

10. What Might Be a Protocell's Minimal "Genome?"

Donald E. Johnson

(Ph.D: Computer & Information Science; Ph.D: Chemistry)

Science Integrity, 5002 Holly Tree Road, Wilmington, NC 28409

Abstract. The origin of life's biggest mystery is the origin of the genome which contains the information to cybernetically control all aspects of cellular life today. Without formal control, no life would exist. The genetics-first and metabolism-first models will be examined, each having characteristics that strain scientific credibility. Major physical science limitations and the formidable information science problems are examined. These problems usually result in over-simplifications in speculative scenarios. More serious are the 11 peer-reviewed scientific null hypotheses that require falsification before any of the naturalistic scenarios can be considered as serious science. Assuming the problems can be resolved, the requirements for a minimal "genome" can be discussed in the areas of initial generation of programmed controls, replication of the genome and needed components that make it useful, regulation of "life's" processes, and evolvability. Life is an intersection of the physical sciences of chemistry and physics and the nonphysical formalism of information science. Each domain must be investigated using that domain's principles. Yet most scientists have been attempting to use physical science to explain life's nonphysical information domain, a practice that has no scientific justification.

Introduction: Pseudo-Scientific Speculations or Science?

A hundred years ago, the title's question wouldn't have been needed since a cell was thought to be bag of plasm [1] originating in a "warm little pond" [2]. Fifty years ago, protein and DNA structures had been determined so science "knew" the secrets of the genome. With the Miller/Urey [3] synthesis, many thought that the origin of life explanation was near. Fifteen years ago, it started to be realized that "junk DNA" was a misnomer. Five years ago, epigenetic control systems largely determined by non-coding DNA began to be discovered. As new knowledge of functional complexity is revealed, we realize that our knowledge of that complexity has been increasing exponentially, with no end in sight. As one layer is peeled back, a new level of functional complexity is exposed. Rather than getting simpler, the more we know, the more we know we don't know! "As sequencing and other new technologies spew forth data, the complexity of biology has seemed to grow by orders of Magnitude" [4]. There seems to be an exponential increase in knowledge, with the target of understanding the origin receding ever faster.

The origin of life (OOL) is unknown and is obscured by the lack of knowledge of the prebiotic conditions that existed as life "developed." "Most of the (bio)chemical processes found within all the living organisms are well understood at the molecular level, whereas the origin of life remains one of the most vexing issues in chemistry, biology, and philosophy" [5]. "The origin of life remains one of the humankind's last great unanswered questions, as well as one of the most experimentally challenging research areas" [6]. Any speculation inevitably involves science as we don't know it. It is metaphysically presumed that since life obviously exists, there must have been a time when non-life developed into life through natural mechanisms. It is also presumed (with no substantiating reasons) that Pasteur's Law of Biogenesis, all life is from life ("Omne vivum ex vivo" [7]), must not have been applicable during life's formation from inanimate material. Pasteur's warning that "Spontaneous generation is a dream" ("La génération spontanée est une chimère" [8]) is perhaps appropriate to consider with the various speculations. It is important to realize that "we don't yet know, but the answers will be coming" isn't a scientific statement, but rather expresses faith in naturalism-of-the-gaps, which is no more scientific than the god-of-the-gaps explanation that most scientists would dismiss out-of-hand.

Speculation on a particular aspect of life may not prove fruitful since all known life is a carefully-orchestrated cybernetic system [9-12]. Without consideration of the origin of cybernetic processes, they are “systems and processes that interact with themselves and produce themselves from themselves” [13]. Michael Polanyi argued that life is not reducible to physical and chemical principles, but rather that, “the information content of a biological whole exceeds that of the sum of its parts” [14]. “A whole exists when it acts like a whole, when it produces combined effects that the parts cannot produce alone” [15]. “Understanding the origin of life requires knowledge not only of the origin of biological molecules such as amino acids, nucleotides and their polymers, but also the manner in which those molecules are integrated into the organized systems that characterize cellular life” [16].

It should be noted that speculation is important within science, since that is the way that new lines of thought are proposed in order to test scenarios for possibility and feasibility [17]. While the dream of becoming a Nobel laureate may encourage wide dissemination of a speculation, it seems appropriate to warn about spreading such speculations outside the scientific community. The public too often views a scientist’s speculation as validated science, so that the speculative nature is overlooked. The public may value a scientist’s view in much the same way that a movie star’s endorsement of a product is seen as important. There seems to be a wide-spread belief in chemical predestination, even though its chief promoter [18] has repudiated its possibility. For example, when signs of water on Mars were discovered, the media proclaimed that there must be life then. Our collective preoccupation with the Search for Extraterrestrial Intelligence [19] illustrates the belief in the inevitability of life.

1. Overview

The approach of this essay will be to consider scenarios for developing the minimal replication and control information (“proto-genome”) for a protocell, since even “protolife” would require self-replication and control ability. Note that the ability to use the “genomic” information for functionality is also critical. Metabolic cycles [20], homochirality [21-24], cell membranes [25-26], and other required components will not be the primary thrust, even though all are

indirectly controlled by today's genome. An excellent review of the organic chemistry for biomolecular origin is available [27]. Each proponent's scenario will be briefly highlighted, with the primary arguments against the scenario coming from proponents of an alternative scenario, typically as quotes. Finally, we'll examine principles that are almost universally ignored in OOL scenarios, but are in critical need of scientific explanation.

1.1 RNA (Genetics) First Scenarios

A ribosome, "a molecular fossil" [28], can join amino acids without additional enzymes except for those that are imbedded in the ribosome itself to make it a ribozyme (enzymes needed to manufacture tRNAs presumably developed later). "An appeal of the RNA world hypothesis is that it solves the 'chicken and egg' problem; it shows that in an earlier, simplified biota the genotype/replicator and phenotype/catalyst could have been one and the same molecule" [29] (but the RNA/enzyme of a ribozyme is another chicken/egg problem). "RNA appears well suited to have served as the first replicative polymer on this planet" [30]. The origin of the RNA World by stringing together optimistic extrapolations of prebiotic chemistry achievements and experimenter-directed RNA "evolution" (a misnomer) has been described as "the 'Molecular Biologists' Dream ... [and] the prebiotic chemist's nightmare" [31]. An excellent review [32] describes the potential and problems of the RNA world. The "difficulties in nucleobase ribosylation can be overcome with directing, blocking, and activating groups on the nucleobase and ribose ... These molecular interventions are synthetically ingenious, but serve to emphasize the enormous difficulties that must be overcome if ribonucleosides are to be efficiently produced by nucleobase ribosylation under prebiotically plausible conditions. This impasse has led many scientists to abandon the idea that a RNA "genome" might have assembled abiotically, and has prompted a search for potential pre-RNA informational molecules" [33]. It has been pointed out that "what is essential, therefore, is a reasonably detailed description, hopefully supported by experimental evidence, of how an evolvable family of cycles might operate. The scheme should not make unreasonable demands on the efficiency and specificity of the various external and internally generated catalysts that are supposed to be involved. Without such a description, acceptance of the possibility of complex non-enzymatic cyclic organizations that

are capable of evolution can only be based on faith, a notoriously dangerous route to scientific progress” [20]. The experimenter-directed “side products would have amounted to a fatal and committed step in the synthesis of a nascent proto-RNA. This problem illustrates a difficulty in non-enzymatic polymerization that must be taken into account when considering how the nature of the synthetic routes to and structural identities of early genetic polymers: irreversible linkages are adaptive for an informational polymer only when mechanisms exist to make them conditionally reversible [34].

“No physical law need be broken for spontaneous RNA formation to happen, but the chances against it are so immense, that the suggestion implies that the non-living world had an innate desire to generate RNA

There is no reason to presume that an indifferent nature would not combine units at random, producing an immense variety of hybrid short, terminated chains, rather than the much longer one of uniform backbone geometry needed to support replicator and catalytic functions” [35]. “The RNA molecule is too complex, requiring assembly first of the monomeric constituents of RNA, then assembly of strings of monomers into polymers. As a random event without a highly structured chemical context, this sequence has a forbiddingly low probability and the process lacks a plausible chemical explanation, despite considerable effort to supply one” [36]. “It has been challenging to identify possible prebiotic chemistry that might have created RNA. Organic molecules, given energy, have a well-known propensity to form multiple products, sometimes referred to collectively as ‘tar’ or ‘tholin.’ These mixtures appear to be unsuited to support Darwinian processes, and certainly have never been observed to spontaneously yield a homochiral genetic polymer. To date, proposed solutions to this challenge either involve too much direct human intervention to satisfy many in the community, or generate molecules that are unreactive ‘dead ends’ under standard conditions of temperature and pressure” [27].

Some [27, 33] believe that inorganic crystals or clay served as a template for the original RNA. The “replication of clay ‘information’ has remained hypothetical, and transfer of replicated clay properties to nucleic acids even more so” [29]. Crystals contain a very small quantity of information in their regular structures, so that any significant information would have to be in irregularities. How would

inanimate nature produce those irregularities to serve as templates for functional information in replicative polymers?

“The reaction system... is a purified reconstituted system in which all of the components and their concentrations are defined. The number of components is amazingly large, yet this is one of the simplest encapsulated systems for carrying out protein translation and RNA replication. With regard to the origin of life, the first living systems would have had functionally identical translation and replication systems, but they must have been simpler and contained machinery for nutrient transport. The complexity of our system implies that extant translation machinery has become highly sophisticated during the evolutionary process” [16].

1.2 Metabolism-First Scenarios

Metabolism-first scenarios involve development of a self-replicating, self-sustaining chemical system that is able to capture energy and that is contained within a protocell [24] or geothermal vent [38-39]. Perhaps energy transfer used an “osmosis first” paradigm [40, 26]. Unlike RNA first, there is no nucleotide genome to control replication or component construction so that selection would have favored “not the best replicator, but the reaction that sucked in fuel the quickest, denying energy to other chemical processes” [41]. The “bag of chemicals” (composome) presumably would grow until it reaches a size that enables it to divide, with each “daughter” inheriting about half the chemical contents. “The origin of life was marked when a rare few protocells happened to have the ability to capture energy from the environment to initiate catalyzed heterotrophic growth directed by heritable genetic information in the polymers ... The origin of life occurred when a subset of these molecules was captured in a compartment and could interact with one another to produce the properties we associate with the living state” [39]. There have been simulations [42-43] in which the composomes “undergo mutation-like compositional changes” that are claimed to illustrate evolution, but these have never been experimentally observed.

Although metabolism-first avoids the infeasibility of forming functional RNA by chance, “replication of compositional information is so inaccurate that fitter compositional genomes cannot be maintained by

selection and, therefore, the system lacks evolvability (i.e., it cannot substantially depart from the asymptotic steady-state solution already built-in in the dynamical equations). We conclude that this fundamental limitation of ensemble replicators cautions against metabolism-first theories of the origin of life” [44]. Concerning the chemical cycles required, “These are chemically very difficult reactions ... One needs, therefore, to postulate highly specific catalysts for these reactions. It is likely that such catalysts could be constructed by a skilled synthetic chemist, but questionable that they could be found among naturally occurring minerals or prebiotic organic molecules...The lack of a supporting background in chemistry is even more evident in proposals that metabolic cycles can evolve to ‘life-like’ complexity. The most serious challenge to proponents of metabolic cycle theories—the problems presented by the lack of specificity of most non-enzymatic catalysts—has, in general, not been appreciated. If it has, it has been ignored. Theories of the origin of life based on metabolic cycles cannot be justified by the inadequacy of competing theories: they must stand on their own” [20].

2. Major Unresolved Difficulties

Nearly all scenarios presented as science during this author’s education using the American Chemical Society’s “From Molecules to Man” have been shown to be incorrect by today’s science. Scientists need to use much caution during speculative dreaming about mechanisms that can be considered as explanations for the observations that are currently available. Some of the major difficulties requiring scientific explanation will be highlighted in this section

2.1 Physical Science Limitations

What natural interactions produced homochirality, α -linkage only amino acids, and non-enzymatic peptide bonds and other dehydration reactions in aqueous solutions to produce proteins and RNAs? What physical laws could integrate biochemical pathways and cycles into a formal protometabolic scheme? How did the enzymes required to level life’s 10^{19} range of uncatalyzed reactions [45] spontaneously polymerize and self-assemble?

2.2 Formidable Information Science Problems

“Biological information is not a substance ... biological information is not identical to genes or to DNA (any more than the words on this page are identical to the printers ink visible to the eye of the reader). Information, whether biological or cultural, is not a part of the world of substance” [46]. “All the equations of physics taken together cannot describe, much less explain, living systems. Indeed, the laws of physics do not even contain any hints regarding cybernetic processes or feedback control” [10]. The argument for abiogenesis “simply says it happened. As such, it is nothing more than blind belief. Science must provide rational theoretical mechanism, empirical support, prediction fulfillment, or some combination of these three. If none of these three are available, science should reconsider that molecular evolution of genetic cybernetics is a proven fact and press forward with new research approaches which are not obvious at this time” [47]. “The challenge for an undirected origin of such a cybernetic complex interacting computer system is the need to demonstrate that the rules, laws, and theories that govern electronic computing systems and information don't apply to the even more complex digital information systems that are in living organisms. Laws of chemistry and physics, which follow exact statistical, thermodynamic, and spatial laws, are totally inadequate for generating complex functional information or those systems that process that information using prescriptive algorithmic information” [48].

It is important to realize that data generated by regular fluctuations (such as seasons or light/dark cycles) have extremely low information content, offering no explanation for life's functional information. Communication of information requires that both sender and receiver know the arbitrary protocol determined by rules, not law. A functioning protocell would have needed formal organization, not redundant order. Organization requires control, which requires formalism as a reality [Chapt 1]. Each protein is currently the result of the execution of a real computer program running on the genetic operating system. How did inanimate nature write those programs and operating systems? The genome would be useless without the processing systems needed to carry out its prescriptive instructions.

2.3 Over-Simplification of Information Requirements

“Whatever the source of life (which is scientifically unknowable), the alphabet involved with the origin of life, by the necessary conditions of information theory, had to be at least as symbolically complex as the current codon alphabet. If intermediate alphabets existed (as some have speculated), each predecessor also would be required to be at least as complex as its successor, or Shannon’s Channel Capacity [49] would be exceeded for information transfer between the probability space of alphabets with differing Shannon capacity. Therefore, life’s original alphabet must have used a coding system at least as symbolically complex as the current codon alphabet. There has been no feasible natural explanation proposed to produce such an alphabet since chance or physicality cannot produce functional information or a coding system, let alone a system as complex as that in life” [50]. Coded information has never been observed to originate from physicality. “Due to the abstract character of function and sign systems [semiotics -- symbols and their meaning], life is not a subsystem of natural laws. This suggests that our reason is limited in respect to solving the problem of the origin of life and that we are left accepting life as an axiom... Life express[es] both function and sign systems, which indicates that it is not a subsystem of the [physical] universe, since chance and necessity cannot explain sign systems, meaning, purpose, and goals” [51]. “The reductionist approach has been to regard information as arising out of matter and energy. Coded information systems such as DNA are regarded as accidental in terms of the origin of life and that these then led to the evolution of all life forms as a process of increasing complexity by natural selection operating on mutations on these first forms of life” [52]. “From the information perspective, the genetic system is a pre-existing operating system of unknown origin that supports the storage and execution of a wide variety of specific genetic programs (the genome applications), each program being stored in DNA. DNA is a storage medium, not a computer, that specifies all information needed to support the growth, metabolism, parts manufacturing, etc. for a specific organism via gene subprograms” [50].

There are many features in current life that are extremely difficult to envision as arising from a protocell. The smallest genome (though not autonomous) found so far is in "the psyllid symbiont *Carsonella ruddii*, which consists of a circular chromosome of 159,662 base pairs... The genome has a high coding density (97%) with many overlapping

genes and reduced gene length" [53]. "The origin and evolution of overlapping genes are still unknown" [54]. Since they are prevalent in the simplest known genome, a big question is how and why did overlapping genes arise? Recently, sub-coded information [55] and a second genetic code [56] characterizing alternative splicing have been discovered. Various transcribed RNAs are mixed and matched and spliced into mRNAs for specifying protein construction and other controls. MicroRNAs regulate large networks of genes by acting as master control switches [57]. Tiny polypeptides (with 11-32 amino acids) can function as "micro-protein" gene expression regulators [58]. Were these features required initially, or by what interactions of nature did they arise later?

Scientists are investigating "the organization of information in genomes and the functional roles that non-protein coding RNAs play in the life of the cell. The most significant challenges can be summarized by two points: a) each cell makes hundreds of thousands of different RNAs and a large percent of these are cleaved into shorter functional RNAs demonstrating that each region of the genome is likely to be multifunctional and b) the identification of the functional regions of a genome is difficult because not only are there many of them but because the functional RNAs can be created by taking sequences that are not near each other in the genome and joining them together in an RNA molecule. The order of these sequences that are joined together need not be sequential. The central mystery is what controls the temporal and coordinated expression of these RNAs" [59]. "It is very difficult to wrap your head around how big the genome is and how complicated ... It's very confusing and intimidating ... The coding parts of genes come in pieces, like beads on a string, and by splicing out different beads, or exons, after RNA copies are made, a single gene can potentially code for tens of thousands of different proteins, although the average is about five ... It's the way in which genes are switched on and off, though, that has turned out to be really mind-boggling, with layer after layer of complexity emerging" [60]. When and how did these features arise? Were any present in the first life?

2.4 Scientific Hypotheses Requiring Falsification

In addition to falsifying Shannon Capacity Theorem [49] if a proposed original information system isn't as complex as today's

codon-based system, the following testable null hypotheses (proposed in peer-reviewed papers) may require falsification. No scenario should be accepted as science if it violates one or more of these unfalsified null hypotheses [60-61, 11-12].

- #1 Stochastic ensembles of physical units cannot program algorithmic/cybernetic function.
- #2 Dynamically-ordered sequences of individual physical units (physicality patterned by natural law causation) cannot program algorithmic/cybernetic function.
- #3 Statistically weighted means (e.g., increased availability of certain units in the polymerization environment) giving rise to patterned (compressible) sequences of units cannot program algorithmic/cybernetic function.
- #4 Computationally successful configurable switches cannot be set by chance, necessity, or any combination of the two, even over large periods of time.
- #5 Self-ordering phenomena cannot generate cybernetic organization.
- #6 Randomness cannot generate cybernetic organization.
- #7 PI (prescriptive information [12]) cannot be generated from/by the chance and necessity of inanimate physiodynamics.
- #8 PI cannot be generated independent of formal choice contingency.
- #9 Formal algorithmic optimization, and the conceptual organization that results, cannot be generated independent of PI.
- #10 Physiodynamics cannot spontaneously traverse The Cybernetic Cut [11]: physiodynamics alone cannot organize itself into formally functional systems requiring algorithmic optimization computational halting, and circuit integration.

3. Could a Protocell Live and Reproduce Without a “Genome?”

Assuming that the problems highlighted in the previous sections can be overcome (including falsifications of 2.4), this section will discuss the key topic of this essay. The protocell will be assumed to have an appropriate boundary (membrane, microtubule, etc.) that separates the “living” protocell from its environment. This section

will highlight what would be required of a “proto-genome,” without regard as to whether such a “genome” is feasible (not operationally falsified). “There seems to be little general agreement as to how the molecular apparatus needed to implement genetics within a cell could have come about. In fact, there seems to be nothing but puzzlement on such questions with virtually no chemically founded suggestions being made at all” [63]. We will be examining the functional requirements of the proto-genome, as opposed to hypothetical implementations. A proto-genome may have little resemblance to today’s DNA-based genome since it will be assumed that life’s origin didn’t involve DNA. Consequently, we will be attempting to examine life as we don’t know it, an exercise that should always be accompanied by healthy scientific skepticism.

It is important to realize that John von Neumann proposed and proved the requirements for a self-replicating automaton long before the discovery of DNA’s information [64]. A self-reproducing system must contain the necessary components of any computer system, as well as the program for its own construction with the hardware needed to accomplish that construction.

3.1 Replication Requirements

A mechanism is needed to divide the protocell into two approximately equal daughters with each daughter being capable of growth and eventual division for exponential population potential. The “proto-genome” with its processing system must replicate itself, along with all cellular controls (functional information and senders/receivers/processors) into each daughter. Unless the “proto-genome” has replisome capabilities included in the “proto-genome” rather than a separate enzyme, the self-contained capability is required to duplicate all other needed components for “life” with high fidelity. Each daughter also needs a replicated (or split) cell boundary.

Science knows that the current replication hardware and software requires all the components to be fully functional for replication to occur at all. All known errors during replication result in a decrease of both Shannon and functional information [65], usually producing a cell that is no longer able to reproduce. Reliable replication is fundamental to life, a characteristic lacking in composites [44].

3.2 Control Requirements

Controlled chemical metabolic networks are needed that can selectively admit “fuel” (redox, heat, photons, etc.) into the cell and process the “fuel” to harness the energy for growth, reproduction, manufacturing of needed components that can’t migrate in, and other useful work. Both sender and receiver of the each control signal are needed, along with knowledge of the protocol rules for correct communication. The manufacturing control for needed cellular components would probably require enzymatic functionality for polymerization, along with producing homochiral components. In addition, control is required for cell division. Without control, organization (as opposed to self-ordering) is impossible, and functionality would disintegrate, with “tar” a likely result.

Cellular control is a cybernetic process, so all of the requirements of the first eight chapters need instantiation into the protocell. While the proto-genome may contain the control instructions, those instructions must be read by other components (unless the proto-genome has expanded capabilities so that it can read itself), and communicated reliably (using “agreed upon” arbitrary protocols between sender and receiver, source and destination) to the components effecting the control operations. This is not an easily-dismissed prerequisite since control in known life is critical to make the chemical components “alive.” In addition, mere physicydynamic constraints cannot generate formal biological controls [66].

3.3 Evolvability Requirements

The system would have to be capable of accurate duplication, but capable of gradual changes that would permit evolution to life-as-we-know-it. A robust information structure that can be self-maintained (including error-correction), such as in a long genetic polymer, would be required. The feasibility of formation of such a polymer has yet to be shown with any prebiotic mixture proposed to date. The enzyme- and template-independent 120-mer polymers recently generated in water at high temperatures [67] are non-informational homopolymers similar to those adsorbed onto montmorillonite clay surfaces [68]. The aqueous polymers are also cyclic and require some experimenter engineering to achieve 120 mer length.

The proto-genome would also need to be able to effect highly accurate duplication of the entire proto-cell, with only an occasional “error” that could produce a very similar proto-cell, still possessing all three of the requirements in section 3. The proto-genome, along with all the proto-cell components, would need to have a feasible path to eventually produce cells with the functional complexity of today’s life. It does little good to speculate a “simple” initial system unless there are feasible scenarios that can traverse from the proposed initial system to life as we know it, including coded information and other features highlighted previously. For example, one could envision dipping a finger into a bottle of ink and flicking the ink toward a white sheet would eventually produce a pattern that looks like an English letter. That would not explain the formal rules and meaningful syntax of letters that you are currently observing in this book, however.

4. Conclusions

While scenarios for the first cell can be envisioned purely from physicality, a “proto-genome” introduces cybernetic aspects that can have no origination from inanimate material. In particular, organization, prescriptive information, and control require traversing The Cybernetic Cut on a one-way CS (Configurable Switch) Bridge [11] that allows traffic only from formalism to physicality. Just as formalism needs recognition as reality, it is also critical to recognize the limits of physical science, such as physics and chemistry, whose spontaneous inanimate mass/energy interaction behavior is constrained by laws, not formal controls. Initial starting constraints chosen by an experimenter become controls for an experiment, but those chosen constraints are instantiations into physicality of nonphysical formalisms.

Life is an intersection of physical science and information science. Both domains are critical for any life to exist, and each must be investigated using that domain's principles. Yet most scientists have been attempting to use physical science to explain life's information domain, a practice that has no scientific justification. Since the chemistry and physics of life are controlled by prescriptive information (not just constrained by laws), biology is really an information science, not a physical science.

References

1. Haeckel E 2004, *The Evolution of Man*, Translation (original 1911), p49.
2. Darwin C 1871,
<http://darwin-online.org.uk/content/frameset?viewtype=side&itemID=F1452.3&pageseq=30>
3. Miller SL 1953, "Production of Amino Acids Under Possible Primitive Earth Conditions," *Science*: 117, p528.
4. Hayden E 2010, "Human Genome at Ten: Life is Complicated," *Nature* 464, 3/31/10, p664-667.
5. Stochel G, Brindell M, Macyk W, Stasicka W, & Szacilowski K 2009, *Bioinorganic Photochemistry* (West Sussex, England: John Wiley & Sons.), p109.
6. Miller SL & Cleaves HJ 2007, "Prebiotic Chemistry on the Primitive Earth" in Isidore Rigoutsos and Gregroy Stephanopoulos, eds., *Systems Biology Volume 1: Genomics* (New York: Oxford University Press), 3.
7. *Encyclopaedia Britannica* 1902, 10th Edition.
8. Pasteur L 1862, <http://www.pasteur.fr/ip/easysite/go/03b-00000j-0e7/institut-pasteur/histoire/>
9. Corning P & Kline S 2000, "Thermodynamics, Information and Life Revisited, Part I: to Be or Entropy," *Systems Research*, 4/7/00, p273-295.
10. Corning P 2005, *Holistic Darwinism*, p330.
11. Abel DL 2008, "The 'Cybernetic Cut': Progressing from Description to Prescription in Systems Theory," *The Open Cybernetics and Systemics Journal* (2), p252-262.
12. Abel DL 2009, "The Biosemiosis of Prescriptive Information," *Semiotica*, 1/4/09, p1-19.
13. Kauffman L 2007, *CYBCON* discussion group 18, 9/20/07, p15.
14. Polanyi M, Quote originally at Michael Polanyi Center site, now at nostalgia.wikipedia.org/wiki/Michael_Polanyi
15. Corning P 2002, "The Re-emergence of 'Emergence': A Venerable Concept in Search of a Theory," *Complexity* 7(6), p18-30.
16. Ichihashi N, Matsuura T, Kita H, Sunami T, Suzuki H, & Yomo T 2010, "Constructing Partial Models of Cells," *Cold Spring Harb Perspect Biol*, 4/28/10, 2:a004945.
17. Abel DL 2009, "The Universal Plausibility Metric (UPM) & Principle (UPP)," *Theoretical Biology and Medical Modelling*, 12/3/09, 6:27.
18. Kenyon D 1969, *Biochemical Predestination*.
19. SETI Website, setiathome.ssl.berkeley.edu/
20. Orgel L 2008, "The Implausibility of Metabolic Cycles on the Prebiotic Earth," *PLoS Biol* 6(1), e18.
21. Toxvaerd S 2009, "Origin of Homochirality in Biosystems," *Int. J. Mol. Sci.*, 3/18/09, p1290-1299.
22. Blackmond D 2010, "The Origin of Biological Homochirality," *Cold Spring Harb Perspect Biol*, 3/10/10, 2:a002147.
23. Lee T & Lin Y 2010, "The Origin of Life and the Crystallization of Aspartic Acid in Water," *Crystal Growth & Design* 10 (4), p1652-1660.
24. Chown M 2010, "Did Exploding Stars Shatter Life's Mirror?," *New Scientist*, 5/19/10.
25. Masy S 2010, "Membrane Transport in Primitive Cells," *Cold Spring Harb Perspect Biol*, 4/21/10 doi: 10.1101/cshperspect.a002188.
26. Lane N, Allen J, & Martin W 2010, "How Did LUCA Make a Living? Chemiosmosis in the Origin of Life," *BioEssays* 32, p271-280.
27. Benner S, Kim H, Kim M, & Ricardo A 2010, "Planetary Organic Chemistry and the Origins of Biomolecules," *Cold Spring Harb Perspect Biol*, 6/1/10.
28. Johnson D & Wang L 2010, "Imprints of the Genetic Code in the Ribosome," *The Scripps Research Institute*, 3/18/10.

29. Yarus M 2010, "Getting Past the RNA World: The Initial Darwinian Ancestor," 4/21/10, Cold Spring Harb Perspect Biol doi: 10.1101.
30. Cheng L & Unrau P 2010, "Closing the Circle: Replicating RNA with RNA," Cold Spring Harb Perspect Biol, 6/16/10, doi: 10.1101/cshperspect.a002204.
31. Joyce G & Orgel L 1999, "Prospects for Understanding the Origin of the RNA World," in *The RNA World*, 2nd ed., p49–77.
32. Orgel L 2004, "Prebiotic Chemistry and the Origin of the RNA World," *Crit. Rev. in Biochemistry and Molecular Biology*: 39, p99–123.
33. Sutherland J 2010, "Ribonucleotides," Cold Spring Harb Perspect Biol, 3/10/10, 2:a005439.
34. Engelhart A & Hud N 2010, "Primitive Genetic Polymers," Cold Spring Harb Perspect Biol, 5/12/10, doi: 10.1101/cshperspect.a002196.
35. Shapiro R 2007, "A Simpler Origin for Life," *Scientific American*, 2/12/07.
36. Trefil J, Morowitz H, Smith E 2009, "The Origin of Life," *American Scientist* 97(2), 3-4/09, p206-210.
37. Hazen R & Sverjensky D 2010, "Mineral Surfaces, Geochemical Complexities, and the Origins of Life," Cold Spring Harb Perspect Biol, 4/14/10, 2:a002162.
38. Schrum J, Zhu T, & Szostak J 2010, 5/19/10, "The Origins of Cellular Life," Cold Spring Harb Perspect Biol doi: 10.1101/cshperspect.a002212.
39. Deamer D & Weber A 2010, "Bioenergetics and Life's Origins," Cold Spring Harb Perspect Biol, 1/13/10, 2:a004929.
40. Luisi, P, Ferri F, et al. 2006, "Approaches to semi-synthetic minimal cells: a review," *Naturwissenschaften* 93, p1-13.
41. Russell M 2009, as reported by John Whitfield in "Nascence man," *Nature* 459, 5/21/09.
42. Segre D, Ben-Eli D, & Lancet D 2000, "Compositional Genomes: Prebiotic Information Transfer in Mutually Catalytic Noncovalent Assemblies," *PNAS* 97 (8), 4/11/00, p4112–4117.
43. Shenhav B, Oz A, & Lancet D 2007, "Coevolution of Compositional Protocells and Their Environment," *Phil. Trans. R. Soc. B* 362, p1813-1819.
44. Vasas V, Szathmáry E, & Santos M 2010, "Lack of Evolvability in Self-sustaining Autocatalytic Networks Constrains Metabolism-first Scenarios for the Origin of Life," *PNAS*, 2/11/10, www.pnas.org/content/107/4/1470.
45. Lad C, Williams N, & Wolfenden R 2003, "The Rate of Yydrolisis of Phosphomonoester Dianions and the Exceptional Catalytic Proficiencies of Protein and Inositol Phosphatases," *PNAS*: 100 (10), 5/13/03, p5607-5610.
46. Hoffmeyer J & Emmeche C 2005, "Code-Duality and the Semiotics of Nature," *J. Biosemiotics* (1), p37-64.
47. Trevors JT & Abel DL 2004, "Chance and Necessity Do Not Explain the Origin of Life," *Cell Biology International*: 28, p729-739.
48. Johnson DE 2009/2010, *Probability's Nature and Nature's Probability: A Call to Scientific Integrity*, p48.
49. Shannon C 1948, "A Mathematical Theory of Communication," *Bell System Technical Journal*: 27, July & October, p379-423 & 623-656.
50. Johnson DE 2010, *Programming of Life*, p37&49.
51. Voie A 2006, "Biological Function and the Genetic Code are Interdependent," *Chaos, Solitons and Fractals*: 28(4), p1000-1004.
52. McIntosh A 2009, "Entropy, Free Energy and Information in Living Systems," *International Journal of Design & Nature and Ecodynamics* 4 (4), p351-385.
53. Nakabachi A, Yamashita A, Toh H, Ishikawa H, Dunbar H, Moran N, & Hattori M 2006, "The 160-Kilobase Genome of the Bacterial Endosymbiont Carsonella," *Science*: 314 (5797), 10/13/06, p267.
54. Veeramachaneni V, Makalowski W, Galdzicki M, Sood R, & Makalowska I 2004, "Mammalian Overlapping Genes: The Comparative Perspective," *Genome Res.*: 14, p280-286.

55. Cannarozzi G, Schraudolph N, Faty M, von Rohr P, Friberg M, Roth A, Gonnet P, Gonnet G, & Barral Y 2010, "Role for Codon Order in Translation Dynamics," *Cell* 141, 4/16/10, p355-367.
56. Barash Y, Calarco J, Gao W, Pan Q, Wang X, Shai O, Blencowe B, & Frey B 2010, "Deciphering the Splicing Code," *Nature*, 5/6/10, p53-9.
57. Lieberman J 2010, "Master of the Cell," *The Scientist*, 4/1/10, p42.
58. Kondo T, Plaza S, Zanet S, Benrabah E, Valenti P, Hashimoto Y, Kobayashi S, Payre F, Kageyama Y 2010, "Small Peptides Switch the Transcriptional Activity of Shavenbaby During *Drosophila* Embryogenesis," *Science* 329, 7/16/10, p336-339.
59. Gingeras T, www.desdeexilio.com/2010/06/28/epigenetica-entrevista-a-thomas-gingeras/
60. Le Page M 2010, "Genome at 10: A Dizzying Journey into Complexity," *New Scientist*, 6/16/10.
61. Abel D & Trevors J 2005, "Three subsets of sequence complexity and their relevance to biopolymeric information," *Theoretical Biology and Medical Modelling*, 8/11/05, 2:29.
62. Abel D & Trevors J 2006, "Self-Organization vs Self-Ordering Events in Life-Origin Models," *Physics of Life Reviews* (3), p211-228.
63. Yana K, Fanga G, Bhardwaja N, Alexander R, & Gerstein M 2010, "Comparing Genomes to Computer Operating Systems in Terms of the Topology and Evolution of Their Regulatory Control Networks," www.pnas.org/cgi/doi/10.1073/pnas.0914771107.
64. von Neumann J, *Theory of Self-Reproducing Automata*, 1966.
65. Abel DL 2009, "The GS (genetic selection) Principle," *Frontiers in Bioscience* (14), 1/1/09, p2959-2969.
66. Abel, DL 2010, "Constraints vs. Controls," *Open Cybernetics and Systemics Journal*, 4, p14-27.
67. Costanzo G, Pino S, Ciciriello F, Di Mauro E 2009, "RNA: Processing and Catalysis: Generation of Long RNA Chains in Water," *J. Biol. Chem.*, 284, p33206-33216.
68. Ferris JP 2002, "Montmorillonite Catalysis of 30-50 Mer Oligonucleotides: Laboratory Demonstration of Potential Steps in the Origin of the RNA World," *Origins of Life and Evolution*

