

The Never-Ending Story—The Origin and Diversification of Life

Marcus Kumala

© Springer Science+Business Media, LLC 2010

Abstract As outlined in the introductory article “The Neverending Story—Using the Narrative as a Fundamental Approach to Teaching Biology and Beyond,” historical storytelling has the potential to add understanding and synthesis to learning that is usually restricted to memorization and skills-based application. The following stories were written in an attempt to reconcile mandated microbiology topics, such as biochemistry and the properties of DNA, with an overarching evolutionary framework that is applicable regardless of the scale of observation. These stories were designed with secondary school biology teachers and students in mind, particularly at the senior-grade level. In an effort to preserve narrative flow, internal referencing has been sacrificed, but a complete set of source material is cited at the end of the compilation.

Keywords Universal narrative · Microbiology · Origin of life · Diversity of life · Co-option · Earth history

Prologue—Food for Thought

We unquestionably have a better understanding of the human body today than at any other moment in history. After all, entire fields have been devoted to deconstructing, fixing, and reassembling a “biological machine” constantly breaking down, or so it would seem. Our mechanics have

done a spectacular job keeping us on the road, but there’s never a lifetime guarantee, and if we were cars, we’d definitely be considered lemons. Perhaps there is a limit to our understanding of the human body when it is treated as an organism without history. Does it matter if you were made in America versus Japan? The parts are undoubtedly similar, but what about fuel efficiency? Are we chimeras—the end product of parts cobbled together and grafted on by the judicious hand of a creator? Or are building materials limited to what’s in stock? It doesn’t seem likely that we’re ever going to come out with a solar model, but why not? Wouldn’t such a design put an end to, say, human hunger? Maybe we’re not so much like machines after all and certain events in the deep past have constrained our futures, but how much so? Sometimes to better understand our present condition, we need to go back to the beginning, and in our case, where better to start than the very beginning.

Chapter I—Life’s Spork

The hardest part of anything is starting. This first sentence, as mundane as it may appear, was my single greatest obstacle to overcome. It was never predestined to succeed, nor can it be attributed to any single causal factor (though writer’s block, sleep deprivation, or something I ate the previous day are among the candidates); written in a different time and place or swapped for a profound quote or question, and the trajectory of this article would have been much different. Yet now written, these surviving keystrokes have set the tone for all the rest. Any major deviation now would destroy whatever continuity has already been established (as tenuous as it may be), and a

M. Kumala (✉)
33 Hillbeck Crescent,
Toronto, ON M1B 2M8, Canada
e-mail: Mkumala@gmail.com

massive rewrite does not appear to be an option. You know, deadlines...

In much the same way, life has been constrained by the events early in its history. While no one can hop in a time machine to literally witness “First Life,” we can still use present-day molecules associated with life as proxies, just as rustling leaves serve as proxies for the wind. And the molecule at the forefront of the origin of life may surprise you: It’s not DNA, but rather its putative ancestor, RNA, ribonucleic acid, the jack-of-all-trades macromolecule.

I’m sure most of you already know the key differences between RNA and DNA: RNA has the nitrogenous base uracil instead of thymine and contains a 2’ hydroxyl group instead of hydrogen. At first glance, these differences may seem rather trivial, but they would prove to be vital in establishing the living world. Briefly, the hydroxyl group makes RNA much more chemically reactive and structurally flexible than DNA, and thus it exists in a multitude of forms: stems, loops, and pins, circular, single-stranded, and double-stranded...sometimes all in one molecule! (Recall the structure of rRNA and tRNA for starters.) Consequently, some RNA molecules possess both secondary and tertiary organization reminiscent of amino acid chains (polypeptides), the basis of proteins. It should come as no surprise that these RNAs are capable of initiating some pretty remarkable processes. For example, “hammerheads” can repeatedly cut new polynucleotide substrates, and RNase P is capable of processing pre-tRNAs to mature tRNAs! Others add or subtract nucleotides, cleave introns, ligate exons, and some even replicate. They are often called “replicases” or ribozymes as a result of this dual ability to replicate and catalyze.

RNA is a robust retort to the teleological notion that “form follows function” because it is a molecular multi-tool, the “spork” to the classic chicken or the egg or fork–spoon scenario perpetuated by the horribly coined “Central Dogma” of molecular biology. You know, the one that goes “Information flows from DNA to RNA to proteins, but if proteins are required to make DNA, which came...” It is not difficult to envision a time when life (or something very close to it) consisted of diverse assemblages of RNA molecules, or “selfish co-operators” if you will¹: splicers, binders, replicators, and even manufacturers. The ribosome is after all fundamentally a ribozyme! Although proteins stabilize and increase the efficiency of modern-day ribosomes, rRNA is the catalyst responsible for forming the peptide bonds that link amino acids together.

Some RNAs were probably capable of manufacturing polypeptide chains in a process predating the later

ribosomal complex, likely utilizing an early version of the genetic code. The *near*, and not absolute, universality of the genetic code is also proof of its potential malleability and lends credence to a time when a common language was still being worked out. Regardless, it’s becoming a lot clearer that RNA was once a large piece of the life puzzle,² and classic textbook *translation* and an early genetic code probably existed well before DNA and DNA replication and transcription. Polypeptides, followed by full proteins, gradually improved the catalytic efficiency of early life, setting the tone for a division of labor where enzymes took on a prominent role as catalysts and DNA eventually replaced RNA as the more stable hereditary template. Nevertheless, RNA is still more than just an intermediary; ribozymes sit at the heart of many processes essential to life as we know it. They represent evolutionary baggage that life was never able to put down, the “molecular scaffolding” left behind. But this still doesn’t explain how you arrive at RNA in the first place. ☺

Chapter II—Soup that Eats Like a Meal!

So how did we arrive at life’s “spork”? For the better part of the last thirty years, scientists from many disciplines have rigorously probed the limits of human resolution to discover not only the recipe for life but also the likely location of its original pot.

Early attempts at discovering a stepwise transition from a prebiotic to a living world emulated a common aspect of bachelor life: raid the fridge, throw whatever you find in a pot, heat it up, and see if it’s edible. One classic experiment seemingly tore a page out of *Frankenstein* by exposing a volatile mixture of gases to lightning and UV radiation—thought to be prominent features on early Earth—and this rather astoundingly produced amino acids, as well as several other precursors to contemporary biomolecules. These fell like manna from heaven to create a “primordial soup” below, imagery that has been difficult to move beyond due to a continued fixation with all things Darwin and the mystique of a “warm little pond.”³ Subsequent research showed promising results once intermediate molecules (i.e., formic acid, formaldehyde, methanol, etc.) were shielded or “buried” from the creative and destructive forces of their birth. It was thought that under certain conditions, these precursors would accumulate and form unique combinations with nitrogen, phosphorus, and sulfur by virtue of their reactive functional groups (i.e., hydroxyls

¹ Early life is more aptly characterized by oxymorons than outright “selfish” labels, which would imply at least a nervous system capable of processing the concept of selfishness, and that won’t appear for a LONG time.

² In fact, uracil requires less energy to make than thymine and actually forms from the degradation of cytosine.

³ Give the man a break, he’s been dead for over a century and is allowed to have been wrong at some point in his life!

and carboxyls), inevitably giving rise to macromolecules like amino acids, fatty acids, nitrogenous bases, and pentose sugars. Over time, these monomers would “self-organize”—a phrase unfortunately wrought with a collective inability to intelligibly explain relevant biochemical interactions to the next generation (myself once a victim and for the time being an ignorant perpetrator)—eventually yielding arguably the most important polymer, RNA. The cumulative results of tinkering with the life recipe would collectively come to be known as “The RNA World Hypothesis” (later, “genes-first”). However, it wasn’t long before rival chemists would issue the first of “Iron Chef” challenges.

It has been known for some time now that the Earth of four billion years ago was both anoxic and reducing, devoid of the oxidizing effects of free oxygen and filled with the noxious fumes of enhanced volcanic activity (i.e., CO, CO₂, CH₄, NH₃, SO₂, H₂S, HCN, etc.). Accordingly, chemical combinations and interactions vital to the emergence of life would have been much more prevalent (and unconsumed), which also helps to explain why *abiogenesis* does not occur today. While scientists are generally in agreement on the key ingredients, the exact proportions and cooking instructions are still far from standardized, and each chef would love to add their own “special” seasoning. So rather than focus on life’s replicative origins, others decided to plunge into the depths of hell to try and explain how life may have overcome its initial energy constraints.

Though Earth appears to be a highly complex and organized form of matter, it was born out of chaos, fire, and brimstone and is still fuming at its core. A primordial heat—the accumulated energy transfers from planetesimal collisions, the sinking of iron, and the continued radioactive decay of isotopes—drives the tectonic forces that reshape our planet (i.e., volcanoes, earthquakes, continental drift, etc.), a phenomenon active only on Earth or at least when considering our present solar system. Coincidentally or not, life may also be unique and, like our planet, exists as merely localized regions of order in a universe increasing in total entropy. In fact, life may owe its beginning to the slow thermal death of its home.

According to the “Iron–Sulfur World Theory” (later, “metabolism first”), life may have originated in the deep oceans where hydrothermal vents called “black smokers” churn out a rich supply of gases like hydrogen, carbon monoxide, ammonia, hydrogen sulfide, and so forth, and still support rich biotic communities today. Transitional metals like iron, nickel, and manganese can act as inorganic catalysts in the presence of hydrothermal gases, resulting in the formation of more complex organic molecules. For instance, soluble metal cations like Fe²⁺ would have been readily available on early Earth due to the absence of free oxygen and the extreme chemical weathering of rocks, which likely colored the waters green. When iron monosulfide (FeS)

reacts with hydrogen sulfide (H₂S), a pyrite crystal (FeS₂), or “fool’s gold,” forms as well as hydrogen gas, releasing energy in the process. Feedback loops can develop when two molecules regulate their respective synthesis. Since the formation of pyrite is an exergonic reaction, it has the potential to drive endergonic reactions like the reduction of inorganic carbon (i.e., CO or CO₂) to simple organic forms such as formic acid (HCOOH). If any of the products from the initial reaction accelerated the rate of reaction while remaining bound to the growing crystal, a rudimentary carbon-reducing cycle could have emerged under the continuous supply of hydrogen sulfide. With increasingly complex ligands attached to an iron–sulfur core, a collection of catalysts and diverse products may have eventually facilitated the synthesis of a nucleic acid like RNA, well after the establishment of an early form of two-dimensional surface metabolism.

So which theory is correct? Well, if solely taken on the merits of their initial laurels, neither: Ribonucleotides have yet to be produced under any prebiotic conditions deemed universally reasonable, and, at up to 400°C, black smokers might be too chemically extreme to serve as the biosynthetic cradle of life. Though hardly worth mentioning, a third kitchen, “panspermia,” merely proposes to shift the primordial diner to another part of the galaxy where life would still have to overcome the same initial start-up problems before being sent out as samplers even became an option. Nevertheless, it would be as inaccurate to portray these theories as competing, mutually exclusive ideas as it is to set them up as straw men to parody. In science, theories are constantly under scrutiny, re-evaluation, and synthesis, leading to modification over time (or at least they should be). No reputable scientist believes that a pool of organic chemicals spontaneously became a microbe with all the parts of a contemporary bacterium through a simple stirring of the pot.

Scientists working on the life recipe have long recognized the inherently speculative nature of their endeavors due to the scarcity of clues beyond the basic ingredients and available molecular comparisons as a result of the deep time scales. But past researchers were not fools. One cannot compare the chemical structure of, say, an adenine nucleotide, ATP, NAD⁺, and coenzyme A and not recognize that these vital components of life may in fact be the result of common descent from a progenitor pool of organic molecules followed by divergence, nor that the startling similarity displayed represents actual *molecular homology*. It is difficult to likewise ignore the fact that the enzymes found in nearly universal metabolic pathways, as well as those likely to be present in some of the earliest microbes, possess either iron–sulfur, iron–nickel–sulfur, or some other transitional metal cores (i.e., hydrogenases, ferredoxins, NADH dehydrogenase, ubiquinone cytochrome *c* reductase, succinate-ubiquinone reductase,

nitrogenase, heme, etc.). But despite a substantial amount of justification, “genes-first” and “metabolism-first” proponents have all at some point recognized the single greatest obstacle to overcome. Namely, “How in the hell do you get the ‘soup’ in a can to begin with?!”

Without some sort of compartment to generate chemical, electrical, and even thermal gradients, any precursor molecules or protons would diffuse into the surroundings and fail to participate in either polymerization or energy coupling. This has led to some rather interesting proposals and even more amusing vernacular: primordial sandwiches, pizzas, and warm mud pies speak to the volume of ideas, some more plausible than others, that have been used to describe how crucial prebiotic molecules could have clustered and interacted favorably on early Earth. Although no one is suggesting that the first compartment had to be composed of phospholipids arranged in a bilayer, it has proved difficult to demonstrate exactly how a continued flow of genetic and/or metabolic precursors gained entry into the available abiotically generated amphiphiles. One of the more promising suggestions involves montmorillonite, a common clay, which is known to cause lipid vesicles (liposomes) to assemble, as well as spontaneously link activated ribonucleotides together to form chains of polynucleotides. It has been argued that within such a pool of random RNA molecules, at least some were bound to be catalytic, and a replicating RNA inside a lipid vesicle does in fact cause it to grow. That’s almost alive, right? Though this still doesn’t explain how you get the ribonucleotides in the first place, it may be as close as we’re going to get. Luckily, new discoveries are always waiting to be seated at the head table.⁴

Life originating in deep-sea vents had several things going for it: For one, it shifted the location of the origin of life away from ongoing debates about the “actual” composition of early earth’s atmosphere by virtue of its own reducing microenvironment. More importantly perhaps, it may have offered life a safe haven from the potentially destructive UV rays and the late heavy bombardment of meteors, which otherwise could have delayed the emergence of life until about 3.9 or 3.8 billion years ago. So while some scoured the skies or simply their pots, others combed the deep once again for subtler alternatives with more explanatory power.

They found one in the form of another type of hydrothermal vent system located a few kilometers away from seafloor spread-zones. Hydrothermal “chimneys” churn out a similar chemically rich effluent as their smoker

cousins but at the more hospitable 40–90°C range and, most importantly, in the presence of stacked microcaverns coated with, you guessed it, catalytic iron and nickel monosulfides (FeS and NiS). When hot, reduced hydrothermal fluid meets cool, CO₂-rich ocean water, several interesting reactions take place.

Dissolved iron (Fe²⁺) and nickel (Ni²⁺) will react with hydrogen sulfide and precipitate on contact, forming mineral mounds composed of tiny microcaverns lined with (Fe, Ni)S semi-permeable walls. As hydrothermal fluid is pumped through these freshly formed caverns, the metal sulfides act as three-dimensional catalysts, facilitating the formation of water and acetate (CH₃COOH) from mixing hydrogen (H₂) and carbon dioxide (CO₂). Acetate is discharged as a “waste” product following the hydrolysis of a high-energy thioester bond (~SH) between versatile acetyl-thiol intermediates (i.e., thioester acetyl methylsulfide, CH₃COSCH₃) manufactured earlier in the reaction. Each reaction in the pathway is exergonic, and a considerable amount of energy is available to drive further endergonic reactions. The production of pyruvates, fatty acids, amino acids, and so forth from further carboxylation of acetyl-thiols under realistic hydrothermal conditions makes it entirely plausible that these important macromolecules were once constructed abiotically inside the chimneys of early Earth. And since such reactions would take place in semi-permeable cavities, the resulting products would not simply diffuse into the surroundings.

These geothermal odes to entropy would also have been present on early Earth and possess all the makings of life’s cradle: a constant source of basic building blocks and energy, temperature and chemical variation for diverse partial reactions, the possibility for exchange between molecules of varying complexity (i.e., monomers, oligomers, and polymers), and the potential for successive steps of cellular evolution in a single structure or “community” that continues to grow. Chemoclines and thermoclines would have been crucial to the formation of polycarbon compounds like nucleotides, amino acids, and pentoses, which need lower levels of gases, like hydrogen, since carbon in its most reduced state is not as versatile. Moreover, it’s been recently discovered that single amino acids can actually catalyze the formation of sugars from simple starting materials with almost enzyme-like specificity! Experiments have already demonstrated that amino acids can be generated within hydrothermal vents, which means that metal-sulfide cavities would not have been the only catalysts available on early Earth. In this way, microcaverns could essentially have functioned as inorganic cells, delaying the necessity for organic *cellularization* at the earliest stages of chemical evolution. Some have gone so far as to propose that chimneys may have served as “hatcheries” where the three “domains” of life—Bacteria, Archaea, and Eukarya—acquired

⁴ This just in! A group of scientists have recently demonstrated that by combining half-sugars and half-bases with a phosphate buffer/catalyst, you can get pyrimidine ribonucleotides to form abiotically! But other chefs are already lining up to dispute the quality of this kitchen’s ingredients. More to follow I’m sure!

cell membranes independently after their respective divergences. This would help explain why archaeal lipids use ether instead of ester-linkages but otherwise share the same essential features with the rest of life (i.e., common metabolic paths, ATP, DNA, ribosomes, etc.). Others consider this to be a stretch.

Hell's Kitchen would have been more of a crude marshmallow roast than a sophisticated five-star dinner. A ribozyme in cell, albeit an inorganic one, is not the molecular juggernaut we see today, but it's a start. If this latest idea happens to be correct, hydrothermal "chimneys" may have provided not only the can but the self-cooker as well!

Chapter III—Flu Season

Although it is impossible to pinpoint with any degree of certainty the earliest possible date for the emergence of cellular life, there is a general consensus among scientists that a complex microbial community was in place by at least 3.5 billion years ago. Of course, the exact nature of this community is largely an inference drawn from the current observable diversity, trace chemical signatures, and the niches that would have been available at the time; fossil microbes after all tend to superficially look a lot like one another. All things considered, such cellular life would have been much more complicated, both genetically and metabolically, compared with a simple ribozyme in a cell. Then how did we arrive at the molecular model of textbook fame? Taking a page from an old cookbook, though likely at the risk of clumsily perpetuating the same dichotomous sins, let us initially address the *molecular diversification of life* from an information, or "genetics," perspective.

You're probably well aware that all present-day organisms utilize *deoxyribonucleic acid*, or DNA, as the molecular basis for heredity. However, as illustrated in "[Chapter I—Life's Spork](#)", it is very unlikely that DNA was the first "living molecule," and the earliest life forms probably utilized the much more versatile RNA (or perhaps something even simpler). Despite not being a molecular multi-tool, DNA still has its virtues: the absence of a 2' oxygen on the ribose sugar makes DNA much more chemically stable, allowing deoxyribonucleotides (dNTs) to form longer chains via the infamous double helix, thereby vastly increasing the amount of stored information, and thus *potentially* expressed complexity. Our modern DNA–RNA cells also boast added protein features like polymerases and helicases, which not only enhance replicative ability but also facilitate repair.⁵ Although these virtues may help explain why DNA–

RNA cells eventually displaced all their competitors, it would be almost Lamarckian to suggest that this highly integrated and complex replicator evolved directly under gradual selection to increase efficiency and fidelity. While these are clearly the reasons we still have a DNA–RNA cell, this does not necessarily explain why one emerged in the first place. That's the difference between an evolutionary *adaptation* and an *exaptation*, a trait's original or gradually refined role versus its maintenance with a new function. Considering the virtues of DNA without hindsight would be the equivalent of suggesting that opposable thumbs were an adaptation to one day conquer the world of video games. So how did life become this leviathan of replication? Well, this may sound strange, but have you ever had the flu?

Parasites have traditionally fallen victim to our irrational hope that life is continually on the up and up, and thus we frequently rebuke losses instead of considering tomorrow's now possible triumphs. Over the last few years, a somewhat controversial theory has been gaining momentum that would elevate the status of viruses, those infamous intracellular parasites, from life degenerates to the major selective force behind the diversification of life, at least in terms of the flow of biological information. I've taken the liberty of dubbing this the "Primordial Flu Hypothesis," or the virosphere origin for DNA and all its fixings. Basically, the Primordial Flu argues that the complexity of the DNA replicating system emerged largely as a result of intense co-selection between RNA viruses and RNA cells during the later stages of the RNA World. An evolutionary arms race gradually selected for more elaborate means of infection and defense, which led to not only the emergence of DNA and its replicating enzymes but also the origin and diversification of the cell wall and membrane. Under this paradigm, the incredibly diverse form and function of viral genomes—double-stranded (ds) DNA (i.e., pox and herpes viruses), single-stranded (ss) DNA (i.e., parvoviruses), dsRNA (i.e., reoviruses), ss-mRNA (i.e., corona and flaviviruses), ss-mRNA template (i.e., influenza viruses, paramyxoviruses, and rhabdoviruses), and ssRNA-DNA template (i.e., retroviruses)—are considered both a product of this ongoing war and, more importantly, an evolutionary homage to the stepwise transition from RNA to DNA–RNA. These great early battles would result not only in plasmids but also the "prokaryotic condition" and the independent origin of the "Three Domains" of life—Archaea, Bacteria, and an early origin for our very own Eukarya—from three separate viruses like something akin to the "Invasion of the Body Snatchers." Do we detect coughing in the background? Well, why don't we get a pot of soup going and see if we can reduce some of these symptoms of disbelief.

Many organisms ranging from cowbirds to tapeworms to teenagers the world over have made parasitism one of the

⁵ Though not efficient enough to prevent inevitable errors from providing the substrate for life's ongoing diversification and yet with more than enough catastrophic failures to dissuade any intelligent designer from ever laying claim to responsibility after the fact!

most efficient and successful life strategies. Freeloading, however, also takes place at the molecular level, and some scientists speculate that even intragenomic parasites, like transposons and introns, can trace their origins all the way back to the earliest competing ribozymes. And why not? It seems logical that the exploitation of those currently in the lead has been a viable strategy from the start, as resources that would normally be tied up in multiple life processes can be invested heavily in specific areas. But that takes us back to our earlier quandary: Did life have to be cellular, in the traditional sense, from the beginning? It's difficult to say because at some point life would have been neither cell nor virus, lacking both the ability to undergo controlled division and manufacture proteins. A case can be made with some confidence, however, that early assemblages of ribozymes were capable of constructing diverse polypeptides, and amidst the ensuing panmixis, some proved better than others at exploiting precursors and degrading or scavenging others for spare parts (much like the online game "Nanowars," as some students gleefully pointed out).

Proponents of the Primordial Flu generally consider the late cellularization of life far-fetched and contend that the earliest ribozymes must have been vesicle-bound within a hydrothermal vent system. Now it would be easy to imagine a scenario—though difficult to unequivocally prove—where much more efficient genomic colonists started to use larger "resource" vesicles as replication factories that would eventually gain greater autonomy. One would be tempted to call the latter "proto-cells" and the former "proto-viruses," but pre-destiny should never be invoked, for lack of a better word.⁶ Sooner or later, the recombinant benefits of horizontal RNA transfer would have been offset by the costs to the functional integrity of the increasingly independent factory, escalating selection for nucleases and membrane surface traffic to combat exposure to damaged and/or intrusive RNA. Likewise, genome modification and protective shielding would provide countermeasures for potential thieves. Such a burgeoning arms race could ultimately have transformed our early free trade communities into mercantile cellular and viral pirate lineages. In the following scrum, one branch would ultimately give rise to the ancestor of our modern ribosome, the early winner of the RNA World, while other less-efficient protein manufacturers may have survived as the parasites of these now numerous hosts. And what better way to begin a career in home invasion than with a minor disguise.

In modern cells, DNA is manufactured from RNA precursors: Ribonucleotide triphosphates are converted to deoxyribonucleotide triphosphates via the enzyme *ribonucleo-*

tide reductase (RNRs), and uracil monophosphates are converted to thymine monophosphates by *thymidylate synthases* (TdSs). RNA can also be transformed into DNA through undoubtedly familiar *reverse transcriptases* (RTs). Viral origin supporters are banking on the likelihood that scientists have previously underestimated the ability of RNA assisted by transcription factors to competently manufacture large enzymes like *RNA polymerase* (RNA Pol), as well as the aforementioned enzymes. If they're right, following the emergence of the ribosome, viral enzymes could have modified RNA into DNA. The resulting nucleic acid would have been initially resistant to all the cellular mechanisms used to destroy the RNA of hostile viruses. Therefore, the original function of DNA, or its primary adaptation, may have been for defensive purposes as opposed to replicative. While this still doesn't explain how we ended up with textbook-style DNA replication, some plausible steps have been proposed.

An initial ssRNA virus could have initially replicated using an RNA Pol and a dsRNA intermediate. Additional RNRs followed by TdSs could have modified ssRNA into first uracil-based DNA (U-DNA) and then the more familiar thymine-based or T-DNA via a DNA/RNA intermediate. Since the vast majority of viruses are not lytic or lysogenic (sometimes remaining within their capsids), RNA cells would have likely contained viral DNA left in a carrier state. If the cell or virus happened to possess a RT, the DNA virus would have been able to progressively integrate host RNA genes into its own DNA genome to help it survive intracellularly or facilitate infection. Further diversification of the DNA lineage could have led to viruses replicating without an RNA intermediate following the emergence of a *DNA polymerase* (DNA Pol), opening the door for dsDNA, which would have been incredibly resistant to host defenses. Requiring only DNA Pol and strand-displacement activity (i.e., binding proteins), replication of dsDNA probably started off asymmetrical (i.e., one strand after the other), but the later addition of a *helicase* would have made this process more efficient. If replication of the two strands became coupled, larger genomes could then be copied but this would have required a DNA primase to initiate replication in the lagging strand and *DNA ligases* to seal the fragments. Finally, rapid replication of very large genomes would have been possible through tight coordination of the leading and lagging strands (i.e., *DNA topoisomerases*). Phew...

But is there any evidence for this? Surprisingly, yes! Believe it or not, there are actually bacterial viruses with U-DNA genomes! And both cytosine deamination to uracil and U-DNA degradation would have provided the ideal substrate for selection of an enzyme-like thymidylate synthase, thereby facilitating the switch to a now repairable T-DNA. More importantly, if we consider the present as a reliable proxy for its past, then the current existence of all

⁶ That is, "mammal-like reptiles" gets a thumbs down when describing our early Synapsid ancestors

these proposed stages in the virosphere is a considerable argument in favor of the viral origin for DNA. But then how did cells end up with only the final DNA all-inclusive package? It is quite possible that viral genomes too large to fit into capsids eventually became plasmids. Since replication of the new DNA plasmid was much more efficient than the host RNA genome, it would eventually replace the host after capturing all the information needed for cellular life. Subsequent diversification of cells equipped with this modern DNA chromosome would eventually lead to the elimination of all remaining RNA cells, finally ushering in the Cellular DNA World.

However, all is not clear in this brave new world; there are a few puzzling patterns we still need to take into account. For instance, many DNA-associated enzymes (i.e., thymidylate synthases, DNA polymerases, primases, helicases, topoisomerases, etc.) appear to be homoplasious in origin, demonstrating analogous functions but lacking convincing evidence of both structural and sequence homology. While it is highly probable that only a subset of all DNA proteins made during early DNA virus evolution has been transferred to cells,⁷ recent phylogenetic reconstructions of homologous proteins likely to be important in the RNA to DNA–RNA transition (i.e., DNA-dependent RNA Polymerase) support the idea that several large dsDNA viruses are more related to Archaea and Eukarya than to Bacteria. This isn't too surprising when one considers that there appear to be two cellular versions of symmetrical DNA replication, bacterial and archaeal/eukaryotic, but the RNA Pol of the particularly massive *mimivirus* (1.2 million base pairs with over 900 protein coding genes!) ended up as the sister group to Eukarya rather than to both. After taking into account the considerable cell membrane differences between Archaea and Eukarya, as well as other additional anomalies, some scientists have suggested that three independent viral infections are in fact responsible for the three domains of life!

Infiltration by simple large viruses, like *poxviruses*, would result in the “prokaryotic condition” displayed by the lineages Archaea and Bacteria, while a complex large virus similar to *mimivirus* infecting an RNA cell with “bacterial-like lipids” and “archaeal-like ribosomes” would lead to the “early-origin” of Eukarya. While it is not very clear in the literature, it would make even greater sense if the viruses infecting the would-be archaeal and Eukarya branches were more closely related than the bacterial invader. It may also be worth noting that viral-induced recruitment of intracellular membranes by a complex virus may have eventually led to the formation of the eukaryote nucleus in a process similar to the one used by large

dsDNA viruses to form viral envelopes. Then again, there's this whole business of some bacteria possessing nuclei, but we're not going to get into that...

So is the Primordial Flu a reality or simply a loss of phylogenetic signal? It's really hard to tell. If the basic narrative is indeed accurate, however, we now know that textbook RNA translation and the genetic code evolved *before* transcription and way earlier than DNA and DNA replication. In other words, the flow of biological information evolved in the exact *opposite* direction than proposed by the “Central Dogma of Molecular Biology” and “reverse transcription” is actually the original condition! Only after integration within the host RNA cell did today's division of labor emerge where DNA plays the role of hereditary molecule, proteins perform the bulk of catalysis, and RNA provides “scaffolding” while it quietly micromanages in the background, pretty remarkable when you think about it.

More importantly, the evolution of the DNA–RNA cell provides us with a great molecular example of some of the major mechanisms that influence the evolution of complexity. Highly organized and integrated biological systems do not evolve solely through gradual selection of minor heritable differences; there are just too many coincidental confluences that must occur to render this a legitimate possibility, even if considering things from the angle of an intense co-evolutionary arms race. To assume that our DNA–RNA cell evolved purely through the process of natural selection is to dream up a “just-so” story no different than any other non-scientific proposal. Clearly, viruses are more than just degenerate bystanders to the molecular diversification of life, but they could not have been the ultimate designers either. Luckily, we can avoid such orthogenetic and ultra-Darwinist positions by turning to indirect evolution and the process of *co-option*.

It is much easier to wrestle with biological complexity from the perspective that all organisms are a mixture of older and newer traits where only a handful might be (or even have been) beneficial, despite our impulses to provide an adaptive explanation for everything. Lots of features are simply holdovers that aren't severely debilitating or currently useful and that's okay! It's like carrying the letters “X” and “Y” in the scrabble game of life for countless turns; you're not likely to be winning outright but perhaps one day you get lucky and cash in with “XYLEM”! (That is, until the next turn). The new functionality of these letters *is* adaptive, but they were not originally. In much the same way, traits that start off with one function and then switch to a new one without being modified are not *adaptations* (“toward” (*ad*) “fit” (*aptus*)) but what we call *exaptations* (“fit” (*aptus*) by reason of (*ex*) “existing form”), and co-option is the general process underlying this change in function (see McLennan 2008). This doesn't mean that natural selection and co-option exist as two mutually

⁷ DNA Pol E, for example, is exclusively encoded by archaeal and bacterial plasmids.

exclusive processes. On the contrary! Particularly complex systems have been influenced by multiple rounds of both mechanisms working in tandem, as well as the “evolutionary baggage” that just wasn’t or can’t be put down. As a result, each integrated and refined system ends up as a mosaic of primary adaptations, exaptations, and secondary adaptations mixed in with a bunch of other stuff with yet to be determined futures. The point is: Nature rarely reinvents the wheel; it just finds a different use for it.

Turning this lens now to the Primordial Flu Hypothesis, strong evidence suggests that DNA may have originally been an adaptation for viral genome defense and an exaptation for later replicative efficiency and fidelity in cells. By the same token, since any reduction in surface area must have been better than being naked, the humble beginnings of cell walls and membrane traffic likely started off as a first line of defense against viral entry (though became equally exploitable after the fact). Life’s ability to tolerate nearly every environment imaginable by virtue of the later diversification of these originally extracellular deterrents (i.e., glycoproteins, lipoproteins, lipopolysaccharides, peptidoglycans, etc.) can thus be considered subsequent secondary adaptations.

Taking everything into consideration, it may simply prove more realistic to treat the early relationship between cells and viruses as one would a real military arms race. Simply put, war is messy; both sides end up with blood on their hands, and it’s not always their own. Cells and viruses have likely exchanged all kinds of biological material rather haphazardly since the beginning, co-opting bits and pieces of “enemy” weaponry left on the battlefield and building up composite stockpiles after the fact. This could be an alternative explanation for the homoplasmy observed, without invoking the romantic gift of three complete viral DNA packages to our waiting RNA cells. Either way, in a pattern to be repeated often in the history of life, cells already armed with the adaptations necessary to take on life’s new and diverse challenges escaped the increasingly hostile, ironclad prisons of the hydrothermal vents and found a world free for the taking, at least temporarily.

Chapter IV—You Going to Pay Me to Eat That?

When addressing the origin of life question, perhaps it is an inescapable fact that a physicist will ultimately look to the skies, a chemist will look to a new recipe, and a virologist will, well, look to raise the stature of their almost-living subjects. As a result, the vast majority of biology “crossover” attempts inevitably steer the question down more familiar paths. The biologist, meanwhile, is forced to wait until the others have presented their cases, and only then can he or she look to piece together a satisfying

enough, albeit watered down, explanation for the one cell to explain all the rest. If this cell was ultimately the product of great battles with viral factions inside the iron–sulfur cells of a hydrothermal vent, then that’s pretty neat. It could have left already armed with the necessary adaptations, or more technically, exaptations (i.e., dsDNA, cell walls, etc.), to contend with a brave new world of intense UV radiation, drastic variations in temperature, salinity, and pH, and surrounded by a potential windfall of unused resources *if*, and that’s a big if, life were capable of exploiting them. But first it needed the energy to escape.

It’s a pretty well known fact that you need oxygen to live, free dioxygen (O₂) to be exact. We’re told this tidbit of information as children and we go on to believe that this is true of all organisms, particularly the animals we keep in close proximity (with the possible exception of our pet fish that somehow “breathes” under water). Eventually our little misconception is shattered when we are told that plants produce oxygen and “breathe” the carbon dioxide we exhale so it’s a good thing we keep them around: Oxygen good, carbon dioxide bad, carbon monoxide worse. This pretty much summarizes the extent of your average person’s knowledge of metabolism, even if they wouldn’t classify this piece of information as being an integral component of metabolism (which is a word strewn about typically in reference to how fast or slow someone “breaks” down something they’ve recently ingested). And because most people don’t know *why* they fundamentally need oxygen or why a plant produces it in the first place, we often feel compelled to try to teach “cell respiration” and “photosynthesis” in the hope that someone will learn to appreciate the wonders of glucose, glycolysis, ATP, the citric acid (or Krebs) cycle, the electron transport chain (ETC), electron “carriers” like NAD⁺, FAD, and NADP, PSII and PSI (unfortunately not the gaming consoles), chlorophylls, cytochromes, quinones, ATP synthase, the Calvin cycle, rubisco, $\text{CH}_2\text{O} + \text{O}_2 \rightarrow \text{CO}_2 + \text{H}_2\text{O} + \text{energy}$ ⁸= cell respiration (and the reverse=photosynthesis), etc. etc... memorize, role-play, cajole, motivate, rinse, and repeat, while clumsily perpetuating a new set of misconceptions.

Most students end up believing the following: We run entirely on glucose, every glucose molecule goes on a magical one-way journey from glycolysis to pyruvate oxidation to the Krebs cycle and finally to the ETC; metabolism refers to only the degradation of complex organic carbon molecules to simpler inorganic forms (catabolism); plants use the Calvin cycle to make sugar, and this is the only way to “fix” or reduce carbon; cell respiration is restricted to animals; and plants are the only “self-feeders.” Who’s to blame for these classic misconceptions?

⁸ Unbalanced equation and showing only net consumption of water

Perhaps no one since metabolism is such an immense topic, and where in fact do you begin? And I honestly don't know. ☺

Well perhaps a safe place to start is with a definition: Metabolism is the sum of all the catabolic AND anabolic chemical reactions in a cell (~1,000) catalyzed by proteins called enzymes. These reactions are distributed among a universal set of about 50 pathways despite whatever additional metabolic specializations may be found amidst the incredible diversity of life on our planet, which also implies the predominantly vertical inheritance of these paths since the last universal common ancestor (or LUCA). The vast majority of these “core” pathways involves carbohydrates (i.e., glycolysis, gluconeogenesis, pyruvate metabolism, the pentose–phosphate cycle, etc.), amino acids, fatty acids (but not membrane lipids), and nucleotides, as well as cofactors and vitamins, the citric acid cycle, and enzymes and coenzymes involved in energy transfer. If we then attempted to create a “metabolic map” of this universal set, we would be left with a proverbial labyrinth of twists and turns that eventually converge at the citric acid cycle. Now when faced with the reality of universal metabolism, one's natural inclination is to play intellectual “hooky” and do one of two things: You will either see this as clear evidence for common descent and accept that these paths evolved “somehow” in the distant past *or* you will see it as an “irreducibly complex” phenomenon that has been mysteriously stamped onto all life forms big and small. Luckily, there are a handful of more curious minds that have taken the challenge of untangling this potential Gordian knot.

Before getting tied down with specifics, let's imagine a simpler scenario where an essential nutrient “A” is being abiotically generated, but consumption has started to exceed production. An organism then capable of manufacturing nutrient A from precursor “B” using *enzyme I* would be at an immediate selective advantage, and soon only organisms with this enzyme could survive. But if precursor B likewise starts to run out, only those with additional *enzyme II* could still make nutrient A from precursors B and now “C.” If we then run this scenario a few more times, the result would be an *anabolic* or biosynthetic pathway (i.e., $5 \rightarrow 4 \rightarrow 3 \rightarrow 2 \rightarrow 1 \rightarrow \text{“nutrient”}$) with the earliest enzyme “downstream” of all the rest. In other words, the metabolic path evolved backwards.

Of course, the opposite would be true for the evolution of a *catabolic* pathway where enzymes would be selected for continued degradation of nutrient “A” to subsequent waste products “B,” “C,” “D,” etc., resulting in the earliest enzyme being “upstream” (i.e., $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow \text{waste}$). Since it would be the most efficient arrangement, such pathway “optimization” makes for a sound starting hypothesis, but nature “rarely reinvents the wheel” and we can expect enzymes to be opportunistically co-opted (or “recruited”) for new metabolic tasks so long as they are good enough for the job.

While this certainly complicates our starting picture, shared enzymes, like any shared, similar character, may provide us with the information we need to unravel the past. The question is, can we reconstruct a metabolic past when there are no suspects to interrogate since all are equally guilty and innocent?

Metabolism, like any aspect of an organism, is subject to “descent with modification.” But there's very little anatomy to compare when all the organisms possess the same set of characters. As a result, scientists studying the evolution of metabolism have worked under the assumption that glucose was the first anaerobic (O_2 free) source of energy and glycolysis the earliest metabolic pathway, given its products and its current position with respect to the citric acid cycle. They based this assumption on the further assumption that amino acid catabolism was unlikely in a world where complex organic molecules were relatively scarce. The citric acid cycle was then believed to have been assembled after glycolysis, functioning originally as an early biosynthetic pathway for amino acids. Unfortunately, glucose has not been well documented in abiogenic experiments, and the basis of these a priori assumptions seems increasingly unwarranted.

Shared enzymes and reactions, like any shared similarity, can be interpreted as the product of common ancestry. The challenge is distinguishing similarities that have arisen through common ancestry (i.e., homology) versus convergence or parallel evolution (i.e., homoplasy). Luckily, phylogenetic systematics and the principle of parsimony provide us with a method of determining homoplasy after the analysis has been completed, thus reducing the number of a priori assumptions we need to make. Since all cellular life possesses the core set of pathways we're interested in reconstructing, a comparison of the metabolic paths of extant organisms would only reflect differences in metabolic regulation rather than differences in structure. To get around this limitation, Cunchillos and Lecointre (2007) had the ingenious idea of treating metabolic pathways as you would organisms (or taxa) while using enzymes and enzymatic functions as characters to build a metabolic phylogeny with an outgroup that obviously possessed nothing. Moreover, they excluded enzymes that operate on triglycerides, phospholipids, coenzymes, and nucleotides, as well as polymers like proteins, RNA, DNA, starch, glycogen, and so forth, focusing instead on the three main kinds of elementary biochemical compounds—amino acids, fatty acids, and monosaccharides—which would later assemble into these more complex molecules after the fact. And the results were quite interesting.

Based on the relationship between universal metabolic pathways, the relative emergence of the main parts of metabolism went as follows: (1) amino acid catabolism, (2) amino acid anabolism? closure of the urea cycle, (3) glycolysis and gluconeogenesis, (4) closure of the pentose–phosphate cycle, (5) closure of the citric acid cycle

and fatty acid metabolism, and finally, (6) closure of the Calvin cycle. Just to clarify, the order of biochemical emergence listed above refers to the *completion* of each major metabolic set. Although glycolysis was completed before the closure of the citric acid cycle, for example, some pathways involved in the current cycle seem to have emerged at relatively the same time. And while the catabolism of a few amino acids evolved before anabolism, it is the opposite in some cases and ambiguous for others. Should it be surprising that amino acids were first degraded for energy before being used to build materials? Not really, when one considers the relative ease of their abiogenesis based on classic experiments, meteor analysis, and newer work on hydrothermal vent chemistry.

Furthermore, while the evolution of amino acid catabolism fits our initial forward evolution hypothesis (i.e., 1→2→3... waste), the reverse anabolic picture is not as clear. Perhaps this is related to the availability of particular amino acids, which in turn has ramifications for the evolution of the genetic code. After all, is it a mere coincidence that the five most prevalent abiotically generated amino acids—glycine, alanine, aspartic acid, glutamic acid, and valine—also tend to be the most thermodynamically favorable *and* appear in the first row of the genetic code? Nevertheless, it may have been rather erroneous to include the Calvin cycle, considering our growing awareness of the diversity of carbon-reducing or “fixing” mechanisms still employed on Earth today. In fact, no evolutionary discussion of metabolism would be complete without a visit back to the hydrothermal vents to explore perhaps the earliest redox reactions and the later metabolic *diversification of life*.

Was LUCA a heterotroph (“consumer”) or an autotroph (“producer”)? Traditionally, glycolysis-like fermentation—a form of *anaerobic* chemoheterotrophy coupled with *substrate-level phosphorylation*—was considered to be the earliest source of organic carbon and energy. However, with the discovery of hydrothermal chimneys, the case for chemoautotrophic origins has been growing steadily. Perhaps the most compelling line of evidence can be traced to several groups of vent-living bacteria and Archaea that utilize a reductive acetyl-CoA pathway to fix carbon from the simplest of reactants: $\text{H}_2 + \text{CO}_2 \rightarrow \text{CH}_3\text{CO}\sim\text{SCoA} + \text{H}_2\text{O}$. “Acetogens,” as we often call them, are essentially being “paid to eat a free meal”; not only do they receive organic carbon, but the reactions are also exergonic and can be coupled to the *chemiosmotic* generation of ATP (*oxidative phosphorylation*). You may be thinking, “Isn’t acetyl-CoA just a decarboxylated product in the short transition from glycolysis to the citric acid cycle?” Nope! And it’s not the only “heterotrophic” pathway that seemingly operates in “reverse” of cell respiration. There are in fact five known carbon-fixing mechanisms in the natural world, including a reductive citric acid cycle and probably the more familiar

Calvin cycle. What sets the reductive acetyl-CoA path apart from the others, however, is that it does not require complicated intermediate biomolecules, making it a strong candidate for one of the oldest metabolic pathways.

In addition, the two enzymes involved contain active sites that bear a striking resemblance to minerals formed in precipitating hydrothermal mounds (i.e., Mackinawite ((Fe,Ni)_xS) and Greigite (NiS₂[Fe₄S₄]S₂Fe)), which also appear quite capable of catalyzing this reaction all on their own (see “Chapter II—Soup that Eats Like a Meal!”). Did life simply *co-opt* these catalysts straight from the source? It sure seems that way, and it’s not the only example: electron-carrying ferredoxins, electron-delivering hydrogenases, and nitrogen-reducing nitrogenases likewise possess metal-sulfide clusters that appear to be intimately connected to similar minerals formed within hydrothermal chimneys. What a strange coincidence, don’t you think?

As you may already know, acetyl-CoA is an incredibly versatile molecule with the potential to serve as an intermediate in a number of pathways, including the synthesis of fatty acids that can then be used to form membranes. Acetyl-CoA possesses a high-energy thioester bond (~SH), and hydrolysis generates a considerable amount of free energy that can be coupled to endergonic processes (much like the breakdown of ATP), in addition to producing acetate (CH₃COOH) as “waste.” Why is this important? Well, because once upon a time, there were no ATPases, and the hydrolysis of thioesters may have served as one of life’s early sources of energy. In fact, if biological membranes even existed at the time, they would have been incredibly leaky to protons (H⁺)! So how or under what conditions did something as impressive as an ATPase evolve?

Unbeknownst to most, ATPases are not all the same and even these “rotary dynamos” have left traces of a transitional legacy. First off, there are two main types of reversible ATPases. They’re reversible? Indeed, ATPases are capable of using an ion current to either hydrolyze or synthesize ATP. The “F”-type is thought to be ancestral for bacteria but can also be found in the mitochondria and chloroplasts of eukaryotes, as well as a few archaeans. The “V”-type, on the other hand, is considered to be characteristic of archaeans but can also be found in eukaryotic cytoplasmic membranes and some bacteria. If the vertical inheritance pattern seems slightly askew, it’s because it has been interrupted on several occasions by lateral or horizontal gene transfer (HGT), an apparent hallmark of single-celled existence.

Like many complex structures, ATPases may owe their existence to the co-option of pre-existing parts: The mushroom-like ATPase “head” appears homologous to the catalytic subunits of helicases, which has led to several proposals that it originated as a helicase transporting RNA through protein channels. This makes sense when considering

the would-be importance of transporting macromolecules like nucleic acids and proteins at a stage when life couldn't divide, and the "seeding" of new resource cells was the only viable replication strategy. At some point, this simple RNA helicase/transport system could have eventually acquired peripheral support proteins to facilitate docking and then eventually evolved into protein transporters. While this level of organization appears homologous to the ATP-driven export of flagellin in the case of bacterial flagellar motors, the non-homologous central stalk of F- and V-type ATPases may be a result of different proteins becoming "trapped" in the process. Only then would the ATPases take on a configuration capable of transporting ions. Although these ideas are still considered speculative, they are not without merit, as some ATPases still participate in moving DNA during bacterial conjugation and cell division, secrete and uptake DNA, and retract and move pili! An ATPase found in double-stranded RNA bacteriophages is actually involved in unwinding, packaging and extruding its RNA!

Most discussions of ATPase evolution and the origin of life inevitably focus on either how acidic or basic the environment was with respect to an organism trying to maintain a nearly neutral cytosol. Thus, in the acidic ocean scenario, the constant pumping out of protons gradually became coupled with an enzyme taking advantage of the return "proton motive force"; or in the alkaline nursery, early life merely took advantage of an interface with a naturally more acidic ocean. Regardless, cellular-bound life would not have started out as the picture of cation chasteness described in textbook style chemiosmosis. On the contrary, only larger ions could have been used to generate an early electrochemical potential, and the discovery of several prokaryote ATPases that actually bind sodium (Na^+) and use a "sodium motive force" to generate ATP lends much credence to this idea. In an ocean increasing in salinity, sodium translocation may have served as the ancestral condition until membranes became tight enough to utilize the much more energy efficient protons available. Luckily, hydrothermal vents come in both acidic and alkaline flavors, and while the case for alkaline origins grows stronger every day, it isn't too difficult to reconcile the growing body of evidence to either scenario, should an eventual consensus emerge.

Discussing origins is ironically never simple. It is very easy to lose sight of the big picture by becoming overwhelmed by even the most pertinent of minutia. It's taken us a while, but it seems we've more or less arrived at a reasonable picture of the one cell to explain all the rest: Battle-tested and opportunistic, it likely possessed double-stranded DNA, plasmids, ribosomes, a cell wall, metalloenzymes (i.e., Fe-S clusters), a cell membrane of uncertain permeability, versatile pili (including conjugal), a reversible sodium-ATPase, and perhaps even a flagellum. Our little LUCA sure has grown

up a lot over a span of a few hundred million years, and, dare we say, is finally poised to take its first flagellar beats away from an increasingly overcrowded and virulent home.

Chapter V—Why Excuse Me! Was it Something I Ate?

But the will to leave the nest is never enough; you also need a way. When we're talking about the diversification of life, finding a "way" is really no different than being financially independent, except now you're dealing with an economy of electrons instead of dollars, and the bottom line is that only a few will leave with an inheritance good enough to get a head start. The rest will have to settle for the good fortune of environmental bursaries or simply return home, while the vast majority inevitably winds up on skid row, too ill-equipped to deal with the new challenges. Early life undoubtedly experienced similar growing pains, and it wouldn't be a surprise if the first fledglings stayed within fridge-raiding distance, never above accepting familiar care packages bubbling up from the seafloor.

Meanwhile, those equipped with metalloenzymes had the advantage of motility and an ocean full of familiar electron donors (synonymous with the "ox" part of redox, reductants, or reducing agents) to consume. This short-lived cornucopia would undoubtedly fuel the basis for a prokaryotic radiation of life, particularly when proliferation and competition for ancestral donors accelerated in regions further away from home. The diversification of early chemoautotrophy would in turn provide substrates for reverse heterotrophy in an increasingly complicated and mutually dependent economy of electron donors and acceptors (see Table 1). For example, the emergence of chemoautotrophic methanogens provided a reduced source of organic "waste" to be oxidized by specific heterotrophs. The release of methane, a far more potent greenhouse gas than carbon dioxide, would have also contributed to already high temperatures on early earth.

We often refer to many of these early metabolic strategies as *lithotrophic* or "rock-eating" since they frequently involve the breakdown of rocks and minerals to gain access to reducing and oxidizing agents. Although primary productivity would remain low by today's standards, any new metabolic solution was poised to trigger dramatic effects on both a biological community still in its infancy and a planet contending with a restless force that began to literally take on a life of its own.

One such solution was the emergence of light-absorbing pigments, which allowed life to trade the entropy of the Earth for the incredible energy cast out by the slow thermal death of the Sun. It is this captured electromagnetic radiation that sustains the diversity of life on Earth today. No one can be for certain how or when exactly, but around

Table 1 Metabolic diversity and redox reactions

Conditions	e ⁻ donor	e ⁻ acceptor	Examples ^a
Anaerobic	H ₂	CO	Acetogenesis – CO + H ₂ → CH ₃ COOH + H ₂ O
	H ₂	CO ₂	Acetogenesis – CO ₂ + H ₂ → CH ₃ COOH + H ₂ O
	H ₂	CO ₂	Methanogenesis – CO ₂ + H ₂ → CH ₄ + H ₂ O
	H ₂	CO ₂	Anoxygenic photosynthesis – CO ₂ + H ₂ + <i>hν</i> → CH ₂ O + H ₂ O
	H ₂	SO ₄ ²⁻	Sulfate reduction – SO ₄ ²⁻ + H ⁺ + H ₂ → HS ⁻ + H ₂ O
	H ₂	S	Sulphur reduction – S + H ₂ → H ₂ S
	H ₂	NO ₃ ⁻	Denitrification – H ₂ + NO ₃ ⁻ → N ₂ + H ₂ O
	H ₂ S	CO ₂	Anoxygenic photosynthesis – CO ₂ + H ₂ S + <i>hν</i> → CH ₂ O + SO ₄ ²⁻
	H ₂ S	NO ₃ ⁻	Nitrate reduction – NO ₃ ⁻ + H ₂ S → N ₂ + SO ₄ ²⁻
	Fe ²⁺	CO ₂	Anoxygenic photosynthesis – CO ₂ + FeO + <i>hν</i> → CH ₂ O + Fe ₂ O ₃
	As ³⁺	CO ₂	Anoxygenic photosynthesis – As ³⁺ + CO ₂ + <i>hν</i> → CH ₂ O + As ⁵⁺ + H ₂ O
	NH ₄ ⁺ (and NO ₂ ⁻)	NO ₂ (and CO ₂)	Annamox – NH ₄ ⁺ + NO ₂ ⁻ + HCO ₃ ⁻ + H ⁺ → N ₂ + NO ₃ ⁻ + H ₂ O + CH ₂ O 1) NH ₄ ⁺ + NO ₂ ⁻ → N ₂ 2) NO ₂ ⁻ + CO ₂ → CH ₂ O + H ₂ O
	CH ₄	SO ₄ ²⁻	Methane oxidation – CH ₄ + SO ₄ ²⁻ → CO ₂ + H ₂ S + H ₂ O
	CH ₂ O	NO ₃ ⁻	Denitrification – CH ₂ O + NO ₃ ⁻ + H ⁺ → N ₂ + CO ₂ + H ₂ O
	CH ₃ CO ₂ ⁻	Fe ³⁺	Iron reduction – Fe ³⁺ + CH ₃ CO ₂ ⁻ + H ₂ O → HCO ₃ ⁻ + Fe ²⁺ + H ⁺
	C ₆ H ₁₂ O ₆	NAD ⁺	Fermentation – C ₆ H ₁₂ O ₆ + NAD ⁺ → C ₂ H ₆ O + CO ₂ + NADH + H ₂
Aerobic	H ₂ O	CO ₂	Oxygenic photosynthesis – CO ₂ + H ₂ O + <i>hν</i> → CH ₂ O + O ₂ + H ₂ O
	(CH ₂ O)	O ₂	Respiration – CH ₂ O + O ₂ → CO ₂ + H ₂ O
	CH ₄	O ₂	Methane oxidation – CH ₄ + O ₂ → HCO ₃ ⁻ + H ⁺ + H ₂ O
	Mn ²⁺	O ₂	Manganese oxidation – Mn ²⁺ + O ₂ + H ₂ O → MnO ₂ + H ⁺
	H ₂	O ₂	Hydrogen oxidation – H ₂ + O ₂ → H ₂ O
	HS ⁻	O ₂	Sulfide oxidation – HS ⁻ + O ₂ → SO ₄ ²⁻ + H ⁺
	NH ₄ ⁺	O ₂	Nitrification – NH ₄ ⁺ + O ₂ → NO ₂ ⁻ + H ⁺ + H ₂ O
	NO ₂ ⁻		NO ₂ ⁻ + O ₂ → NO ₃ ⁻
	Fe ²⁺	O ₂	Iron oxidation – Fe ²⁺ + H ⁺ + O ₂ → Fe ³⁺ + H ₂ O

CH₂O refers to an idealized reduced organic carbon (including pyruvate, acetate, formate, fumarate, methanol, etc.)

hν electromagnetic radiation

^a Equations are unbalanced, and some of the chemoautotrophic pathways do not illustrate carbon reduction, as many of these organisms are mixotrophs, employing both autotrophic and heterotrophic mechanisms based on context

3.5 billion years ago, some organisms—perhaps first protected by various UV-screening proteins—began to utilize solar energy to photooxidize electron donors. Did this occur in a bubbling terrestrial pond after colossal waves flung out castaways in all directions? Or did life just gradually move up the water column? Based on the structural similarities between so-called bacteriochlorophylls and “plant” chlorophylls, we’re quite certain that these pigments share a common ancestry, but the incredible diversity in absorption spectra indicates a rich history of niche differentiation.

For example, while bacteriochlorophylls and chlorophyll “a” typically absorb light in the 700–900- and 400–700-nm range respectively, some cyanobacteria use chlorophyll “d” instead of “a” (720 nm) and a species of purple bacteria (*Blastochloris viridis*) can use wavelengths as long as 1,025 nm! While one may contend that the absorption of these longer wavelengths in the near infrared is an adaptation to occupying murky water and sediments or contending with light filtered through overlying organisms, it is equally possible that they represent an ancestral light-absorbing condition somewhere lower in the water column before life was capable of taking on the high-energy extremes of the near surface.

Although bacteriochlorophylls tend to absorb lower energy wavelengths, the free energy gains were far greater than any previous lithotrophic reliance on the redox disequilibria of geological processes, and this growing independence allowed phototrophs to flourish and expand their ecological ranges. Before long, organisms possessing bacteriochlorophylls held in reaction centers and carbon-fixing enzymes began to dominate the microbial landscape. This increased primary productivity many times over by providing both organic matter and electron acceptors for likewise expanding heterotrophic niches (i.e., SO_4^{2-} , NO_3^- , Fe^{3+} , etc.), the more electronegative of which would allow further oxidation of reduced carbon sources and hence greater energy return.

The most productive of these early communities would have likely been based on electrons donated from hydrogen and iron(II), but some photoautotrophs even made use of arsenic(III)! With the advent of photosynthesis, it perhaps became feasible for life to reduce its own nitrogen (N_2) instead of relying on the contributions of lightning and hydrothermal outgassing. As mentioned earlier, nitrogenase is a metalloenzyme containing iron–sulfur clusters (19 to be exact) that bear a striking resemblance to minerals found within hydrothermal vents. At an incredible cost (~16 ATP!), nitrogenase reduces a single atom of nitrogen to ammonia ($\text{N}_2 + 8\text{H}^+ + 8\text{e}^- \rightarrow \text{H}_2 + 2\text{NH}_3$), which in turn is converted to ammonium where it can then be utilized to manufacture proteins. Without efficient photosynthesis and subsequent *photophosphorylation*, it is very unlikely that life could have absorbed these high-energy costs.

Moreover, based on the distribution of nitrogenase, it appears to have been lost and swapped multiple times across many lineages, and the same can be said for enzymes involved in carbon-fixation (i.e., rubisco of the Calvin cycle). This isn’t too surprising when you weigh the benefits of being able to supply your own reduced carbon and nitrogen against the great manufacturing costs of maintaining the machinery for the job. And if you can avoid it, why make something expensive when someone else will? If we were to take a snapshot of the biological world approximately 3.5 billion years ago, it would be easy to describe it as an idyllic time where diverse prokaryotes exchanged tools and dined pleasantly together as they slowly filled every niche imaginable. But a pleasant evening out at Earth’s surface diner was about to be ruined by a table full of loud, gaudy, and gassy customers.

As you may already know, reaction centers (RCs) are part of a much greater protein complex that also includes light-harvesting units and other pigments like carotenoids, which subsequently form an incredible structure known as a photosystem (PS). There are two types of photosystems based on differences in reaction centers: *type I* RCs typically draw electrons from hydrogen sulfide for linear-electron transport and possess iron–sulfur clusters, while *type II* RCs are quinone-based and drive a cyclic electron transport pathway.

Additionally, we can divide photoautotrophy into two main groups: organisms that possess reaction centers in either a PSII-like (i.e., purple non-sulfur bacteria) or PSI-like complex (i.e., green sulfur bacteria), but never both, perform what is known as anoxygenic photosynthesis. Anoxygenic phototrophs rely on longer wavelengths of light to oxidize electron donors like H_2 , H_2S , and Fe^{2+} , eventually yielding products like water, sulfate, or iron (III). But somewhere between 3.4 and 2.3 bya, the ancestors of today’s cyanobacteria did something incredible: They split H_2O , stole its electrons, and released O_2 as a by-product. How did they accomplish this feat? Well, in possibly one of two ways.

The ancestor of cyanobacteria either possessed one photosystem, and then a second one was transferred via HGT giving rise to a “chimeric” photosynthetic apparatus, or there was a gene duplication event followed by subsequent divergence into two different photosystems. In the latter scenario, the bacterium likely expressed one or the other depending on the situation via a regulatory switch; a mutation disabling this switch under the right environmental setting and time would generate a smooth opportunistic path toward what is known as oxygenic photosynthesis.

Functionally, a cluster of four manganese and one calcium ion (MnCaMn_3) is responsible for extracting four electrons from water one at a time in what is known as the

oxygen evolving complex (OEC) found on the surface of PSII. Only when this cluster is a full four electrons short does it rip them from two water molecules, producing the waste product oxygen (O_2); the calcium ion may stabilize the first oxygen molecule until the second one is added. Believe it or not, the manganese cluster bears a striking similarity to minerals like calcium manganite ($CaMn_4O_8$) and hollandite. Could this be yet another example of mineral co-option? Perhaps. When exposed to UV light, manganese becomes photooxidized and loses an electron, leaving it in a more oxidized state. As a result, the absorption of UV by Mn atoms in the OEC causes them to break away from PSII (“photoinhibition”), preventing the continued oxidation of water. In a pre-water-splitting world, an organism expressing PSII and living in an aquatic, Mn-rich environment would constantly have to contend with electrons flowing into its photooxidized chlorophyll. This logjam would be deadly unless a second photosystem (PSI) was expressed at the same time to bleed off the surplus and create the flow of electrons we observe today in all oxygenic photosynthesizers.

Fine tuning of this arrangement had obvious benefits: Water is literally everywhere on our planet, and any organism capable of exploiting it would have a truly unfair advantage. Numerous fossil stromatolites are a testament to the gaudy success of the cyanobacteria, and through their new industry, a waste product would begin to accumulate and slowly but steadily transform Earth into an aerobic world.

Nature rarely reinvents the wheel. Then again, it also makes up the wheel as it goes along. Needless to say, even though the finished product may look perfect after being run through generations of abrasives and polishes, it is still tethered to the existing parts and impromptu selections made in the past; life doesn’t have a “Magic 8 Ball” or a chop shop to go through a catalog of all the best pieces life has to offer and then “soup-up” a cell (well, beyond a certain degree). That said, the splitting of water, though a boon to cyanobacteria, set off a cascade of busts that is truly remarkable to comprehend.

Geologically speaking, the rise in dioxygen literally caused the earth to rust. Banded iron formations gave way to “red beds,” and the soluble Fe^{2+} of which many life forms were now dependent became the much more insoluble Fe^{3+} . Since other transitional metals also existed in a reduced state, it took some time—a span of perhaps a few hundred million years—for O_2 to become a significant component of the atmosphere. At some point, O_2 levels were high enough to convert much of the ammonium into nitrite and nitrate, which are substrates for both denitrifying and anammox bacteria, but too low to carry out effective oxygen-dependent nitrification. Consequently, life may very well have experienced a nitrogen crisis where the

fixed nitrogen supply in the environment became very unstable.

To make matters worse, nitrogenase is incredibly oxygen sensitive! At present O_2 concentrations, cyanobacteria lose approximately 30 percent of their nitrogenase activity as the Fe–S clusters become oxidized and the solution is to simply make more of the enzyme, which further strains the nitrogen economy. The same can be said about the enzyme rubisco, where upwards of 40 percent of its activity is oxygenase in a typical C3 plant and the solution is once again to make more! Unsurprisingly, it may be the world’s most abundant protein. Even the proteins involved in the water-splitting complex of PSII (i.e., D1) struggle with the task at hand and become irreversibly damaged after exposure to oxygen and must be replaced about every 10,000 electron transfers. These examples are often referred to as “frozen metabolic accidents,” and ironically this is not far from the truth.

It’s difficult to tell whether the single greatest (but typically forgotten) extinction in the history of life occurred before, during, or after the destruction of a methane greenhouse, but the rise in oxygen undoubtedly led to the death and marginalization of countless obligate anaerobes in what has been dubbed “The Oxygen Catastrophe.” Although rubisco is far from perfect, it still handles oxygen much better than any of the other available carbon-reducing enzymes, which are more or less rendered useless in high oxygen environments. Those not poisoned to death immediately may have eventually perished under the harsh, nutrient-poor conditions of a Snowball Earth (the “Makganyene”) approximately 2.2 billion years ago. Scientists still debate whether or not the planet completely froze over or remained a “slushball,” but the geological evidence suggests that it was quite frigid all the way down to the equator. Many believe the smoking gun rests in the photosystems of flourishing cyanobacteria, but if this is indeed the case, a young Sun with lower luminosity is also to blame. Thankfully for our ancestors, the primordial restlessness of our planet has always churned at its core, and ongoing volcanic emissions would eventually rescue the planet from a perpetual deep freeze. Following the great thaw, surviving cyanobacterial strains continued where they left off; oxygen levels steadily rose again and gave birth to the planet’s contemporary *biogeochemical cycles*. But why on earth would an organism chum out a by-product that would nearly end its own existence?

Life is indifferent to the future and is tethered by the past. It does not reach some sort of illusory “balance” with the environment because of some intrinsic harmonious drive to be at one with Earth, quite the opposite in fact. It is because life routinely overshoots its carrying capacity that we observe such diversity today. Without this indifference, there would be no struggle for existence, no selection, and no modification of the next generation. The fact that

cyanobacteria flourished without foresight and in doing so killed off innumerable competitors, ruined the nitrogen economy, nearly committed suicide, and perhaps ushered in an ice age of global proportions is proof that there is no intelligent plan but simply survival of the good enough. The release of free oxygen was essentially *primordial flatulence* that skunked not only anaerobes but also exposed life's imperfection.

Evolution can only tinker, it cannot engineer. It is not a mad scientist in a lab whimsically putting together the best parts from a limitless inventory. The fact that enzymes vital to the sustainability of life as we know it remain oxygen sensitive and have not been “fixed” in billions of years is not proof of natural selection's ineptitude (since natural selection *cannot* create in the first place!) but rather how unnecessary it is to be the “fittest.” The energy of the Sun is so great that these costs are affordable to the already rich, and you only need to be good enough to survive. Regardless, the tale of the cyanobacteria, perhaps the world's first mega-polluter, really puts our own climate issues into perspective.

To the majority of life forms on early earth, the splitting of water was like opening Pandora's box, but life was far from doomed. The same indifference that exposes life's poor planning and imperfection also generates life's resiliency. As evidenced by our current biodiversity, a few organisms were able to adjust and even profit from this new pollutant. Proof-positive, purple photosynthetic (non-sulfur) bacteria, while photosynthetic under reducing conditions, are remarkably able to operate their electron transport chains in “reverse” under aerobic conditions to oxidize organic carbon! In other words, the proton gradient did not have to be reinvented. By using O₂ as a terminal electron acceptor under aerobic heterotrophic conditions, organisms receive a *huge* increase in free energy from the oxidation of glucose when compared with anaerobic mechanisms like fermentation. For a purple non-sulfur bacterium, this also means it can survive when the lights are turned out during, oh, let's say, a global deep freeze. Likewise, anoxygenic photoheterotrophs (i.e., purple sulfur bacteria) are able to use the energy of the Sun to offset heterotrophic costs even though they are unable to fix enough carbon on their own, which gives them an advantage over strict heterotrophs when dissolved organics are plentiful. “Dual fuel” is a prime example of the metabolic flexibility enjoyed by the microbial world, and this is undoubtedly one of the main reasons why the reign of bacteria continues even after billions of years and countless organisms have come and gone.

In the end, we breathe oxygen for the same reason that other heterotrophs need sulfate, nitrate, or iron (III). If you can handle it, oxygen is simply a powerful electron acceptor. However, the question of why you need oxygen to live is a lot less interesting than “Why do you breathe

oxygen?” for the need would not exist without the opportunity. And as we've just witnessed, the question of opportunity, unlike a question of current functionality, is a question of origins, an inquiry into a historical event, a transition in time; in short, a story to tell, a story that links all life forms in a single drama and one not just restricted to the current winners.

Chapter VI—Guess Who's Coming to Dinner? And Staying...

As we wind down our journey through deep time, perhaps it is fitting that our previous five chapters carry with them five lessons: *Nature is a tinkerer, not an engineer*—natural selection can only operate on existing variation leading to the gradual refinement of parts (i.e., RNA diversity, the origin of DNA, ATPase, etc.), not the sudden emergence of a complex whole. *Nature rarely reinvents the wheel*—existing parts (even inorganic) can be co-opted for new functions with or without subsequent secondary refinement, which can lead to overall greater complexity (i.e., DNA replication machinery, metalloenzymes, oxygenic photosynthesis, etc.). *Nature will take the path of least resistance*—life is lazy and will optimize towards the best-worst option available (i.e., the one that yields slightly more benefits than costs). Diversity has always been at the mercy of the shifting trade-offs between “do-it-yourself” and “ready-made” strategies; it's often home cooking versus take out (i.e., facultative heterotrophs). *Nature makes it up as it goes along*—combinations of refinement and co-option in the past are made without knowledge of the future and limited by the *available* options, which can lead to suboptimal design. In other words, life is historically constrained and survivability is historically contingent (i.e., oxygen sensitivity of vital enzymes). If you got it, flaunt it; if you don't, you may sadly be out of luck. And last but not least, *life goes on*.

Imagine that you're at a huge family dinner spanning generations. Naturally, your guest is that special belle or beau you've been seeing for some time, but you're afraid that his or her chosen lifestyle might a bit too extreme for your conservative family. After a glorious meal, everyone decides to show their appreciation with modest burps of carbon dioxide; great granddad lets out a little bit of hydrogen sulfide and everyone giggles. It's surprisingly going okay. But then all of sudden, your date does the unthinkable and downs a cool glass of water in front of everyone and drops the single most putrid example of metabolic leftovers the world has ever known. Grandma and grandpa keel over instantly; your cousins, aunts, and uncles writhe on the floor, while your siblings hide in the closet; mom and dad are left with a frozen look of horror on their faces and you, well, you're actually not that bothered

by it; sure, it's a little strange but tolerable. Your family, well, what's left of them, is horrified; they instantly pack up and move back to the south where conditions are familiar and you're left struggling to make ends meet.

That's pretty much what happened about 2.2 bya, except the consequences were a tad more severe: the killing of millions of terrestrial and photic zone obligate anaerobes, food shortages, global freezing, etc., all because an organism was completely indifferent to the well-being of others. How rude! It's too early to tell whether our industrious choices, mere extensions of our old family ties to aerobic chemoheterotrophy, will lead to a major ecological catastrophe on par with life's first flirtation with dioxygen, but history tells us that there's a pretty good chance life will rebound as it always has after being pushed to the near brink.

During or shortly after the meltdown, the survivors hosted a far quainter affair that would prove to be no less controversial. As strange as it may sound, we're not sure who started the event—either the would-be hosts failed to make meals out of the now permanent guests or the would-be guests failed to make meals out of the now permanent hosts. Either way, the relationship forged between them would shock the world by overcoming one of the most notorious social barriers.

If reports are true, it appears that a run-of-the-mill archaeon either engulfed or was infiltrated by a relative of present-day α -proteobacteria, perhaps similar to extant *Rickettsia* species or purple non-sulfur bacteria. Once inside, either the host archaeon failed to digest the α -proteobacterion or the α -proteobacterion failed to complete its parasitic lifestyle and became a permanent resident. The descendants of this strange matrimony would give rise to the Eukarya (under the eukaryotes-late hypothesis⁹), and in some of our better-known lineages, the descendants of these α -proteobacteria would become our mitochondria. Yes, believe or not, the organelle that “super-charges” your cells originated from a symbiotic alliance that occurred anywhere between 2 to 1.5 billion years ago! And later, at a more exclusive gala, it appears that this eukaryote conglomerate successfully proposed a similar merger with a cyanobacterium, and its descendants would one day become the plastids (chloroplasts, in particular) of algae, plants, and other photosynthetic “protists” whose lineages independently went “shopping” for additional accessories. To think, *organelle*-level cellular organization emerged from perhaps the most extreme example of co-option: endosymbiosis. But as fantastic as this might sound, is there

any proof of these events? Surprisingly, given the timescale we're looking at, there's quite a bit!

Where to begin? Well, both mitochondria and chloroplasts possess a double membrane with the innermost one made up of a chemical composition that is similar to bacterial plasma membranes, which is to be expected if a larger cell engulfed a smaller bacterium. By the same token, it shouldn't come as a surprise that both organelles have F-type bacterial ATPases, while the eukaryote cytoplasmic ATPases are of the V-type found normally in archaeans (“Chapter IV—You Going to Pay Me to Eat That?”). This would corroborate the suggestion that an archaeon played the role of host cell. In addition, mitochondria and chloroplasts replicate in a manner eerily reminiscent of bacterial binary fission and oftentimes independent of the host cell (though now entirely dependent on the host for overall survival). They also have small ribosomes (70S) where translation begins with the amino acid *N*-formylmethionine instead of “regular” methionine just as it does in bacteria. But perhaps the most provocative evidence is the fact that both organelles possess a single, circular chromosome, and phylogenetic reconstructions indicate that mitochondrial (mt) and chloroplast DNA is most closely related to living α -proteobacteria and cyanobacteria, respectively. In fact, it appears that their genomes have been reduced due to massive horizontal gene transfer into the host nucleus, which is also the reason why various mt-diseases do not demonstrate a maternal inheritance pattern.

More specifically, genes coding for enzymes like mtDNA and mtRNA polymerases are located in the host nucleus, while those coding for mt rRNAs, tRNAs, and of course proteins involved in the electron transport chain, remain in their ancestral home. Perhaps because they were later acquisitions, chloroplast DNA continues to retain a copy of its own RNA polymerase. In short, eukaryotic nuclear DNA is a real life chimera!

But what would motivate such a union in the first place? Traditionally, the rise in O₂ has been viewed as a causal factor in the archaeon- α -proteobacteria merger (or hostile takeover depending on your perspective); the benefits accrued from the far more efficient O₂-mediated oxidation of organic matter being too good of a deal to pass up (i.e., organic carbon oxidized all the way to CO₂ and H₂O). However, oxygen concentrations were less than 1 percent of present atmospheric levels, so it is incredibly unlikely that this was the original reason. A stronger case can be made that the host was just a regular heterotroph, which just so happened to engulf an anoxygenic photoautotroph similar to a purple non-sulfur bacterium. For reasons unknown, the symbiont was not digested and actually started to thrive within its living shelter. To overcome the costs of harboring such a guest, the proto-mitochondrion likely provided a rich supply of reduced carbon products, while also serving as a “septic tank” that would recycle the

⁹ Under the eukaryotes-early hypothesis, Eukarya emerged at around the same time as Bacteria and Archaea through a different type of viral infection. So instead of an archaeal host, it would be a eukaryote host under the same guest scenario.

host's nitrogenous wastes (NH_4^+) into a shared supply of new amino acids and nucleic acids, while maybe even directly fixing atmospheric N_2 . But the benefits of receiving ready-made macromolecules would have paled in comparison to the fundamental conservation of fixed nitrogen, particularly during a Nitrogen Crisis! (“Chapter V—Why Excuse Me! Was it Something I Ate?”).

But the proto-mitochondrion was in for quite a rude awakening when one (or several) branches of eukaryotes successfully propositioned a cyanobacterium to enter the fold. This must have been the metabolic equivalent of a spouse finding out for the first time that their partner went off and started a second family; it probably wasn't pretty. The oxygen released by the proto-plastid (or proto-chloroplast) very likely disrupted the oxygen sensitive processes of the proto-mitochondrion (i.e., carbon and maybe even nitrogen fixation). But under the right set of circumstances, even a polygamous relationship can work out. Although the proto-mitochondrion probably lost its photosynthetic capabilities, like purple non-sulfur bacteria, it may have been able to run its electron transport chain in “reverse” and become a facultative heterotroph. Once oxygen levels were high enough, this arrangement would prove to be an incredibly powerful combination, and it's no wonder that the organisms that continue to carry this legacy dominate our world today. As a testament to the rewards of such metabolic team-ups, similar interspecies marriages remain prominent features of our natural world (i.e., zooxanthellate cnidarians, lichens, various clams and nudibranchs, etc.), but let's not get too far ahead. It would be remiss of us not to briefly comment on some of the unexpected consequences of these wholesale mergers.

One of the most conspicuous features of a eukaryote is of course the nucleus. Now, as mentioned in passing earlier, the presence of nuclei is not unique to Eukarya, and it's unclear how the few examples of bacterial nuclei are related to the more familiar form found in eukaryotes. Traditionally, the origin of the nucleus has been linked to the acquisition of the proto-mitochondrion and the ensuing challenges of managing and controlling the incredible influx of information flowing between host and symbiont. But there may be a more sinister reason.

The genomes of “prokaryotes” contain far more protein-coding regions than eukaryote genomes, which are littered with non-coding sections known as introns. Some scientists have speculated that the proto-mitochondrion was a carrier of self-splicing (group II) introns, which subsequently jumped into a naive host genome. As they began to spread, it became imperative for the host to remove these offensive “breaks” in the code to prevent the manufacture of aberrant proteins. In response to this “Intron Catastrophe,” both the nuclear envelope and the spliceosome may have emerged to temporarily segregate and process the initial mRNA

products of transcription. The growing necessity to efficiently access, analyze, and manipulate the eukaryotic genome may also have led to the emergence of linear chromosomes and telomerase to combat the loss of information in the lagging strand. While it is difficult to imagine how a symbiont set off this chain reaction, it probably wasn't too difficult to assemble something like a spliceosome since it is fundamentally a *ribozyme*, and life has had a long and successful history finding odd jobs for catalytic RNA (“Chapter I—Life's Spork”). We must say though, it's kind of disheartening to learn that life has somehow failed to internalize such an obvious lesson in spite of almost two billions years of practice: If you're going to pick up a stranger, you should always practice safe...

One of the major misconceptions is that all eukaryotes are obligate aerobes that rely on oxygen as a terminal electron acceptor. As it turns out, quite a few eukaryotes, including various amoebas, fungi and even animals (often those associated with hydrothermal vents or parasitic lifestyles), are anaerobic and capable of using a wide range of electron acceptors including nitrate, nitrite, sulfur, and metabolic intermediates like fumarate. Recent findings indicate that the deep oceans were persistently anoxic until around 550 million years ago (mya) so the vast majority of our history has been spent under oxygen-poor conditions. As a result, it is very likely that a form of anaerobic metabolism represents our ancestral condition and it isn't too surprising that our α -proteobacteria symbiont has experienced many different fates depending in large part on the conditions of its captivity.

Our “mitochondrial family” includes great morphological and functional diversity, while sharing the common traits of a double membrane and enzymes involved in the biosynthesis of Fe-S clusters. *Hydrogenosomes*, for example, are considered to be anaerobic mitochondria that produce hydrogen and ATP but no longer possess a vestigial genome (unlike aerobic mitochondria). Moreover, in cases where the proto-mitochondrion appears to have been lost entirely or greatly reduced (i.e., *mitosomes*), the continued presence of mitochondrial genes in the nucleus of all observed eukaryotes suggests its ancestral presence.

While it is true that a single species of purple non-sulfur bacteria demonstrates more metabolic diversity than all the eukaryotes combined, our greater acceptance of the far more restricted role that oxygen has played in our collective history has helped us realize that specialized aerobes like ourselves are the oddity, not the norm. But since we happen to carry the aerobic form of mitochondria, the bulk of past research has naturally focused on our travels, and we shall continue to journey along this course. So what in fact caused oxygen to reach a level where our specialist strategy began to pay off?

Chapter VII—Life's Smorgasbord

The story of oxygen has always coincided with the story of life. Any substantial change in the concentration of oxygen is going to be intimately linked to an equally significant change in Earth's biota since the overwhelming source of O₂ comes from the photobiological oxidation of water. The eukaryote acquisition of a cyanobacterium, the progenitor of present-day chloroplasts, would have added another oxygen producer to the mix. Although a eukaryote carrying photosynthetic organelles would superficially appear to be a “higher” level of organization compared with a simpler “prokaryotic” existence, it may be surprising to learn that multicellularity was not a eukaryotic innovation.

Cyanobacteria can in fact form multicellular chains where a specialized cell called a heterocyst provides the colony with fixed nitrogen while remaining somewhat protected from the damaging effects of its oxygen-producing neighbors. This is a distinct, albeit two-cell, example of a division of labor. Multicellularity has evidently been gained and lost multiple times in the history of life. “Why”, might you ask? Well, assembling such a team-up isn't as simple as going “Hey, if you're not doing anything later, would you like to go split some sugar?”, though communication (cell–cell signaling) and hooking up (cell–cell adhesion) are a vital part of the process (look up the “slime mold” *Dictyostelium discoideum* for an incredible example of overcoming social barriers). The most difficult hurdle actually appears to be metabolic.

Although cell respiration (aerobic chemoheterotrophy) yields far more ATP than glycolysis (32 mol of ATP versus 2 mol of ATP per oxidized glucose), anaerobic fermentation occurs at a rate that's up to 100 times faster! Therefore, a metabolic switch to a higher yield but lower rate represents a form of co-operation. But how in the world can a multicellular, metabolic slowpoke overcome the costs of reduced mobility and essentially more mouths to feed when competing with diverse single-celled organisms capable of gobbling up resources at a much faster rate?

While fast inefficient metabolism always dominates when resources are abundant, the disadvantages of decreased motility become less pronounced under conditions where even single-celled organisms are restricted in movement. Models seem to suggest that it is only favorable for an undifferentiated mass of cells to stick together and switch to a slower, efficient metabolism under conditions of resource scarcity and overcrowding. Given these specific challenges, it isn't surprising that permanent multicellularity was difficult to establish in non-photosynthetic lineages where they couldn't make their own food to offset the costs of not being able to gather resources quickly enough.

From an oxygen perspective, both single- (i.e., phytoplankton) and multi-celled photosynthesizing eukaryotes

flourished in the aftermath of the first Snowball Earth. By about 1.2 bya, multicellular red algae appear on the scene followed shortly by a general expansion of algal species, including “green algae” around 750 mya, which would drive atmospheric oxygen levels up to around 8 percent. The absence of terrestrial multicellular life suggests that UV levels were still quite high, but it is perhaps at around this time that the ozone layer started to form. Just as things began to stabilize, however, two more snowballs would smack life right in the face—the Sturtian (710 mya) and the Marinoan (640 mya)—and the resulting makeover would be incredible.

Imagine you're at a New Year's Eve party with a room full of people. There's only a couple pieces of cake left, but everyone's been doing a lot of drinking and eating so there's no rush. With so many distractions, no one really notices the snowstorm and heaping mounds piling up outside. Suddenly the lights go out. It's pitch black. No one panics, but the minutes turn to hours, and all of a sudden you're officially snowed in. Those last few pieces of cake somewhere in the room start to skyrocket in value; the problem will not only be finding but also controlling this limited resource, and a lot of things can happen when no one is looking.

Out of the cold, rumors swirl of rampant cannibalism and new alliances. Prior to the second and third Snowball Earth events, single-cell on single-cell violence was mostly passive-aggressive by today's standards—lots of chemical interference and a few poisonings but not a whole lot of mechanical stabbing and smashing. Sure, larger cells were capable of engulfing smaller ones (i.e., *phagocytosis*), but grazing pressures must have been limited due to restrictions on size, and you could probably characterize life as mostly a drinking competition hosted by the local photoautotrophs. But the competition must have intensified during successive deep freezes, which would have created a situation where single-celled organisms experienced high cell densities and resource limitations, apparently the only scenario where multicellularity and the efficient use of a resource becomes a viable strategy, particularly for heterotrophs. Out of these dire conditions, it appears that one lineage of eukaryotes, one we're particularly fond of, formed a specific partnership that would transform the traditional “Guess who's coming to dinner?” to more of a “Guess who's going to be dinner?” kind of affair.

Now historically, Eukarya has been divided up into a ragtag group of typically unicellular organisms called “protists” and the three “super” groups made popular by their multicellular members: green plants (Viridophytes), fungi, and animals (Metazoa). With improved phylogenetic reasoning and methodology, we now know that protists are an unnatural grouping, and various “protist” lineages are more related to plants, animals, and fungi than they are to one

another. Focusing on our lineage, the Metazoa, scientists are fairly confident that the single-celled/colonial group of eukaryotes known as the Choanoflagellates represent our living “sister group,” meaning we share a relatively recent common ancestor. Comparing the natural history of choanoflagellates with the oldest surviving branches of the Metazoa,¹⁰ the sponges, it’s pretty safe to assume that the earliest animals were multicellular filter-feeders, gorging on a bounty of single-celled takeout with ready-made amino acids, nucleic acids, complex carbohydrates, and lipids. Whether this started off just before, during, or after the last snowball event is still a matter of molecular debate, but the fossil record unambiguously confirms that by no later than 570 mya (the “Ediacara”), animals reminiscent of multicellular sponges and cnidarians (i.e., jellies, sea pens, anemones, etc.) exhibiting a *tissue-level*¹¹ of organization, plus a few soft-bodied peculiarities, were already firmly established in the deep oceans.

The increasing levels of oxygen generated by the continuing algal radiation probably had a lot to do with the improved ability of animals to produce more rigorous but metabolically expensive adhesion molecules like collagen, allowing greater organizational complexity and hence grazing ability to emerge. But sticking closer together also has its drawbacks: Tissues can become disturbed, releasing individual cells which can then revert to anaerobic fermentation. The subsequent buildup of greater waste products can lead to further disturbances; more cells are liberated and cell–cell adhesion becomes compromised. In other words, a tissue level of organization establishes the pre-conditions for *cancer*, which is by definition a multicellular affliction and one often associated with lower oxygen levels. It is difficult to determine whether a crucial oxygen threshold was reached prior to the infamous Cambrian “explosion”—more mega-tonnage after a long fuse and hand grenades thrown in the pre-Cambrian than a single detonation—but by around 540 mya, a diverse collection of bilaterally symmetrical, *organ-carrying* animals were well on their way to establishing a “dog-eat-dog world.” If it turns out that the molecules are correct, and the origin of the metazoans lies somewhere in the neighborhood of 800–700 mya, perhaps things got incredibly ugly in the dark. But so far, no one is talking.

Although some enigmatic fungi appear to have planted the first flag (~540 mya), by about 460 mya, the descendants of a group of “green algae” made their way onto terra firma, establishing the first land plants (Embryophytes), which were probably reminiscent of today’s mosses, liverworts, and

hornworts. Based on a persistent charcoal record from 450 mya to present, oxygen levels were probably around 12 percent at this stage, enough to sustain fire, and a stratospheric ozone layer was likely well established by now. Mutation rates must have dropped considerably as a result, and the increase in morphological complexity of plant, animal, and fungal groups shortly after points to the greater role amplification mutations would play in generating genetic diversity through redundancy (i.e., gene, chromosome, and whole-genome duplications). In fact, it probably isn’t a coincidence that some of the major radiations of metazoan life have occurred shortly after whole-genome duplications. Of course, with increased morphological complexity, we see a decrease in metabolic diversity and flexibility, as well as a considerable decrease in the incidences of horizontal gene transfer. As a result, the genealogical relationships between species of plants, fungi, and animals became much more “tree-like” than the early “Network of Life.”

In the Carboniferous period (360–200 mya), plants achieved their own version of organ-level organization with the incredible predominance of vascular plants (Tracheophytes) and early seed plants (Spermatophytes). The unbridled success of ferns and early trees doubled the planet’s primary productivity, causing oxygen levels to soar over 30 percent. Success, however, may have come at the cost of starting an ice age possibly due to the excessive removal of carbon dioxide. Millions of years later, the deaths of these plants would ironically spark the Industrial Revolution in a reversal of climate fortunes. The Permian–Triassic boundary bore witness to a slew of extinctions in the aftermath where oxygen levels plummeted back down to 10–12 percent. A group of scaly tetrapods in possession of an air-sac breathing system would be far better adapted to these conditions than any of their contemporaries. And as oxygen levels rose to 23 percent, some of their descendants would grow to absolutely monstrous proportions, while others would get to experience low oxygen levels all over again during mid-flight. Flowers first opened bloom to a parade of insects around 144 mya, but it’s only in the last 200,000 years that modern humans have had a chance to stop and smell the roses.

Is it a coincidence that life emerged in a much more cooperative state after every disaster? Are partnerships naturally forged in cold, dark recesses, and will the same thing occur (or perhaps the reverse) when life is left exposed to heat and high intensity UV radiation? Such events are nearly singular in nature, so it is rather difficult for even a relatively introspective creature to gain assurances from any perceived pattern. Can we expect a similar collaborative response should conditions turn for the worse? Maybe it is foolhardy to idealize the history of life on our planet. After all, the pivotal merger between an α -proteobacterion and an archaeon was more of an example of inept predation than a

¹⁰ I highly recommend comparing the anatomy and physiology of various colonial Choanoflagellates like *Proterospongia* to contemporary sponges.

¹¹ Check out “Placozoans” for insight into the transition from multicellular to tissue-levels of organization.

kumbaya while cuddled in an ice shelter. Who really got the better end of the deal anyway? One partner was imprisoned and forced to be a battery while the other was permanently infected with a genomic disease. We may be wrong, but failed grazing attempts rarely make for happy endings.

Epilogue—Satisfied?

If we had to summarize the history of life up to now, it would be as follows: Life emerged from thermodynamically favorable chemical interactions on early Earth, perhaps within a hydrothermal vent somewhere, sheltered and maybe protected from the violent ebbs and flows of a planet wrestling with growing pains. Before long this cradle became a prison, and it is here where the molecular expansion of life first started. Through divergent selection and co-option within the hydrothermal chimneys, the flow of biological information would first diversify and after the nest was abandoned, new challenges would select from an increasingly complex pool of potential. A period characterized by diverse metabolic solutions to many of the same

problems would then pass through a dark age as the rise in oxygen would force life through a harsh filter roughly two billion years into our story.

The next two billion years would see life undergo far more stabilizing selection than in the past, particularly after the emergence of the ozone layer when mutation rates began to fall and life likely experienced an increase in amplification mutations (i.e., gene and genome duplications), which coincidentally or not, tend to precede a gradual increase in morphological expansion. The first of many major disasters to come may have encouraged the single most important merger of all time, and while entrenched in a second and third, various groups of organisms independently negotiated a permanent social contract in the face of increasing competition for scarce resources. We would emerge as the descendants of one of these proceedings some six hundred million years ago, give or take a few million, our identities now heavily layered and just one of a multitude of sister species that have graced a planet that has barely batted an eye after giving birth to us some four billion years ago.

After delving into the origin of life, Koch and Silver (2005) once humbly quipped: “How do we assemble all

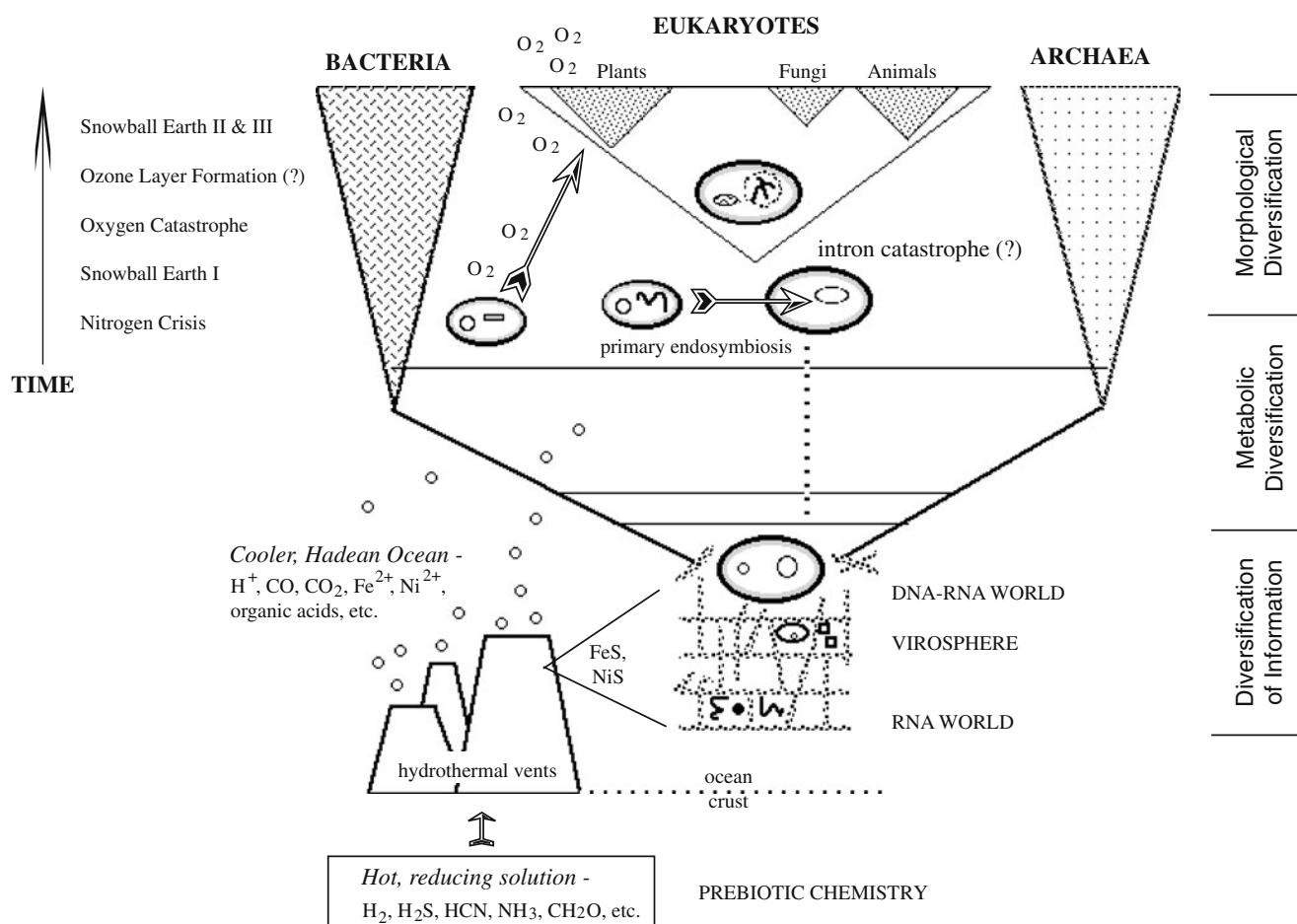


Fig. 1 The origin and diversification of life

these processes, components, and thoughts coherently? We do so in a barely convincing manner today, but made an attempt and hope that our efforts serve useful purposes, especially introducing this topic to a broad audience of microbiologists.” Likewise, I admit this isn’t the most convincing review but I hope it has (re)introduced both old and new topics in a format that is accessible to all (Fig. 1). Any weaknesses in story telling or misconceptions are of course entirely my own. I hope this leads not to dismissal but rather garners further support for teachers who are often faced with the challenge of presenting such topics to an audience that extends far beyond future microbiologists. And although I’ve reached the end of my tales from the deep, *The Neverending Story* will continue on without me; new tales will undoubtedly emerge, and old ones rewritten with new twists and turns. Some things will remain the same, though: Natural selection and co-option will continue to tweak and repackage, while life will continue to evolve, albeit with lenses too preoccupied with the past to plan for the future.

References

- Cunchillos C, Lecointre G. Ordering events of biochemical evolution. *Biochimie*. 2007;89:555–73.
- Koch AL, Silver S. The first cell. *Adv Microb Physiol*. 2005;50:227–59.
- McLennan DA. The concept of co-option: why evolution often looks miraculous. *Evo Edu Outreach*. 2008;1:247–58.
- Further Reading**
- Aledo JC. An early and anaerobic scenario for the transition to undifferentiated multicellularity. *J Mol Evol*. 2008;67:145–53.
- Allen JF, Martin W. Out of thin air. *Nature*. 2007;445:610–2.
- Brooks DR, McLennan DA. The nature of diversity: an evolutionary voyage of discovery. 1st ed. Chicago: University of Chicago Press; 2002.
- Canfield DE, Rosing MT, Bjerrum C. Early anaerobic metabolisms. *Phil Trans R Soc B*. 2006;361:1819–36.
- Carroll SB. Chance and necessity: the evolution of morphological complexity and diversity. *Nature*. 2001;409:1102–9.
- Falkowski PG, Godfrey LV. Electrons, life, and the evolution of Earth’s oxygen cycle. *Phil Trans R Soc B*. 2008;363:2705–16.
- Falkowski PG, Isozaki Y. The story of O₂. *Science*. 2008;322:540–2.
- Feng S, Tian G, He C, Yuan H, Mu Y, Wang Y, et al. Hydrothermal biochemistry: from formaldehyde to oligopeptides. *J Mater Sci*. 2008;43:2418–25.
- Forterre P. The two ages of the RNA world, and the transition to the DNA world: a story of viruses and cels. *Biochimie*. 2005;87:793–803.
- Freilich S, Goldovsky L, Ouzounis CA, Thornton JM. Metabolic innovations towards the human lineage. *BMC Evol Biol*. 2008;8:247–60.

- Gabaldon T, Huynen MA. From endosymbiont to host-controlled organelle: the hijacking of mitochondrial protein synthesis and metabolism. *PloS Comput Biol*. 2007;3:2209–18.
- Gogarten JP, Fournier G, Zhaxybayeva O. Gene transfer and the reconstruction of life’s early history from genomic data. *Space Sci Rev*. 2008;135:115–31.
- Gregory TR. The evolution of complex organs. *Evo Edu Outreach*. 2008;1:358–89.
- Hackstein JHP, Tjaden J, Huynen M. Mitochondria, hydrogenosomes, and mitosomes: products of evolutionary tinkering! *Curr Genet*. 2006;50:225–45.
- Higgs PG, Pudritz RE. A thermodynamic basis for prebiotic amino acid synthesis and the nature of the first genetic code. *Astrobiology*. 2009;9:483–90.
- Howe CJ, Barbrook AC, Nisbet RER, Lockhart PJ, Larkum AWD. The origin of plastids. *Phil Trans R Soc B*. 2008;363:2675–85.
- Jalasvuori M, Bamford JKH. Structural co-evolution of viruses and cells in the primordial world. *Orig Life Evol Biosph*. 2008;38:165–81.
- Kiang NY et al. Spectral signatures of photosynthesis. I. Review of earth organisms. *Astrobiology*. 2007;7:222–51.
- Kirschvink JL, Kopp RE. Palaeoproterozoic ice houses and the evolution of oxygen-mediating enzymes: the case for a late origin of photosystem II. *Phil Trans R Soc B*. 2008;363:2755–65.
- Koonin EV. The origin of introns and their role in eukaryogenesis: a compromise solution to the introns-early versus introns-late debate. *Biol Direct*. 2006;1:22–45.
- Koop RE, Kirschvink JL, Hilburn IA, Nash CZ. The paleoproterozoic snowball Earth: a climate disaster triggered by the evolution of oxygenic photosynthesis. *PNAS*. 2005;102:11131–6.
- Kulp TR. Arsenic (III) fuels anoxygenic photosynthesis in hot springs biofilms from Mono Lake, California. *Science*. 2008;321:967–70.
- Kump RL. The rise of atmospheric oxygen. *Nature*. 2008;451:277–8.
- Mentel M, Martin W. Energy metabolism among eukaryotic anaerobes in light of proterozoic ocean chemistry. *Phil Trans R Soc B*. 2008;363:2717–29.
- Mulkidjanian AY, Makarova KS, Galperin MY, Koonin EV. Inventing the dynamo machine: the evolution of the F-type and V-type ATPases. *Nat Rev*. 2007;5:892–9.
- Mulkidjanian AY, Galperin MY, Makarova KS, Wolf YI, Koonin EV. Evolutionary primacy of sodium bioenergetics. *Biol Direct*. 2008;3:13–31.
- Robinson R. Jump-starting a cellular world: investigating the origin of life, from soup to networks. *PloS Biol*. 2005;3:1860–3.
- Rokas A. The origins of multicellularity and the early history of the genetic toolkit for animal development. *Annu Rev Genet*. 2008;42:235–51.
- Rothschild LJ. The evolution of photosynthesis... Again? *Phil Trans R Soc B*. 2008;363:2787–801.
- Russell MJ. The alkaline solution to the emergence of life: energy, entropy, and early evolution. *Acta Biotheor*. 2007;55:133–79.
- Russell MJ, Martin W. The rocky roots of the acetyl-CoA pathway. *Trends Biochem Sci*. 2004;7:358–63.
- Saul JM, Schwartz L. Cancer as a consequence of the rising level of oxygen in the Late Precambrian. *Lethaia*. 2007;40:211–20.
- Sleep NH, Bird DK. Evolutionary ecology during the rise of dioxygen in the Earth’s atmosphere. *Phil Trans R Soc B*. 2008;363:2651–64.
- Szostak JW. Systems chemistry on early earth. *Nature*. 2009;459:171–2.
- Thauer RK. A fifth pathway of carbon fixation. *Science*. 2007;318:1732–3.