sequence of these amino acids—just as the meaning of a word is carried (for us) by a sequence of letters drawn from the English alphabet.[29] The only way to reliably build the right sequence(s) of amino acids to make the proteins of metabolism is to follow genetic instructions, one code-word (*codon*) at a time. In other words, for more than 3 billion years, everything living has needed proteins to make genetic information—and needed genetic information to specify how these proteins are to be made.

At the time of discovery, this system looked like something that ID proponents might call *irreducibly complex*: an irreducibly complex system is one that cannot evolve from simpler precursors, because any simplification would lose the entire functional value of the system. This perception of an unevolvable code was further enhanced by the discovery that the same exact genetic code is at work in organisms as different as human beings and *E. coli* bacteria. (Refer back to figure 1. This is about as genetically different as living organisms can be!) Scientists at the time came to think that one genetic code was universal for all living systems on our planet. This led Francis Crick to propose that the genetic code is a "*frozen accident*" of evolution,[30] universal across life precisely because once it had formed (by some unknown event), it was so fundamental to all biochemistry that it could never change again. Specifically, he pointed out that any change to the rules of genetic coding would be equivalent to a simultaneous mutation in every single gene in the organism (box 1).[31] While evolutionary theory requires that occasional small mutations produce a better fit to the environment, the simultaneous mutation of thousands of genes seems extreme even by the standards of macromutationism. However, subsequent science has developed at least three major lines of research that undermine the concept of a frozen accident (and irreducible complexity) for genetic coding.[32]
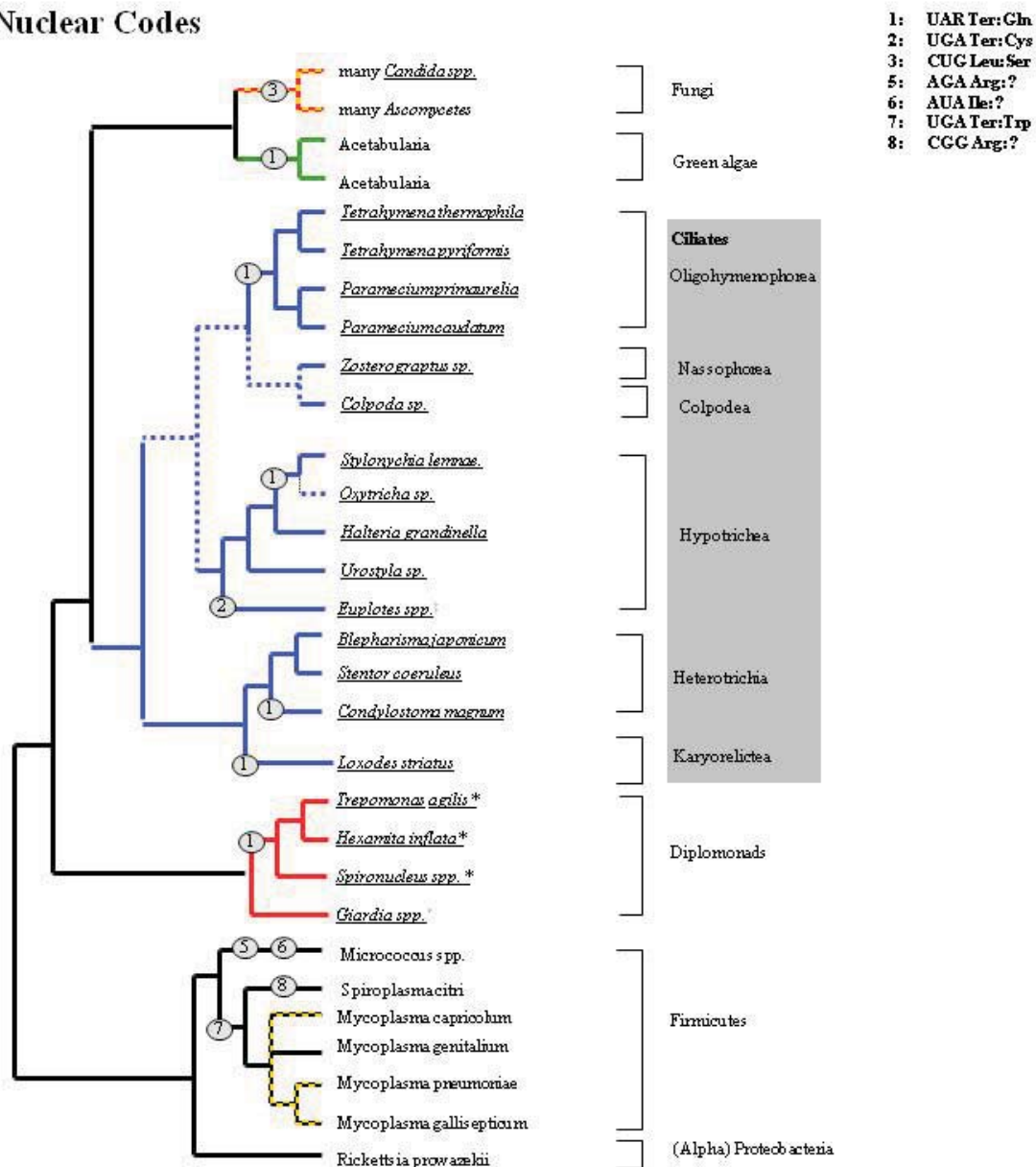
First, it has been discovered that the genetic code is not universal. Around a dozen or so minor variations exist.[33] These variations are mostly codes in which one or more genetic codons have altered their

amino acid "meanings." Some involve a more significant change—the addition of a twenty-first or twenty-second amino acid.[34] Everything indicates that these genetic codes evolved from the standard genetic code during the past few hundred million years, and continue to evolve today.[35] Arguments for the evolvability of the code are strengthened by the finding that amino acids are assigned to genetic code-words nonrandomly. In particular, codons are assigned to amino acids in such a pattern that common mutations produce minor variations as proteins are decoded. A growing body of evidence connects this feature of the code to the idea that considerable evolution by natural selection had gone into shaping this system.[36] Everything suggests that the genetic code is evolved and evolvable after all.[37]

The second major insight into the origins of genetic coding is that multiple, independent lines of evidence suggest that the standard amino acid *alphabet* of twenty building blocks grew from a smaller earlier alphabet corresponding to an earlier stage in genetic code evolution. Many variations have been proposed.[38] Most derive their views by considering only one or two types of evidence: sophisticated calculations of the amino acid sequences of truly ancient proteins, the repertoire of amino acids found in meteorites; simulations of an early, prebiological planet Earth; and so on. What is interesting is an unlooked for match between the broad findings of these different approaches. In particular, different approaches end up dividing the twenty amino acids of modern organisms into ten that were around in the earliest systems, and ten that arrived later, as by-products of early biological evolution. The members of each group are remarkably consistent,[39] hinting directly at the process by which the genetic code evolved, growing more complex over time from simpler beginnings. Recent findings are also starting to make sense of why natural selection created this particular alphabet of building blocks.[40]

The third line of insight takes us backwards to the possible origins of genetic coding. Some scientists have used the SELEX approach that is described in a companion article by Jonathan Watts to define mini-sequences of RNA that specifically bind to a particular amino acid.[41] Although results have been patchy, some amino acids seem to associate with surprising choosiness to the code-words assigned to them in the standard genetic code. This association

Figure 3. Some of the nonstandard genetic codes that have been discovered since the time of the frozen accident hypothesis, together with their evolutionary relationships (adapted from R. D. Knight, S. J. Freeland, and L. F. Landweber, "Rewiring the Keyboard: Evolvability of the Genetic Code," *Nature Reviews Genetics* 2 [2001]: 49–58). Recent work that has examined the process by which genetic coding evolves is reviewed by G. R. Moura, J. A. Paredes, and M. A. Santos, "Development of the Genetic Code: Insights from a Fungal Codon Reassignment," *FEBS Letters* 584 (2010): 334–41.

suggests that the earliest steps in genetic coding may have been nothing more than simple physical affinities between two types of chemical.

Between them, these insights represent significant progress from the impossibly self-referential system viewed by Crick and those around him just fifty years ago. This half-century of research indicates that the standard genetic code at work in modern cells may be a product of substantial evolution that had taken place by around 3 billion years ago. But perhaps the most interesting progress is that few scientists still regard the emergence of life's central dogma as the origin for genetic information.

The Deepest Origins of Genetic Information

The observation that RNA sequences can bind amino acids hints at something very important: proteins are not the only type of molecule that can spontaneously fold into shapes with interesting properties. As described in the companion article by Watts, sequences of RNA can exhibit protein-like behavior.[42] Technologies first developed in the 1980s and 1990s have been used to lab-evolve a wide variety of molecules, dubbed *ribozymes* in deference to the previously known class of protein catalysts known as enzymes. These ribozymes now cover most steps of fundamental biochemistry (such as linking together carbon atoms to make important biological molecules).

Proteins are much less necessary for life than they seemed a couple of decades ago. This observation finds unlooked-for synergy with another line of scientific discoveries. In modern living systems, not all RNA performs the simple role of carrying genetic information from DNA to be decoded into proteins. A handful of the genes that are faithfully copied from DNA into RNA fold up into complex three-dimensional shapes that act as if they were proteins. Interestingly, these natural ribozymes tend to occur in the most ancient metabolic pathways—those shared by bacteria, humans, and everything else alive today. Aspects of biology that have not changed much in billions of years of evolution are likely still with us because they have been doing their job very well throughout this period. In other words, this type of RNA behaving like a protein is exactly what one might expect to see if the ribozymes produced

by SELEX resemble a stage of our truly ancient evolutionary past when genetic coding of proteins was far less important (if it was present at all).

Oddly enough, Crick (of the frozen accident) had suggested something similar to this concept of *molecular fossils* when he looked at how genetic decoding works. He noticed that the *adaptor molecules* responsible for decoding individual genetic code-words into specific amino acids are nothing more than folded-up RNA. He also noticed that the biggest and most complex molecular machine involved with genetic decoding (the *ribosome*) seemed to be made of RNA with a few proteins thrown in for good measure. Three decades later, new technology allowed researchers enough precision in their study of the ribosome's structure to confirm that this is correct: although proteins are embedded within the tangled, folded RNA, they appear to offer little more than structural enhancements.[43] At its core, the *ribosome* is a *ribozyme*. It seems likely that a primitive ribosome could function without any encoded proteins, exactly what we would expect if genetically encoded proteins emerged from a simpler, earlier world in which only RNA existed.

Of equal interest, everything points toward DNA being the last arrival out of the three fundamental biomolecules: DNA, RNA, and protein.[44] DNA is made by complex, genetically encoded protein enzymes without a ribozyme in sight. The individual building blocks of DNA (*deoxynucleotides*) are made by taking and modifying a nucleotide of RNA. Again, all this is exactly what we would expect if DNA evolved from RNA, after genetically encoded proteins had already entered the picture. Indeed, DNA is a more chemically inert version of RNA—better for safe storage of genetic information, worse for folding up into a catalyst. This is what you might expect if it emerged after RNA had already handed off the job of catalysis to genetically encoded protein enzymes. The RNA would end up sandwiched in the middle of DNA and proteins, just where we find it today.

Observations that expand on all of these themes continue to accumulate and are beginning to sketch a framework that was completely unknown in the mid-1960s. At its best, this "RNA-world" hypothesis solves much of the puzzle for the origin of living systems. One molecule, RNA, is its own catalyst and

information carrier. However, many puzzles remain. For instance, the universe seems quite good at making amino acids without life. They have been found in meteorites, formed in simulations of the conditions of interstellar space, and turn up reliably in just about every possible simulation of our planet's early conditions. For nucleotides, the building blocks of RNA, the exact reverse is true. It seems relatively simple to make the nucleobases (such as adenine and guanine)—but these must be chemically linked to a ribose sugar and a phosphate in order to make a single nucleotide in processes that are antagonistic to those in which the bases form: there are real chemical difficulties in forming the individual nucleotide building blocks, and even bigger difficulties for linking them together into sequences that do not also contain all sorts of unwanted molecular garbage.[45] If RNA came first, then why is it so much easier to make amino acids than RNA from nonbiological scratch?

Scientists are relatively confident that our world, in which DNA genes are copied into mRNA transcripts *en route* to protein translation, was preceded by a simpler biology that comprised only RNA and (genetically encoded) proteins. Every clue that we can find supports this conclusion. A more mysterious question is how this earlier RNA-protein world emerged. One broad class of ideas asserts that we have simply failed to discover some set of conditions that encourages sequences of RNA to form spontaneously. Mineral surfaces are often mentioned here, as they can catalyze many chemical reactions. For example, in 2004, the mineral borate was shown to catalyze the notoriously difficult synthesis of ribose—an essential component of the chemical structure of every single nucleotide.[46] Perhaps other minerals will be found to help other steps in nucleotide synthesis, and for linking nucleotides into sequences. Certainly chemists, geologists, and biologists are talking more than ever before as they seek to add up their knowledge of the ways in which life, chemistry, and the planet interact. Among them, increasing attention is coming to focus on hydrothermal vents as a good place to look next in the search for the origin of life.[47] Here, hot water full of interesting chemicals is forced to flow over richly diverse minerals. This can produce a slew of chemical reactions, most of which are still poorly understood.

Another view is that scientists searching for nonbiological origins for RNA are looking in the wrong place. Instead, genetic information, at least in the form that we think of (polymerized nucleotide sequences), was itself an evolutionary invention of an earlier metabolism, a pre-RNA world. Perhaps significantly, proponents here are also drawn to minerals and to hydrothermal vents because the same conditions that might aid nucleotide synthesis produce a wide diversity of interesting and newly discovered chemical reactions.[48]

It might even be that these two views will meet up one day. Since the mid-1960s, a scientist by the name of Graham Cairns-Smith has been proposing that minerals *were* the original genetic information.[49] Crystalline minerals show the interesting property of harnessing energy from the environment to grow by making copies of themselves. As they do this, they are creating chemical order from chaos. That is exactly what a salt crystal is doing as you watch saltwater evaporate in a glass or a rock-pool. Crystalline minerals also show the potential to catalyze specific chemical reactions on their surface according to their exact atomic composition.[50] In effect, they might carry simple genetic information that starts to trap the energy flowing through the system into a chemical reflection of the environment. But by now we are talking about one of the swarm of competing ideas at the edge of Category 2. Here they will compete and rise or fall according to the evidence that can be gathered through careful and ingenious tests.

Summary

Evolutionary theory, like any other branch of science, achieves progress by testing new ideas. Some of these ideas will go on to change what we thought we knew, others will be found incorrect, and some will stagnate as they fail to gather clear evidence, for or against. For evolutionary theory, many suggestions have been made for new causal factors that are required to explain how genetic diversity has arisen. ID, for example, proposes that some types of genetic information cannot evolve through natural processes unless we admit a role for an intelligent designer. This proposition claims testability by using a definition of information that usually refers to creation by an intelligent agent. Meanwhile, many biologists perceive that they are able to understand exactly where

life's genetic information comes from (the local environment) by thinking in terms of more fundamental and well-established definitions of information that do not involve intelligent design. A related suggestion is that current evolutionary theory cannot explain how natural processes could produce a genetic information system in the first place. I agree that we are far from a full understanding, but I choose to outline some major themes in the scientific progress made since the discovery of life's central dogma in 1966 to provide a context for readers to judge for themselves.

It would be remiss to finish an article in this journal without some comment on the theology of all this. If we accept the evolutionary explanations sketched above, then science is taking major steps toward understanding the mechanism by which life came into the universe. Some famous advocates of this science claim it presents a logical connection to an atheistic worldview.[51] Many others (myself included) perceive that any connection between evolution and spirituality is an act of faith—and faith in atheism is only one of many options.[52] For my part, I find excitement and challenge in the search to unravel this marvelous mystery. I choose to associate that inspiration with a loving creator God whose universe I am exploring. I agree with Dawkins (and Darwin) that from a human standpoint, the suffering and death implicit to natural selection form questions for my faith—and I am grateful that scientists and theologians are able to discuss such issues in forums such as this,[53] where I can read, learn, and grow my relationship with God through an exploration of science.   ✌

Notes

[1] This definition appears, for example, within the classic textbook for undergraduates, D. J. Futuyma, *Evolution* (Sunderland, MA: Sinauer Associates, 2005).

[2] For an accessible discussion of this topic, see Neil Shubin, *Your Inner Fish: A Journey into the 3.5-Billion-Year History of the Human Body* (New York: Vintage, 2009).

[3] For example, see the review by K. Omland and D. Funk, "Species Level Paraphyly and Polyphyly," *Annual Reviews in Ecology, Evolution and Systematics* 34 (2003): 397–423; Christopher E. Bird, Brenden S. Holland, Brian W. Bowen, Robert J. Toonen, "Diversification of Sympatric Broadcast-Spawning Limpets (*Cellana spp.*) within the Hawaiian Archipelago," *Molecular Ecology* 20 (2011): 2128.

[4] For example, see Gary Parker, *Creation: Facts of Life* (Green Forest, AR: New Leaf Press, 2006), 75–148. This text is freely available online at: http://www.answersingenesis.org/articles/cfl/.

[5] For an excellent review of the history by which evolutionary thought absorbed and dismantled these ideas to reach the "(neo-)Darwinian synthesis," see P. J. Bowler, *Evolution: The History of an Idea* (Berkeley, CA: University of California Press, 1983), chapter 9, "The Eclipse of Darwinism."

[6] P. Senter, "Using Creation Science to Demonstrate Evolution: Application of a Creationist Method for Visualizing Gaps in the Fossil Record to a Phylogenetic Study of Coelurosaurian Dinosaurs," *Journal of Evolutionary Biology* 23 (2010): 1732–43. For a more accessible overview of this article, see http://www.bbc.co.uk/blogs/wondermonkey/2011/07/faith-versus-science-does-crea.shtml.

[7] A good, recent summary is presented by Jerry Coyne, *Why Evolution Is True* (New York: Viking Penguin, 2009).

[8] Natural mutations can sometimes have large effects, particularly in genetic regions that influence deep developmental pathways of multicellular organisms (i.e., the genes that control how other genes are switched on and off to build an adult organism from a single fertilized egg cell). However, these changes are generally deleterious to the organism, and are therefore unusual components of an evolutionary lineage. A deeper discussion of this type of mutation can be found in S. B. Carroll, "Homeotic Genes and the Evolution of Arthropods and Chordates," *Nature* 376 (2005): 479–85. I would draw the reader's attention to the broader context: these sorts of mutations are limited to relatively few events on one small branch of the tree of life. In terms of general macroevolution for life on our planet, biologists do not view these events as typical in the formation of new species.

[9] Some of these suggestions for "skyhooks" and "cranes" that would like to lift the natural processes of evolution to produce higher levels of genetic change are discussed in chapter 3 of Daniel Dennett, *Darwin's Dangerous Idea: Evolution and the Meaning of Life* (New York: Simon and Schuster, 1995), 61–84.

[10] See, for example, Simon Conway Morris, *Life's Solution: Inevitable Humans in a Lonely Universe* (Cambridge: Cambridge University Press, 2003).

[11] See, for example, Stephen Jay Gould, *Wonderful Life* (New York: W.W. Norton, 1989). Gould is extreme in his view, but is closer to the position of mainstream evolutionary science, as can be seen from reviews of the books in which Morris (*Life's Solution*) argues for inevitable humans (e.g., the review by the National Center for Science Education, http://ncse.com/rncse/30/review-deep-structure-biology).

[12] For example, see Simon Conway Morris, ed., *The Deep Structure of Biology: Is Convergence Sufficiently Ubiquitous to Give a Directional Signal?* (West Conshohocken, PA: Templeton Foundation Press, 2008).

[13] For example, see William Dembski, *Intelligent Design: The Bridge between Science and Theology* (Downers Grove, IL: InterVarsity Press, 1999), chapter 6.

[14] William Dembski, "Intelligent Design as a Theory of Information" (1998). Web material copyrighted to William Dembski, http://www.arn.org/docs/dembski/wd_idtheory.htm.

[15] Nothing in evolutionary theory suggests that there *must* be an increase in the length or complexity of a DNA molecule over time: for example, many bacteria and viruses appear to

have undergone extensive natural selection to reduce the size of their genetic material as a specific adaptation to make copies of themselves faster than their competitors. For a recent example, see: N. Nikoh, T. Hosokawa, K. Oshima, M. Hattori, T. Fukatsu, "Reductive Evolution of Bacterial Genome in Insect Gut Environment," *Genome Biology and Evolution* 3 (2011): 702–14, doi:10.1093 /gbe/evr064.

[16]R. Redon, S. Ishikawa, K. R. Fitch, L. Feuk, G. H. Perry, T. D. Andrews, H. Fiegler, M. H. Shapero, A. R. Carson, W. Chen, E. K. Cho, S. Dallaire, J. L. Freeman, J. R. González, M. Gratacòs, J. Huang, D. Kalaitzopoulos, D. Komura, J. R. MacDonald, C. R. Marshall, R. Mei, L. Montgomery, K. Nishimura, K. Okamura, F. Shen, M. J. Somerville, J. Tchinda, A. Valsesia, C. Woodwark, F. Yang, J. Zhang, T. Zerjal, J. Zhang, L. Armengol, D. F. Conrad, X. Estivill, C. Tyler-Smith, N. P. Carter, H. Aburatani, C. Lee, K. W. Jones, S. W. Scherer, and M. E. Hurles, "Global Variation in Copy Number in the Human Genome," *Nature* 444 (2006): 444–54.

[17]For an excellent, evolving review of the interesting topic of genome sizes, see T. R. Gregory, *Animal Genome Size Database* (2005), http://www.genomesize.com/statistics.php.

[18]Dembski, "Intelligent Design as a Theory of Information."

[19]If you would like to consider the implications of combinatorial language in greater detail without any formal mathematics, try reading Jorge Luis Borges's famous short story entitled "*The Library of Babel.*" Available in English, *Labyrinths: Selected Stories and Other Writings,* ed. D. A. Yates and J. E. Irby (New York: New Directions Publishing, 1964), 51–9.

[20]In fact, what is harder is to deduce which of the many routes is most likely, if you assign slightly different probabilities to each different type of step. This explains why the past couple of decades have seen considerable research effort go into developing computer algorithms that estimate the most likely series of mutation steps that separate two versions of genetic material. To understand the level of complexity here, consider some different routes by which a series of letter-mutations could transform the word "evolution" into "creation," and then scale that challenge upwards to do something similar for two sentences, two paragraphs, two novels. A good, recent overview is given in K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar, "MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods," *Molecular Biology and Evolution* 28 (2011): 2731–9.

[21]Randy Isaac, "Information, Intelligence, and the Origins of Life," *Perspectives on Science and Christian Faith* 63, no. 4 (2011): 219–30.

[22]For an accessible, eloquent discussion of where this can lead, see Richard Dawkins, *The Blind Watchmaker* (New York: W. W. Norton & Company, 1986), 43–77.

[23]On a different note, it is interesting to see how this same line of thought parallels theological examination of the famous biblical text that humanity was created in the image of God (Gen. 1:27). If each of us is built in the image of God, and each of us is different, then it follows that each of us is capable of developing a different relationship with God based on the unique perspective granted us. This observation provides a logical check to any theologies that assert necessary submission to a single, all-embracing interpretation of God's revealed truth. Within the Gospels, Jesus's personal encounters show a consistent emphasis on the unique point of connection between an individual's perspective and God's greater truth (e.g., compare John 3:1–7; John 4:1–29; Mark 17:10–22; Matt. 8:5–13; Luke 23:33–43), together with a consistent wariness toward group ideologies (e.g., Mark 12:18–27; Matt. 12:1–9; Matt. 15:1–11).

[24]This reductionist description of evolution contains little that is new (scientifically) precisely because the aim of this article is to explain how classic neo-Darwinian orthodoxy addresses the issue of the origin of (new) genetic information. This view of evolution is probably best known through the popular works of writers such as Dawkins, and everything written here is in true alignment with insights expressed in his books such as *The Selfish Gene, The Blind Watchmaker,* and (most relevant to criticisms of reductionism) *The Extended Phenotype.* Behind these works lies an extensive primary research literature that has developed these ideas, before and after, with respect to genomics, genetics, biological development ("embryology"), animal behavior, morphology, life history strategies, and so on. This reductionist view does not overlook the existence of phenotype as the filter through which the environment passes its information into DNA—thereby explaining why *The Extended Phenotype* is the most relevant popular work to discuss in this context—but as Dawkins explains so clearly in *The Selfish Gene*, environmental pressures that do not create a corresponding "match" within DNA are irrelevant to evolution precisely because heritability is one of the three tenets (variation, heritability, and competition to reproduce) that lead to Darwin's inescapable conclusion: heritable variations which increase the reproductive success of a lineage will, over time, accumulate.

[25]For a fascinating and accessible discussion of the incorrect ideas that paved the way for these discoveries, see B. Hayes, "The Invention of the Genetic Code," *American Scientist* 86 (1998): 8–14.

[26]J. D. Watson and F. H. C. Crick, "A Structure for Deoxyribose Nucleic Acid," *Nature* 171 (1953): 737–8.

[27]L. Frisch, ed., "The Genetic Code," *Cold Spring Harbor Symposia on Quantitative Biology* (1966): 1–747.

[28]More accurately, "A", "T", "G," and "C" refer to the four bases used in genetic coding. Bases are part of a whole nucleotide—the base must be added to a molecule of ribose and a phosphate to form a nucleotide. The ribose-phosphate construction is used as a universal scaffolding for joining together sequences of bases. This technical differentiation becomes important to the origin of genetic information because bases are relatively easy to produce under prebiotic conditions, full nucleotides much less so. This and other subtleties are described further in a later section, explained well in Robert Shapiro, "A Simpler Origin for Life," *Scientific American* (June 2007): 47–53.

[29]This key insight brought Christian Anfinsen a Nobel prize in 1972, and a brief overview is found in his classic paper, C. B. Anfinsen, "Principles That Govern the Folding of Protein Chains," *Science* 181 (1973): 223–30.

[30]F. H. C. Crick, "The Origin of the Genetic Code," *Journal of Molecular Biology* 38 (1968): 367–79.

[31]Given that there are really only 64 different rules for converting genetic information into proteins, and an individual protein can be several hundred amino acids in length, most genes use each of these rules many times over.

[32]For a much more thorough and technical version of this section, including several hundred references to the primary scientific literature, see S. J. Freeland, "'Terrestrial' Amino Acids and Their Evolution" in *Amino Acids, Peptides and Proteins within Organic Chemistry* 1, ed. A. B. Hughes (Weinheim, Germany: Wiley-VCH, 2009).

[33]R. D. Knight, S. J. Freeland, and L. F. Landweber, "Rewiring the Keyboard: Evolvability of the Genetic Code," *Nature Reviews Genetics* 2 (2001): 49–58.

[34]For a brief overview, see J. F. Atkins and R. Gesteland, "The 22nd Amino Acid," *Science* 296 (2002): 1409–10.

[35]Knight, Freeland, and Landweber, "Rewiring the Keyboard: Evolvability of the Genetic Code."

[36]For an accessible overview of this topic, see S. J. Freeland and L. D. Hurst, "Evolution Encoded," *Scientific American* 290 (2004): 84–91. A more technical and more recent treatment of this topic can be found in A. S. Novozhilov and E. V. Koonin, "Exceptional Error Minimization in Putative Primordial Genetic Codes," *Biology Direct* 4 (2009): 44.

[37]For a detailed review, see E. V. Koonin and A. S. Novozhilov, "Origin and Evolution of the Genetic Code: The Universal Enigma," *IUBMB Life* 61 (2009): 99–111.

[38]More than fifty models are considered in E. N. Trifonov, "Consensus Temporal Order of Amino Acids and Evolution of the Triplet Code," *Gene* 261 (2000): 139–51.

[39]Compare the similarities in two recent reviews: P. G. Higgs and R. E. Pudritz, "A Thermodynamic Basis for Prebiotic Amino Acid Synthesis and the Nature of the First Genetic Code," *Astrobiology* 9 (2009): 483–90; and H. J. Cleaves, "The Origin of the Biologically Coded Amino Acids," *Journal of Theoretical Biology* 263 (2010): 490–8.

[40]G. K. Philip and S. J. Freeland, "Did Evolution Select a Nonrandom 'Alphabet' of Amino Acids?" *Astrobiology* 11 (2011): 235–40.

[41]Jonathan K. Watts, "Biological Information, Molecular Structure, and the Origins Debate," *Perspectives on Science and Christian Faith* 63, no. 4 (2011): 231–9. The current status of data here is reviewed in M. Yarus, J. J. Widmann, and R. Knight, "RNA-Amino Acid Binding: A Stereochemical Era for the Genetic Code," *Journal of Molecular Evolution* 69 (2009): 406–29.

[42]Watts, "Biological Information, Molecular Structure, and the Origins Debate," 235.

[43]For an introduction and references to more detailed material, see T. R. Cech, "The Ribosome Is a Ribozyme," *Science* 289 (2000): 878–9.

[44]A brief review of data here is found in S. J. Freeland, R. D. Knight, and L. F. Landweber, "Do Proteins Predate DNA?" *Science* 286 (1999): 690–2.

[45]Readers who are interested in this particular subtopic are encouraged to read Shapiro, "A Simpler Origin for Life." Shapiro's passionate emphasis represents the best traditions of scientific skepticism, ruthlessly pointing out some very real problems with all current attempts to explain how a nonliving universe could have produced RNA. In particular, widespread enthusiasm for the RNA-world has become so fashionable that even high-profile scientific publications which explicitly seek to demonstrate prebiotic origins for the RNA-world continue to ignore well-understood and long-standing criticism of the problems. For example, one recent high-profile paper claims to demonstrate prebiotic plausibility for synthesis of nucleotides: M. W. Powner, B. Gerland, and J. D. Sutherland, "Synthesis of Activated Pyrimidine Ribonucleotides in Prebiotically Plausible Conditions," *Nature* 459 (2009): 239–42. This interesting work shows that exactly the right purified solution of linear organic molecules can cyclize under the right conditions to present activated nucleotides. However, it entirely misses Shapiro's "garbage bag" point—that one of the biggest challenges for understanding the evolution of an RNA-world is to understand how building blocks form into oligonucleotides when they are coming from any sort of messy molecular organic broth (rather than from a purified solution of exactly the right reactants under exactly the right conditions). There is no chemical reason why nucleotides should form and stick to one another rather than to other chemicals produced in the same broth—such as amino acids, alcohols, esters, etc. Of further note, the chemistry reported in this *Nature* paper bears no resemblance to the reactions by which living organisms have been making nucleotides for more than 3 billion years. Maybe early life changed its metabolic pathways beyond recognition—but as yet we have absolutely no evidence for this whatsoever: prebiotically *possible* and prebiotically *plausible* are subtly different concepts.

[46]For example, see the comment by Steve Benner on page 52 of Knight, Freeland, and Landweber, "Rewiring the Keyboard: Evolvability of the Genetic Code." For a more detailed treatment, see H. J. Kim, A. Ricardo, H. I. Illangkoon, M. J. Kim, M. A. Carrigan, F. Frye, and S. A. Benner, "Synthesis of Carbohydrates in Mineral-Guided Prebiotic Cycles," *Journal of the American Chemical Society* 133 (2011): 9457–68.

[47]A very readable overview of this topic can be found in Nick Lane, *Life Ascending: The Ten Great Inventions of Evolution* (New York: W.W. Norton & Company, 2009), chapter 1.

[48]One recent example is given by R. E. Mielke, M. J. Russell, P. R. Wilson, S. E. McGlynn, M. Coleman, R. Kidd, I. Kanik, "Design, Fabrication, and Test of a Hydrothermal Reactor for Origin-of-Life Experiments," *Astrobiology* 10 (2010): 799–810.

[49]Although Cairns-Smith's ideas date back to the mid-1960s, they are most accessibly presented in his later book, *Seven Clues to the Origin of Life* (New York: Cambridge University Press, 1985).

[50]For a broad introduction to this progress, as of 2001, see Robert M. Hazen, "Life's Rocky Start," *Scientific American* 284 (2001): 77–85.

[51]For example, see Richard Dawkins, *River Out of Eden* (New York: Basic Books, 1995), 95–135.

[52]For example, see the letter(s) and signatories of the Clergy Letter Project, http://www.theclergyletterproject.org/.

[53]For example, the excellent pair of articles, Junghyung Kim, "Naturalistic versus Eschatological Theologies of Evolution," *Perspectives on Science and Christian Faith* 63, no. 2 (2011): 95–107; and Keith Miller, "'And God Saw That It Was Good': Death and Pain in the Created Order," *Perspectives on Science and Christian Faith* 63, no. 2 (2011): 85–94.