DARWINIAN EVOLUTIONARY THEORY AND THE LIFE SCIENCES IN THE 21ST CENTURY

Roland F. Hirsch

This essay was originally published in "Uncommon Dissent" (ISI Books, 2004) edited by William Dembski.

Evolutionary theory has had a major impact on the development of biology since the appearance of *On the Origin of the Species* in 1859. Over the century following publication of that book, experiments and field observations led to successive refinements of the Darwinian theory of evolution, and it was confidently proclaimed as the foundation of biology in the Darwin Centennial year of 1959.

Such confidence is not warranted today. New technologies developed in the past four decades have revealed to us the chemistry underlying biological processes. These technologies have revealed that life is far more complicated than was imagined in 1959, and that much of its complexity cannot easily be addressed by existing evolutionary theory. Indeed some of the major discoveries in the life sciences presented in this article were hardly anticipated by evolutionary theory, but instead came out of advances in experimental technologies. The physicist Freeman Dyson put it this way in his book of essays *Imagined Worlds*:

> There are two kinds of scientific revolutions, those driven by new tools and those driven by new concepts.... The effect of a concept-driven revolution is to explain old things in new ways. The effect of a tool-driven revolution is to discover new things that have to be explained. In almost every branch of science, and especially in biology and astronomy, there has been a preponderance of tool-driven revolutions. We have been more successful in discovering new things than in explaining old ones.[1]

This essay is a personal account of how the author's assessment of evolutionary theory has been changed by advances in biology. During the past 15 years my involvement in analytical

chemistry, genome sequencing technologies and structural biology has brought home to me the great impact of new technologies in these fields on how we understand life.[2] Much has been revealed about the nature of the genome; the complicated networks that regulate expression of genes; the chemical transformations that expressed proteins undergo after synthesis in the cell; the structure of proteins in three dimensions; and how proteins and other biological molecules form the complex machines that carry out most of the functions of a cell. Much has also been learned about how cells interact, especially about how microbes interact in communities. We will discuss several recent, technology-driven advances in the life sciences that have convinced the author that the standard evolutionary theory is inadequate, and that evolution should no longer hold the place once claimed for it as the foundation of the life sciences.

What genome sequences tell about evolution

The genome of an organism consists of one or more segments of deoxyribonucleic acid (DNA). DNA consists of two long, complementary strands each of which is composed of a backbone onto which are bound the four bases that comprise the "alphabet" of the genetic code, adenine, cytosine, guanine and thymine. The sequence in which these bases are organized determines both the composition of proteins that carry out the functions of the organism, and whether or not a particular protein will be expressed or not at any given time. Advances in analytical chemistry, automation, bioinformatics and other fields during the last two decades changed the process of obtaining the sequence of bases in a genome from one that was slow and expensive to one that is rapid and inexpensive.[3] As a result, base sequences for many complete genomes have been determined, more than 100 as of this writing.[4]

These complete genome sequences have revealed several complexities that Darwinian evolutionary theory did not anticipate. Four of these will be discussed here: the major role played by transfer of genes from one species to another as opposed to inheritance from ancestors; the fact that bacterial species do not evolve solely in a random fashion, but show a bias toward deletion of

genetic material; the discovery that much of the portions of the genome that do not code for proteins is not "junk DNA" but in fact has a critical function; and the observation that expression of genes is controlled by regulatory circuits that are as complicated and as precisely arranged as the most sophisticated engineering diagrams.

The established Darwinian concept of inheritance is in a vertical direction, that is, from the older generation to the younger generation, extending down in time from a single ultimate ancestor cell. As time proceeds, branchings repeatedly occur that represent the formation of new species. This is commonly represented symbolically by a diagram that looks like a tree. The National Academy of Sciences booklet *Teaching about Evolution and the Nature of Science* (published in 1998) shows a diagram of a tree of descent and states: "The ability to analyze individual biological molecules has added great detail to biologists' understanding of the tree of life. For example, molecular analyses indicate that all living things fall into three domains—the Bacteria, Archaea, and Eucarya—related by descent from a common ancestor."[5]

Yet just a year after this book was published it was clear that this was not so; the availability of complete genome sequences for two dozen microbial species led to a surprising discovery, one that previous information had only hinted at.[6] The genome sequences revealed that for at least two microbial domains (bacteria and archaea), much of the inheritance is in a horizontal direction. Significant portions of the genome of a particular species came not from its ancestors but arrived at various times from the species that were its neighbors through what is termed "horizontal gene transfer" (hgt, sometimes also called "lateral gene transfer" or lgt). The relationships among species could no longer be arranged in a tree of descent, for the genome of a species was only partially inherited from an ancestor species. Each gene in the DNA of a particular species that arose through hgt would trace a different ancestry, as it could have come from a different neighboring species in a different generation of the species being studied.

Thus there is no single "tree of life" but rather a "web" or "net" of interconnections that are both vertical and horizontal. As stated by Doolittle, "If, however, different genes give

different trees, and there is no fair way to suppress this disagreement, then a species (or phylum) can 'belong' to many genera (or kingdoms) at the same time: There really can be no universal phylogenetic tree of organisms based on such a reduction to genes."[7] And it also seems likely that—contrary to the assertion in the National Academy of Sciences document cited above—there was no "common ancestor cell," but rather it is now thought more likely that there was a pool of cells that changed communally over a long period of time.[8]

Indeed, so much transfer has occurred across the three domains of life that they must be defined not on the basis of ancestry but on the basis of functional properties. If the focus was on evolutionary ancestry of species during the 20th century, clearly the focus will be on function (how the component parts of a cell or organism work) rather than ancestry during the 21st century. The construction of evolutionary trees does continue as it enables identification of logical functional groupings of species. However, the goal no longer is construction of a single unique tree of life, but rather development of statistical best fits to genome data. In some cases individual genes are used to construct the trees in order to identify organisms that share a particular property coded for by that gene, regardless of whether they share any other common properties. In other cases, an average "whole-genome tree" will be constructed to give an estimate of overall similarities among and differences between species. Even the choice of which type of tree to use is the subject of considerable controversy in phylogenomics, the study of family relationships among species.[9]

A particularly significant example of horizontal gene transfer involves the photosynthetic microorganisms. Photosynthetic bacteria are found in five quite different phyla. The Darwinian inheritance model would predict that either the five phyla diverged from a common ancestor or that the photosynthetic function evolved independently five times. A study of the genomes of a representative from each phylum revealed that neither of these predictions is correct: the genomic comparisons show that significant numbers of genes relevant to photosynthetic function must have been transferred horizontally among the species studied.[10]

Horizontal gene transfer was identified in bacteria and archaea already by 1990. It was thought for the following decade that hgt was not a significant factor for the more complex eukaryotes. Recent studies of complete genome sequences however have identified substantial hgt in many of the branches of the eukaryotic domain. Examples include hgt involving the protists (unicellular eukaryotes) such as the parasite *Giardia Lamblia*[11], flowering plants,[12] algae,[13] fungi,[14] and nematodes, the most abundant of all metazoans, including the worm *Caenorhabditis elegans* as well as many plant and animal parasites.[15] The latter group is especially significant for it involved relatively complex, multicellular organisms, and because some of the transferred genes are critical to the functioning of the recipient species.

Thus horizontal gene transfer must be considered a significant source of genetic variation in all three domains of life. The purely vertical pattern of inheritance axiomatic to Darwin's theory of evolution from his own writings down to the present clearly is inadequate to explain the observed complexities of the origin of genes.

The second discovery that throws doubt upon the Darwinian evolutionary paradigm also comes out of the increasing availability of complete bacterial genome sequences. While horizontal gene transfers would increase the size (number of base pairs (bp)) of a genome, bacterial genomes turn out to be remarkably small. The genome size for the bacteria and archaea ranges from about 500,000 bp (e.g., for *Mycoplasma genitalium)* to under 10 million bp (e.g., about 7 million bp for *Mesorhizobium loti*). Yet through many generations of a species, numerous instances of hgt can be documented, as well as cases of gene duplication. Why, then, are the genomes uniformly small in these two domains of life? A persuasive answer is that there is a bias toward deletion of genetic material in bacteria, that is, that portions of the DNA tend to fall away when the bacterium reproduces: "The obvious answer is that lineages must undergo the inactivation and loss of genes, and the elimination of the corresponding DNA that made up the genes. This could result if the mutational process driving the structural evolution of chromosomes is biased towards DNA loss."[16] Thus the standard Darwinian mechanism of random mutation and

natural selection is inconsistent with the observed fact of a non-random bias toward deletion of DNA.

The third unexpected discovery deals with regions of a genome that do not carry the code for a protein. At the time that the project to sequence the complete human genome began, it was already realized that most of the genome did not code for a protein. As little as 2 percent (and probably even less) of the human genome of some 3 billion base pairs was thought to be in the actual genes. The rest was thought to be largely "junk DNA," that is, stretches of DNA that had no function. The evolutionary theoretical basis for the idea of "junk DNA" was that if a segment of DNA was not part of a gene that carried the code for a protein, then there was no mechanism for natural selection to act on this segment. The segment would be hidden from the action of natural selection because it would not be expressed in any form that affected the functional properties and hence the survival of the organism. Instead the segment would be affected by periodic random mutations that would scramble any code that originally might have been carried by the segment.

Today the experimental evidence suggests that much, though probably not all, of the non-coding regions of the genome have critical roles in the development and function of an individual. Completion of the human and mouse genome sequences in particular has resulted in useful insights into the function of non-coding DNA.[17] There appear to be fewer genes in the human genome than the more than 100,000 that many specialists thought were present before the completion of the human genome project. Most current estimates range from 30,000 to 60,000 genes, with a few going higher. The small number of genes suggests that the non-coding regions must have key roles to play, including even repetitive portions of the DNA.[18] As the authors of a recent review point out, "From genomic analysis it is evident, however, that with increase of an organism's complexity, the protein-coding contribution of its genome decreases...."[19] Clearly the non-coding regions must have crucial roles in accounting for this complexity.

This leads to the fourth point. It is clear that the mechanism by which expression of many of these genes is controlled is much more complicated than was previously believed. It is not a simple matter of the code carried by a gene being decoded to make a messenger RNA (mRNA) to enable the synthesis of a protein. Instead, a group of genes encodes a set of regulatory factors that controls the transcription of the gene that is to be decoded for synthesis of a protein by the cell (translation of the genome code into a polypeptide). And, in many cases, the expression of several genes is controlled by a single set of regulatory genes, since a cellular function may depend on a number of different proteins being produced at the same time. The set of genes and associated regulatory factor(s) in the genome required for a particular cellular function forms an operon. The components in an operon act in a precisely timed fashion so that the entire process is best visualized as a circuit diagram that resembles closely the diagrams used in the design of processes in chemical engineering or the design of circuits in electrical engineering or electronics. Even the language is similar: "oscillators," "feedback loops," etc.

There indeed is considerable interest in using engineering principles to understand gene networks and to design new ones. The authors of a recent review of this topic note that interest in mathematical modeling of gene regulatory networks appeared already in the 1970s and state: "[R]ecent experimental advances have reignited interest in the development of circuit analysis techniques for describing complex gene networks."[20]

While gene regulation applies to expression of practically all genes at all times, the regulation of gene expression during development is of especial interest. The genome of an animal is fixed at conception, yet expression of the genome changes constantly from stage to stage of development both before and after birth. The genomic code that regulates these changes works both by encoding factors that regulate transcription of the genes for structural and functional proteins, and by encoding the regulatory network that controls expression of these factors. The diagrams of these gene regulatory networks visualize the high level of

interconnection and interdependence of the individual genes and regulatory factors for each component of development.[21]

The study of regulation of gene expression has led to an appreciation that variation in gene sequence does not explain all variation in function among members of a species. Modifications of the genome that do not affect the gene sequence are called *epigenetic* phenomena. Such changes are responsible for cells in a particular organ of a multicellular organism showing the appropriate characteristics of that organ even though the cells carry the complete genome and not just the components required for becoming the specialized cells of that organ. It is also possible that development of cancer and other diseases depends on epigenetic modifications in the DNA of cells rather than mutations in the base sequence itself.[22] And there is speculation that Lamarck's concept of inheritance of acquired traits may be valid in some cases.[23]

The study of gene expression has been greatly aided by new developments in technologies such as mass spectrometry and gene microarrays. The study of protein composition in a cell using these high resolution techniques has been called "proteomics."[24] Mass spectrometry now allows identification of large numbers of proteins in a cell grown under particular conditions and then comparison of the amounts of each protein relative to the amounts in the same species grown under different conditions.[25] The effects of a stress on expression of a large number of proteins can thus be assessed, and the changes in gene expression with time can be determined to obtain insight into the complexities of this process that are not revealed by the genome itself. Gene microarrays make it possible to follow expression patterns across an organism, determining which genes are induced and which are repressed at various points in the life cycle of a cell.[26]

To sum up this section, Darwinian evolutionary theory failed, in this author's view, to anticipate several key discoveries about genetics, inheritance, and gene expression and development. In each case, evolutionary theory should have guided researchers to make these discoveries, but in fact the opposite seems true: changes were made in evolutionary theory after

the fact to account, for example, for the significance of horizontal gene transfer or to explain the complexities of regulation of gene expression.

What proteins tell about evolution

We have seen that the relationship between the genome and the function of an organism is much more complex than can be accounted for by the traditional Darwinian-oriented concept of

one gene → one protein → one function.

But we have only looked at the first part of this sequence; there is more to consider. Additional complexities come into play after translation of the genetic code into a protein molecule, complexities that again seem beyond the reach of evolutionary theory. In this section I will discuss how an expressed protein is chemically transformed into the actual molecule that participates in cell functioning, how it folds into the shape required for exhibiting its characteristic activity, and how the transformed, folded protein then must become part of a multi-protein complex that is actually responsible for the function in which that protein participates.

Proteins are not ready to perform their function as they are synthesized in the cellular protein factory called the ribosome (about which more shortly). Each protein must undergo a series of changes in shape and composition before it can "do" anything. It is likely that a majority of proteins in most cells are involved in chemical reactions that change the covalent structure of the protein molecule into its active form, reactions such as methylation, acetylation, phosphorylation, deamidation, oxidation and dozens of others. An interesting example of this is a study of modifications of proteins in lens tissues of eyes, seeking to understand the changes that are responsible for the formation of cataracts. The authors note that 90 percent of the protein in the lens is in one family of proteins, the crystallins, and that these proteins are long-lived and thus subject to post-translational modification during aging and cataract formation. The cataract studied was congenital, so a linkage to the genome of the patient exists, even though the linkage is likely to be complex given the large number of modifications discovered.[27]

The protein molecule must fold into the shape in which it is active. Folding is a complicated process that requires some steps to occur on the picosecond time scale while others may take milliseconds to seconds.[28] Portions of the process are often mediated by chaperone or cofactor[29] proteins that are required to guide folding into the proper shape or to prevent misfolding.

The folded shape of a protein is determined by several factors, among them internal covalent bonds (such as disulfide bridges between cysteine units in the chains), hydrogen bonds, and hydrophilic and hydrophobic interactions with the solvent surrounding the protein molecule. Two proteins with very different amino acid sequences can fold into a closely similar shape and have a similar function. The three-dimensional design of the resulting protein is thus more important than the sequence in explaining function. There is much interest in classifying folds of proteins to better understand function and to identify the likely function of newly discovered proteins.[30] The shape is also affected by interaction of the protein with the other molecules, large and small, that participate with it in carrying out a particular activity in a cell.

This folding process is possible only because it is guided. A process of folding in which the protein chain bends entirely in random ways could not achieve the functional fold of that protein in any useful period of time. Several models have been developed to describe the folding process.[31] "Misfolding" of proteins also is of considerable interest, as it may be a significant factor in the onset of a variety of diseases, such as Alzheimer's syndrome. Proteins that can pass from one molecule to another through changes in folding conformation from the normal conformation (without involvement of the genome) include the prions, protein molecules implicated in a number of diseases.[32]

Most cellular functions are carried out by highly organized protein machines that are assembled with what amounts to engineering precision. This was not realized until relatively recently, in part because experimental technologies were inadequate to detect the structure of the machines, but also—in the author's view—because the Darwinian concept of evolution by

random mutation and natural selection encouraged (practically required) treating each protein (gene product) as a distinct unit in the functioning of an organism. For how could a function requiring multiple proteins in a cellular machine ever arise through the required random mutations that developed one protein molecule at a time and in a stepwise manner, and gave no intermediate product with any function that would allow Darwinian natural selection to work?

Thus the traditional view of protein function was that a protein catalyzes the reaction of a substrate to form product(s). Other proteins would be involved in preceding or following steps in a metabolic cycle but not in a complex with the protein in question. There is no doubt that much progress was made in enzyme biochemistry using this concept. However, proteins usually do not carry out functions in isolation, and the development of electron microscopy, macromolecular crystallography, mass spectrometry and other technologies from the 1950s enabled visualization of the complex protein machines that actually are responsible for most of the properties of a cell.

Recognition of the true complexity of cellular processes has led to a major change in the vocabulary used to describe cells. Words such as "machine", "factory" and "motor" are in routine use,[33] and a cellular function is best explained in terms of design of the machine(s) responsible for the function. As biologist (and President of the National Academy of Sciences) Bruce Alberts explained it:

> We have always underestimated cells. Undoubtedly we still do today. But at least we are no longer as naive as we were when I was a graduate student in the 1960s. Then, most of us viewed cells as containing a giant set of second-order reactions: molecules A and B were thought to diffuse freely, randomly colliding with each other to produce molecule AB—and likewise for the many other molecules that interact with each other inside a cell.…
>
> But, as it turns out, we can walk and we can talk because the chemistry that makes life possible is much more elaborate and sophisticated than anything we students had ever considered. Proteins make up most of the dry mass of a cell. But instead of a cell dominated by randomly colliding individual protein molecules, we now know that nearly every major

> process in a cell is carried out by assemblies of 10 or more protein molecules. And, as it carries out its biological functions, each of these protein assemblies interacts with several other large complexes of proteins. Indeed, the entire cell can be viewed as a factory that contains an elaborate network of interlocking assembly lines, each of which is composed of a set of large protein machines.[34]

Thus the traditional, evolution-influenced concept of function of a protein acting rather randomly and in relative isolation from the other proteins in the cell has given way to the view that protein function can only be defined through its interactions with other proteins and smaller molecules in cellular machines.

Many of these machines are extraordinarily complex, including those machines responsible for the essential core functioning of a cell. The ribosome, which is responsible for the synthesis of all cellular proteins, itself is comprised of two distinct subunits that contain in all some 55 proteins and three ribosomal RNAs in the simplest (bacterial) form and about 75 proteins and 4 ribosomal RNAs in the eukaryotic form. Considering the ribosome as the equivalent of an assembly line in a factory makes sense when one describes what it does: bind the messenger RNA (mRNA) that contains the code for the protein to be built; bind a transfer RNA (tRNA) that holds the first amino acid in the protein chain, based on the first three nucleotides of the mRNA (the codon, or "triplet" code specifying which amino acid to select); bind another tRNA carrying the second amino acid in the chain; form the peptide bond between the two amino acids in the proper order; move the peptide chain forward so that the next amino acid can be added; finally, recognize the stop codon on the mRNA that indicates that all the amino acids have been added, and release the completed protein. All this is done with extremely high accuracy.

Imaging the structure of ribosomes has been an extraordinary challenge. At present remarkably good pictures of the three-dimensional structure of a bacterial ribosome have been produced through a combination of technologies, such as neutron scattering, electron microscopy, and x-ray diffraction.[35] This structural information shows how tightly interlocked the components

of the machine are. No simpler machine is known or even imagined that could carry out all of the steps in protein synthesis with such accuracy and speed. Yet no living cell can exist without the means to rapidly and continually synthesize hundreds of proteins over and over again with high fidelity to the code in its DNA. Even the formation of the ribosome itself requires a large number of synchronized steps and more than 100 proteins and 100 small RNA segments.[36] Clearly the traditional idea that organisms evolve from simple to complex does not apply to the protein synthesis machinery.

Much of the functioning of a cell is carried out by similar machines (even if few have as many components as the ribosome). RNA polymerase II, which synthesizes messenger RNA in eukaryotes, is another critical machine. It is a complex of 12 proteins that holds the DNA strand to be transcribed and puts together the mRNA that will guide the ribosome in producing a protein. The mass of the complex is more than 500,000 daltons, involving more than 3500 amino acid residues. The structure of RNA polymerase II from yeast has been solved in an elegant series of experiments. The references describing these experiments should be read if only to see how the complex is put together from its components by the cellular machinery with a precision exceeding that of the most complicated devices designed and engineered by humans.[37]

Another important category of protein complexes, and one that also has implications for evolutionary theory, comprises the molecular chaperones. These are complexes that contain several proteins and perform a number of cellular functions, such as assisting in the folding of proteins[38] and controlling the binding of hormones to hormone receptors.[39] The former group of chaperones functions to prevent misfolding or aggregation of a protein as it folds after transcription and appears to be required for proper folding of many proteins. Failure of the system of chaperones to guide folding of a protein properly is implicated in neurodegenerative diseases such as Alzheimer's and Parkinson's.[40] The latter help stabilize the receptor in a shape that allows access by the appropriate hormone and, following binding of the hormone to the receptor, helping start transcription of the segment of DNA that is controlled by the hormone.

Another category of chaperones helps control the movement of metal ions in cells. Many of the metals (e.g., copper, iron, zinc) are incorporated in key biochemical constituents of cells. Each cell contains a significant number of ions of these metals, yet it appears that all of the atoms of these metals are tightly bound to regulatory and chaperone proteins.[41] Breakdown of this metal control system allows abnormal binding of a metal ion to any of a large number of proteins where it normally might not attach, shutting down the function of those proteins and potentially leading to disease. Malfunctions of copper metabolism, for example, are implicated in familial amyotrophic lateral sclerosis and may be involved in several other diseases.[42]

In this section I have discussed two levels of complexity of cellular chemistry that extend beyond the control of the gene that codes for a protein: post-translational chemical modification and folding of proteins into their active form and the requirement that most proteins be incorporated into large, multi-protein complexes in order to participate in a cellular function. Both of these phenomena were known in 1959, but the extensiveness of each was unexpected. Certainly, evolutionary theory did not alert biologists to the significance of these modifications and the protein machines, given the one-gene/one-protein outlook that random mutation and natural selection strongly encouraged. Can Darwinian evolution really be as fundamental to biology as is claimed by its proponents if it gave no guidance to scientists about the extraordinary complexity and high degree of organization of cellular chemistry that new experimental tools were to discover?

What microbes and microbial communities tell about evolution

The impact of the discovery of horizontal gene transfer on the Darwinian picture of evolution has already been discussed. There are some important cases of gene transfer into and within the eukaryotic domain,[43] but as far as we know these are rather isolated and infrequent. On the other hand, many of the species in the archaeal and bacterial domains that have been characterized have been affected by this phenomenon.

But there are several other aspects of the biology of microbes that have been discovered since 1959 that also seem to the author to be inconsistent with Darwinian evolutionary theory as it is commonly understood. As with the genome and protein discoveries already discussed, these surprising properties of microbes are the result of greater levels of complexity than can be explained by evolutionary theory.

The Darwinian evolutionary theories are based upon the assumption of continual competition among species and among individuals of a given species. Darwin himself stated that his reading of the theory of Malthus, who argued that a population will increase to the limit of available resources, was the critical step in formulating his own theory of natural selection. Successive generations of biologists refined the theory in various ways but retained the principle of competition as the basis for natural selection. Even the "selfish gene" hypothesis of Dawkins is based on the idea that reproduction is the measure of (competitive) fitness of an organism.

Yet research in microbiology over the past three decades has shown that this often is not so. Microbes of many different species do not under normal conditions compete with each other. Rather, they form stable communities in which each individual has a role and is dependent on other individuals of the same and of other species in the community. Microbes studied in pure cell cultures, the traditional method of microbiology, therefore behave differently from microbes in real environments. This discovery has had great practical implications, for example, in understanding human health and treating diseases, and in describing contaminated environments and discovering how to clean up the contaminants.

One common form of microbial community is called a biofilm—a population of cells growing on a surface and enclosed in a porous polymer matrix. Microbes are commonly found as part of a biofilm, whether on the surface of a soil particle, the root of a plant, a tool or container, or inside a higher organism such as a human.[44] As stated in a review of the implications of biofilms for clinical medicine, "We now understand that biofilms are universal.... Using tools such as the scanning electron microscope and, more recently, the confocal laser scanning

microscope, biofilm researchers now understand that biofilms are not unstructured, homogeneous deposits of cells and accumulated slime, but complex communities of surface-associated cells enclosed in a polymer matrix containing open water channels."[45]

Biofilms are thought to play an important role in many human infections. Biofilms formed by pathogenic bacteria also may be responsible for some of the antibiotic resistance that is observed in bacteria. The common explanation for development of resistance is that a portion of the population of the pathogenic organism is a mutant form that is not affected by the antibiotic. Individuals of normal form of the pathogen are killed by the antibiotic, leaving the resistant, mutant form behind to proliferate. However, cells that are not antibiotic-resistant mutants also can and do survive and reproduce to recreate the original population size. If these *persistent* cells are treated again with the same antibiotic they are killed by it to the same extent as in the original treatment, demonstrating that they are not an antibiotic-resistant mutant. This mechanism of resistance is not an evolutionary one. Persistence appears common in bacteria that form biofilms and needs to be taken into account in handling difficult-to-treat infections.[46]

Microbial communities, whether or not they are fixed in biofilms on surfaces, show a great deal of cooperation. Communication among members of these communities is an active field of research. Chemical signals are released by bacteria and influence the behavior of the entire community. Quorum sensing is an important type of signaling that induces a specific behavior in the community when the numbers of a particular species is sufficiently high (hence the name for this function).[47] The behavior is induced by the concentration of a signaling compound (the autoinducer) exceeding a specified level. At that point transcription of previously inactive genes occurs (is induced by the autoinducing compound). Several of these compounds may be released within a community of microbes, controlling a number of different functions of the members of the community. Quorum sensing is a widespread phenomenon and appears to play a critical role in regulating a variety of human bacterial infections, such as those that cause the complications of cystic fibrosis.[48]

Thus the behavior of a group of bacteria has a level of complexity beyond that of the individual members. Where more than one species is present, the growth and functioning of the community is therefore controlled at a higher level than the functioning of the individual species. The behavior of the species in a community is thus not reducible to the sum of the behaviors of the species isolated in pure cultures. As a review article about the formation of biofilms by human oral bacteria states, "A successful search for genes critical for mixed-species community organization will be accomplished only when it is conducted with mixed-species communities."[49]

The communal behavior of microbes has another important consequence. It means that the simplest microbes—those with the smallest genomes—cannot survive in the absence of other species that contribute needed nutrients or remove waste products. There is much interest in determining the smallest genome that an organism can have and still be viable. But this means viability only in a suitable microbial community. It is important to remember this as we look at the question of the simplest microbial genome.

One reason for being interested in this topic is that it determines the complexity that would have been required for the very first living cells on the Earth. Under the *abiogenesis* hypothesis, life arose from non-living chemicals on the early earth. The requirement for this prebiotic chemistry is that it must have been able to generate all of the components of the simplest living cell. Much thought has gone into developing possible mechanisms for the generation of these chemical species on the early Earth.[50] While it is estimated that the earth is about 4.5 billion years old, the first few hundred million years were inhospitable to life—indeed inhospitable to accumulation of any but the simplest organic chemicals due to repeated sterilizations by impacts of huge objects onto the Earth. Some evidence points to the first appearance of life within as little as a hundred million years from the end of the last massive impacts, and most experts agree that at most three hundred million years or so were available for the complexity of the first living cell to be achieved starting with no more than the basic

inorganic chemicals. Thus, it is important for evaluating abiogenesis theories to have an idea of what the end result required.

The smallest genomes yet found in a living organism (which excludes viruses) include those of *Nanoarchaeum equitans*, which has about 500,000 base pairs (bp) in its genome,[51] *Buchnera* species, some of which have genomes as small as 450,000 bp,[52] and *Mycoplasma genitalium*, which has a genome of about 580,000 bp.[53] These species are quite dependent on the host species in which they are found, being symbiotic or parasitic. Thus, they cannot serve to define the minimum genome size for the very first organism to appear under the abiogenesis theory. However, they do provide an idea of the magnitude of the challenge. Earlier we saw the bias toward minimization of the size of a bacterial genome. There is very little in the genomes of these organisms that is not critical to their functioning, and probably not less than about 300 of their genes in each case are necessary.

Not only must the 300 or so genes be present in the genome, they must code for proteins that are properly matched to their functions as part of the large cellular machines, such as the ribosome, which are required even in these simplest of organisms. And the expression control system must be able to produce the proteins when they are needed, and in the amounts that are needed, in the presence of the reagents needed to transform the proteins into active forms and under conditions favoring proper folding of each protein.

It is controversial whether "evolution" of the first living organism on Earth is a necessary part of the theory of evolution. However, the improbability of this having occurred through the required chance variations in chemistry is so great that it argues against any purely naturalistic history of life on Earth. To the author of this essay, the doubts about the mechanism for life getting started on Earth motivate him to question other naturalistic theories about the early history of life on Earth. They should not be easily accepted when they are based on so many premises based on minimal actual evidence.

In this section I have explained how microbial species do not tend to compete to out-reproduce one another but tend instead to cooperate in stable communities which are more complex that the sum of the properties of the individual species. They engage in intercellular signaling that controls many aspects of the behavior of each species in a community. I have also noted that even the simplest known microbes are exceedingly complicated, requiring a set of several hundred genes of just the right kinds to function. This complexity was not known to Darwin, who lived at a time when microbes were just beginning to be characterized, and indeed little of it was known at the time I consider the peak of Darwinian influence on biology in the 1950s and 1960s. It requires many twists and turns in evolutionary theory to accommodate these complexities, few of which were anticipated by such theories. Indeed, the more that is learned about microbes, the clearer it becomes that while evolutionary theory plays a role in understanding life, it is a much smaller role than its proponents would have the world believe.

Conclusion

I have no doubt that these and other technology-driven advances in the life sciences present a serious challenge to the validity of the main principles of Darwinian evolutionary theory. Much of what was taught forty years ago has had to be unlearned or has become irrelevant; much of what today's experiments and field research reveal about life cannot be explained by the evolutionary theory of the past.[54] Life as revealed by new technologies is more complicated than the Darwinian viewpoint anticipated. Thus evolutionary theory, which was considered to be a key foundation of biology in 1959, today has a more peripheral role. Adam S. Wilkins, the editor of the review journal *BioEssays*, put it this way in introducing an issue of his journal devoted to evolution in December 2000:

> The subject of evolution occupies a special, and paradoxical, place within biology as a whole. While the great majority of biologists would probably agree with Theodosius Dobzhansky's dictum that "nothing in biology makes sense except in light of evolution," most can conduct

their work quite happily without particular reference to evolutionary ideas. “Evolution” would appear to be the indispensable unifying idea and, at the same time, a highly superfluous one.[55]

Perhaps the reader will recognize from the preceding examples that to assume all one needs to know about an organism is contained in its genome is an unsatisfactory way to study biology. The much anticipated completion of sequencing the human genome—and of many other genomes—has only revealed that life is more complex than the previously dominant gene-oriented evolutionary theory led scientists to believe. Biologists are now increasingly turning to a systems approach to study biology, using, for example, the concepts of engineering and design.[56] There is good reason to believe that this trend will continue as the 21st century progresses. In the view of this author, modern science makes it possible to be a scientifically informed doubter of Darwinian theories of evolution.

Acknowledgement

I have been fortunate to hear lectures by, receive reports of advances in research from, and to have discussions with many leading scientists and colleagues in the disciplines relevant to this topic. Some of their work is referenced here, along with that of many others, though much more could have been included. I appreciate the many opportunities they have given me to learn about key developments in chemistry and biology. However, the interpretations in this essay of the research results are mine, and it should not be assumed that the scientists responsible for a particular discovery would agree with my assessment of its significance for Darwinian evolutionary theory. I also wish to credit the authors of the other essays in this volume for educating me by raising important questions in their writings over the past 15 years. Their work has encouraged me to look closely at the justification for the Darwinian approach in light of the wealth of new knowledge about life that is being produced by the new technologies. Finally, this article is adapted in part from an address given in August 2000 for the American Chemical

Society Division of Analytical Chemistry Award for Distinguished Service in the Advancement of Analytical Chemistry, sponsored by the Waters Corporation. I would like to thank the Division and the sponsors for this honor and for the opportunity to express my ideas about evolution in public for the first time.

References:

Most of the basic concepts in biology discussed in this essay are covered in recent textbooks. The references given here are examples of current developments in biological research selected from a much larger number of similar advances, and offer a place to start reading about a particular topic. In many cases, review articles were selected instead of the original research papers as the reviews will be more easily understood and will give references back to the primary literature. These articles are generally limited to ones that the author has read, but there are other articles that would cover many of the topics equally well.

---

[1] Dyson, F. *Imagined Worlds* (Cambridge, Massachusetts and London: Harvard University Press, 1997), pp. 49ff.

[2] In addition to the articles referenced below, see, for example, R.F. Hirsch, "Analytical Chemistry and the life sciences," *Analytical Chemistry* 73 (2001): 117A; J. Handley, C.M. Harris, "Great ideas of a decade: Analytical chemists recall the birth of what are now key fields of research," *Analytical Chemistry* 73 (2001): 660A–666A; R.A. Keller, et al., "Analytical applications of single-molecule detection," *Analytical Chemistry* 74 (2002): 316A–324A; S.A. Hu, N.J. Dovichi,. "Capillary electrophoresis for the analysis of biopolymers," *Analytical Chemistry* 74 (2002): 2833–2850; D.M. Cannon, Jr., N. Winograd, A.G. Ewing, "Quantitative chemical analysis of single cells," *Annual Review of Biophysics and Biomolecular Structure* 29 (2000): 239–263; "Synchrotron Supplement," *Nature Structural Biology* 5 (August 1998): 614–656.

[3] E. Zubritsky, "How analytical chemists saved the Human Genome Project … or at least gave it a helping hand," *Analytical Chemistry* 74 (2002): 23A–26A.

[4] Resources at The Institute for Genomic Research: http://www.tigr.org/tigr-scripts/CMR2/CMRHomePage.spl; resources at the National Center for Biotechnology Information: for eukaryotes: http://www.ncbi.nlm.nih.gov/PMGifs/Genomes/EG_T.html; for archaea and bacteria: http://www.ncbi.nlm.nih.gov/PMGifs/Genomes/micr.html for plants: http://www.ncbi.nlm.nih.gov/PMGifs/Genomes/PlantList.html

[5] *Teaching about evolution and the nature of science*. (Washington, D.C.: National Academy Press, 1998).

[6] That horizontal gene transfer existed was known for some time (see Syvanen, M., "Molecular Clocks and Evolutionary Relationships: Possible Distortions Due to Horizontal Gene Flow," *Journal Molecular Evolution,* 26 (1987):16-23), but the pervasiveness was obviously unexpected by those who wrote the National Academy of Sciences book.

[7] W.F. Doolittle, "Phylogenetic classification and the universal tree," *Science* 284 (1999): 2124–2128. W.F. Doolittle, "Uprooting the tree of life," *Scientific American* (February 2000): 90–95. C.R.

[8] C.R. Woese, "On the Evolution of Cells," *Proceedings of the National Academy of Sciences (USA),* 99 (2002): 8742–8747. C.R. Woese, "Interpreting the Universal Phylogenetic Tree," *Proceedings of the*

National Academy of Sciences (USA), 97 (2000): 8392–8396. W.F. Doolittle, "The Nature of the Universal Ancestor and the Evolution of the Proteome," *Current Opinion in Structural Biology* 10 (2000): 355–358.
[9] R.L. Charlebois, R.G. Beiko, and Ragan, M.A. "Microbial phylogenetics: Branching out," *Nature* 421 (2003): 217–218.
[10] J. Raymong, O. Zhaxybayeva, J.P. Gogarten, S.Y. Gerdes, and R.E. Blankenship, "Whole-genome analysis of photosynthetic prokaryotes," *Science* 298 (2002): 1616–1620; E. Pennisi, "Bacteria shared photosynthesis genes," *Science* 298 (2002): 1538–1539.
[11] J.O. Andersson, A.M. Sjogren, L.A. Davis, T.M. Embley, and A.J. Roger, "Phylogeneic analyses of diplomonad genes reveal frequent lateral gene transfers affecting eukaryotes," *Current Biology* 13 (2) (2003): 94–104. J.O. Andersson and A.J. Roger, "Evolution of glutamate dehydrogenase genes: evidence for lateral gene transfer within and between prokaryotes and eukaryotes," *BioMed Central Evolutionary Biology* 3 (2003): article 14.
[12] E. Bergthorsson, K.L. Adams, B. Thomason, and J.D. Palmer, "Widespread horizontal transfer of mitochondrial genes in flowering plants," *Nature* 424 (2003): 197–201.
[13] J.M. Archibald, M.B. Rogers, M. Toop, K. Ishida, and P.J. Keeling, "Lateral gene transfer and the evolution of plastid-targeted proteins in the secondary plastid-containing alga *Bigelowiella natans*," *Proceedings of the National Academy of Sciences (USA)* 100 (2003): 7678–7683.
[14] U.L. Rosewich and H.C. Kistler "Role of horizontal gene transfer in the evolution of fungi," *Annual Review of Phytopathology* 38 (2000): 325–363.
[15] E.H. Scholl, J.L. Thorne, J.P. McCarter, D. McK. Bird, "Horizontally transferred genes in plant-parasitic nematodes: a high-throughput genomic approach," *Genome Biology* 4 (2003): paper R39.
[16] A. Mira, H. Ochman, and N.A. Moran, "Deletional bias and the evolution of bacterial genomes," *Trends in Genetics* 17 (2001): 589–596.
[17] C. Dennis, "Mouse Genome: A Forage in the Junkyard," *Nature* 420 (2002): 458–459.
[18] J.A. Shapiro, "Repetitive DNA, Genome System Architecture and Genome Reorganization," Research in Microbiology 153 (2002): 447–453.
[19] M. Szymanski and J. Barciszewski, "Beyond the Proteome: Non-Coding Regulatory RNAs," *Genome Biology* 3 (2002): 0005.1–0005.8.
[20] J. Hasty, D. McMillen, and J.J. Collins,. "Engineered gene circuits," *Nature* 420 (2002): 224–230.
[21] E.H. Davidson, D.R. McClay, and L. Hood, "Regulatory gene networks and the properties of the developmental process," *Proceedings of the National Academy of Sciences (USA)* 100 (2003) 1475–1480; http://www.its.caltech.edu/~mirsky/endomes.htm
[22] C. Dennis, "Altered States," *Nature* 421 (2003): 686–688. A.P. Feinberg, "Cancer epigenetics takes center stage," *Proceedings of the National Academy of Sciences (USA)* 98 (2001):392–394. M.A. Goldman, "The Epigenetics of the Cell," *Genome Biology* 4 (2003): 309. J.A. Shapiro, "Genome Organization and Reorganization in Evolution: Formatting for Computation and Function," *Annals of the New York Academy of Sciences* 981 (2002): 111–134.
[23] Y.O. Chernoff, "Mutation processes at the protein level: is Lamarck back?" *Mutation Research/Reviews in Mutation Research* 488 (2001): 39–64.
[24] W.S. Hancock, "The Challenges Ahead," *Journal of Proteome Research* 1 (2002): 9. D.F. Hunt "Personal Commentary on Proteomics," *Journal of Proteome Research* 1 (2002): 15–19. R. Aebersold, "Constellations in a Cellular Universe," *Nature* 422 (2003): 115–116. B. Marte, "Proteomics," *Nature* 422 (2003): 191, and the review articles that follow.
[25] M.S. Lipton, et al., "Global analysis of the Deinococcus radiodurans proteome by using accurate mass tags," *Proceedings of the National Academy of Sciences (USA)* 99 (2002): 11049–11054. Y. Shen, et al., "Packed Capillary Reversed-Phase Liquid Chromatography with High-Performance Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry for Proteomics," *Analytical Chemistry* 73 (2001): 1766–1775. J. Rappsilber and M. Mann, "Is Mass Spectrometry Ready for Proteome-wide Protein Expression Analysis?" *Genome Biology* 3 (2002), comment 2008.
[26] H.J. McCune and A.D. Donaldson, "DNA Replication: Telling Time with Microarrays," *Genome Biology* 4 (2002), 204. R.W. Ye, T. Wang, L. Bedzyk, and K.M. Croker, "Applications of DNA Microarrays in Microbial Systems," *Journal of Microbiological Methods* 47 (2001), 257–272. M.T. Laub, H.H. McAdams, T. Feldblyum, C.M. Fraser, and L. Shapiro, "Global Analysis of the Genetic Network Controlling a Bacterial Cell Cycle," *Science* 290 (2000): 2144–2148. A.T. Revel, A.M. Talaat, and M.V.

Norgard, "DNA Microarray Analysis of Differential Gene Expression in *Borrelia burgdorferi*, the Lyme Disease Spirochete," *Proceedings of the National Academy of Sciences (USA)* 99 (2002), 1562–1567.
[27] M.J. MacCoss, et al., "Shotgun identification of protein modifications from protein complexes and lens tissue," *Proceedings of the National Academy of Sciences (USA)* 99 (2002): 7900–7905.
[28] An experimental study is found in E. Rhoades, E. Gussakovsky, and G. Haran, "Watching proteins fold one molecule at a time," *Proceedings of the National Academy of Sciences (USA)* 100 (2003), (published online on February 28, 2003).
[29] P. Wittung-Stafshede, "Role of Cofactors in Protein Folding," *Accounts of Chemical Research* 35 (2002): 201–208.
[30] J. Hou, G.E. Sims, C. Zhang, and S.H. Kim, "A global representation of the protein fold space," *Proceedings of the National Academy of Sciences (USA)* 100 (2003): 2386–2390.
[31] J.W.H. Schymkowitz, F. Rousseau, and L. Serrano, "Surfing on protein folding energy landscapes," *Proceedings of the National Academy of Sciences (USA)* 99 (2002): 15846–15848.
[32] S.M. Uptain and S. Lindquist, "Prions as protein-based genetic elements," *Annual Reviews of Microbiology* 56 (2002): 703-741
[33] R.D. Vale, "The Molecular Motor Toolbox for Intracellular Transport," *Cell* 112 (2003), 467–480.
[34] B. Alberts, "The cell as a collection of protein machines: Preparing the next generation of molecular biologists," *Cell* 92 (1998): 291–294.
[35] A. Yonath, "The search and its outcome: High-resolution structures of ribosomal particles from mesophilic, thermophilic, and halophilic bacteria at various functional states," *Annual Reviews of Biophysics and Biomolecular Structure* 31 (2002): 257–273. V. Ramakrishnan, "Ribosome structure and the mechanism of translation," *Cell* 108 (2002): 557–572. J. Frank, "The ribosome—a macromolecular machine par excellence," *Chemistry & Biology* 7 (2000): R133–R141. H.F. Noller and A. Baucom, "Structure of the 70S ribosome: implications for movement," *Biochemical Society Transactions* 30 (2002): 1159–1161. P.B. Moore and T.A. Steitz, "After the ribosome structures: How does peptidyl transferase work?" *RNA* 9 (2003): 155–159. J.A. Doudna and V.L. Rath, "Structure and Function of the Eukaryotic Ribosome: The Next Frontier," *Cell* 109 (2002): 153–156.
[36] J.R. Warner, "Nascent Ribosomes," *Cell* 107 (2001): 133–136.
[37] P. Cramer, D.A. Bushnell, and R.D. Kornberg, "Structural basis of transcription: RNA polymerase II at 2.8 Ångstrom resolution," *Science* 292 (2001): 1863–1876. A.L. Gnatt, P. Cramer, J. Fu, D.A. Bushnell, and R.D. Kornberg, "Structural basis of transcription: An RNA polymerase II elongation complex at 3.3 Ångstrom resolution," *Science* 292 (2001):1876—1882. A. Klug, "A marvelous machine for making messages," *Science* 292 (2001): 1844–1846.
[38] S. Walter and J. Buchner, "Molecular chaperones—cellular machines for protein folding," *Angewandte Chemie International Edition in English* 41 (2002): 1098–1113. F.U. Hartl and M. Hayer-Hartl, "Molecular chaperones in the cytosol: From nascent chain to folded protein," *Science* 295 (2002): 295, 1852–1858. D. Thirumalai and G.H. Lorimer, "Chaperonin-mediated protein folding," *Annual Review of Biophysics and Biomolecular Structure* 30 (2001): 245–269.
[39] R.I. Morimoto, "Dynamic remodeling of transcription complexes by molecular chaperones," *Cell* 110 (2002): 281–284.
[40] P.J. Muchowski, "Protein misfolding, amyloid formation, and neurodegeneration: a critical role for molecular chaperones?" *Neuron* 35 (2002): 9–12.
[41] M. Rouhi, "No pools of free zinc in cells," *Chemical & Engineering News* 79 (September 17, 2001): 53. H.M. Baker, B.F. Anderson and E.N. Baker, "Dealing with Iron: Common Structural Principles in Proteins that Transport Iron and Heme," *Proceedings of the National Academy of Sciences (USA)* 100 (2003): published on-line March 17, 2003
[42] A.C. Rosenzweig, "Copper delivery by metallochaperone proteins," *Accounts of Chemical Research*, 34 (2001): 119–128.
[43] N. Kondo, et al., "Genome fragment of *Wolbachia* endosymbiont transferred to X chromosome of host insect," *Proceedings of the National Academy of Sciences (USA)* 99 (2002): 14280–14285.
[44] M.E. Davey and G.A. O'Toole, "Microbial biofilms: From ecology to molecular genetics," *Microbiology and Molecular Biology Reviews* 64 (2000): 847–867.
[45] R.M. Donal and J.W. Costerson, "Biofilms: Survival mechanisms of clinically relevant microorganisms," *Clinical Microbiology Reviews* 15 (2002): 167–193.
[46] K. Lewis, "Riddle of biofilm resistance," *Antimicrobial Agents and Chemotherapy* 45 (2001): 999–1007.

[47] B.L. Bassler, "Small talk: Cell-to-cell communication in bacteria," *Cell* 109 (2002): 421–424. M.B. Miller and B.L. Bassler, "Quorum sensing in bacteria," *Annual Reviews of Microbiology* 55 (2001): 165–199.
[48] D.L. Erickson, et al., "*Pseudomonas aeruginosa* quorum-sensing systems may control virulence factor expression in the lungs of patients with cystic fibrosis," *Infection and Immunity* 70 (2002): 1783–1790.
[49] P.E. Kolenbrander, et al., "Communication among oral bacteria," *Microbiology and Molecular Biology Reviews* 66 (2002): 486–505.
[50] See for example L.E. Orgel, "The origin of life—a review of facts and speculations," *Trends in Biochemical Sciences* 23 (1998): 491–495. G. Wächterhäuser, "The origin of life and its methodological challenge," *Journal of Theoretical Biology* 187 (1997): 483–694. C. de Duve, *Vital Dust*. New York: Basic Books, 1995.
[51] H. Huber, et al., "A new phylum of Archaea represented by a nanosized hyperthermophylic symbiont," *Nature* 417 (2002): 63–67.
[52] R. Gil, et al., "Extreme genome reduction in *Buchnera* spp.: Toward the minimal genome needed for symbiotic life," *Proceedings of the National Academy of Sciences (USA)* 99 (2002): 4454–4458; R.C.H.J.van Ham, et al., "Reductive genome evolution in *Buchnera aphidicola*," *Proceedings of the National Academy of Sciences (USA)* 100 (2003): 581–586.
[53] C.A. Hutchinson, et al., "Global transposon mutagenesis and a minimal mycoplasma genome," *Science* 286 (1999): 2165–2169.
[54] R.L. Carroll, "Towards a New Evolutionary Synthesis," *TRENDS in Evolution and Ecology* 15 (2000): 27–32; 205–206.
[55] A.S. Wilkins, "Evolutionary processes: a special issue," *BioEssays* 22 (2000): 1051–1052.
[56] L. Hood, "A Personal View of Molecular Technology and How It Has Changed Biology," *Journal of Proteome Research* 1 (2002): 399–409. See also the web site of the Institute for Systems Biology, http://www.systemsbiology.org/.