Title

Understanding the Early Major Transitions in Evolutionary History Part 2: Ancient Evolution of Biological Systems and the Biosphere

Authors and affiliations

Primary Author:
Greg Fournier, e:g4nier@mit.edu, t:617-324-6164, Massachusetts Institute of Technology, Department of Earth, Atmospheric, and Planetary Science

Co-authors:
Johann Peter Gogarten, e:gogarten@uconn.edu, t:860-465-6267, University of Connecticut, Department of Molecular and Cell Biology
Aaron D. Goldman, e:agoldman@oberlin.edu, t:440-775-8749, Oberlin College, Department of Biology, Blue Marble Space Institute of Science, Prebiotic Chemistry and Early Earth Environments Consortium
Anton S. Petrov, e:anton.petrov@biology.gatech.edu, t:G404-894-8338, Georgia Institute of Technology, School of Chemistry and Biochemistry, Center for the Origins of Life
Lynn Rothschild, e: lynn.j.rothschild@nasa.gov, t:650-604-6525, NASA Ames Research Center, Space Science and Astrobiology Division
Daniel Segrè, e:dsegre@bu.edu, T:617-358-2301, Boston University, Department of Biology, Bioinformatics Program, Biological Design Center, Microbiome Initiative
Eric Smith, e:dsmith470@gatech.edu, t:720-364-6394, Georgia Institute of Technology, School of Biological Sciences, Center for the Origins of Life
Loren Williams, e:loren.williams@chemistry.gatech.edu, T:(404) 385-6258, Georgia Institute of Technology, School of Chemistry and Biochemistry, Center for the Origins of Life, Prebiotic Chemistry and Early Earth Environments Consortium

Recommendation We recommend that early evolution research, the excavation of the historical record of biology, continue to be a high priority of the Astrobiology Program because of its role in understanding the origin and diversity of life on Earth, its importance in guiding prebiotic chemistry research, and contribution toward understanding the ability of life to adapt to diverse environments. Early evolution research is key to, and must be integrated into the design of, life detection strategies. Finally, there remain keys to alternate trajectories that life could have taken in early evolved clades, providing insights into the possibility of other life forms elsewhere.

Motivation

All extant life on Earth shares a common biochemistry based on a relatively small set of organic molecules (Kluyver and Donker, 1926), the same mechanisms of information storage and inheritance (Woese, 1965), RNA, DNA and protein, enzymatic cofactors and ATP energy currency, a dependence on water, a related cellular organization, and a handful of core metabolic pathways with reactions performed by proteins with a shared ancestry. These central features of life as we know it all evolved in the period between the origin of life and the last universal common ancestor (LUCA) of all extant life. Understanding universalities, their evolutionary elaborations and early major transitions in evolutionary history has is an important component of astrobiology research and has increasing potential thanks to improved research strategies and cross-disciplinary collaborations, as well as advances in molecular and cellular biology, increasing power of computational resources and breadth of bioinformatics databases.

Extant biology contains detailed and interpretable records of pre-biological processes and molecules and of early biology. The excavation of the historical record (top down origins of life research) uses a variety of information including universal features of biology, phylogenetic divergence and the geochemical record. The early history of life on earth provided by top down approaches provides unique information and important perspectives for NASA's astrobiology and exobiology programs. Top down research on early evolution has revealed extensive information on geochemistry of earth, chemical evolution, pre-LUCA biological phenomena and early evolution of bacteria, archaea and eukarya. Top-down guides bottom up approaches and vice versa. Our goal is to help drive the golden spikes in the seams between prebiotic chemistry and early biology, and to guide life detection in other systems.

Understanding how and under what circumstances life increased in complexity, became capable of populating new environments and altering the planet and its atmosphere can inform our expectations of when a planet or moon may be inhabited. Furthermore, in the search for life elsewhere, we only have one biosphere to inform our search for extraterrestrial life. That biosphere was very different in early evolutionary history. Thus, understanding the early biosphere gives us an extended set of features which may be informative of biological processes across a broader time window. For example, multiple pathways of carbon fixation exist in ancient life forms that were abandoned as lineages evolved, modified bases are found in the DNA of several protist lineages, and non-ribosomal protein synthesis provides an extant example of an alternate pathway for peptide synthesis. Many examples of life adapting to extreme environments come from early evolved clades.

The study of early evolutionary history can, therefore, enhance the search for life beyond Earth in several important ways. It can help in developing an explicit theory of the origins of biology. Currently we cannot distinguish transcendental features of life from those that are quirks of evolutionary history of Earth. But this distinction is essential for designing sufficiently broad life-detection methods. Below are several examples of current research areas and future directions within the field of early evolution that center on the evolution of complex molecular systems and their effects on the biosphere.

Earliest synergies of emerging metabolism, bioenergetics, and catalysis

Explaining the emergence of a metabolism that could support macromolecular evolution requires discovering possible synthetic routes to biomolecules and more. Metabolism operates through an interdependence of organic synthesis, free energy transduction, and catalysis, and is organized as a highly integrated network with extensive mechanistic redundancy and multiple uses for intermediates. The complex macromolecular systems that characterized the LUCA, and patterns in the genetic code, represent metabolic order much like that reflected in biology today, suggesting that these interdependencies and at least part of the specification of terrestrial biochemistry evolved jointly with the functionality of macromolecules themselves. Three research themes address the emergence of metabolism as a system above and beyond the inventory of its components and mechanisms:

Which parts of universal metabolism and bioenergetics on Earth were inherited from analogous geochemical processes, and which became possible only when catalysts could be recruited or manufactured to couple exergonic and endergonic reactions that were not geochemically coupled? Today, all biological chemistry is organized around two recycled bioenergetic currencies: one carrying redox potential and mediated by sulfur often in association with metals, the other carrying dehydration potential and mediated by anhydrous phosphates. The emergence of such an organization must have depended on how either substrate coupled to planetary disequilibria, how they came to drive organosynthesis, and when, how, and in what direction of causation they came to be connected to each other, forming the most crucial bridge in bioenergetics. As our understanding of metabolism across living systems today is gradually extended to larger and larger machine-searchable maps [KEGG, BioCyc, KBase, Hatzimanikatis Atlas], it becomes possible to start asking specific questions about plausible steps to have led from geochemistry to biochemistry. The connectivity and stoichiometry of metabolic networks, together with information about the pH and temperature-dependent thermodynamic feasibility of specific reactions and the geological availability of redox gradients can be used to constrain the feasible paths of metabolic evolution. Understanding what reactions constitute thermodynamic bottlenecks along these paths, and what coupling of exergonic and endergonic reactions could overcome these barriers, would help us solve one of the most profound mysteries of the emergence of life, that is, the rise of “energy currency” and of macromolecules as energy transducers capable of using these energy currencies to drive otherwise uphill reactions. While biochemical details of these processes may be different in extraterrestrial life, it is likely that this feature of living systems is fundamental and universal across any chemical life form.

What factors that selected the first peptide folds (likely in association with RNA) brought stable catalytic functions, and what were the earliest feedbacks through which these could have harnessed redox and phosphate energetics, enhancing or controlling organosynthesis? Metabolism emerged as a self-supporting system when synthetic and dissipative reactions became linked through catalysis by macromolecules (Branscomb et al., 2013), possibly as the size and functional diversity of catalysts evolved interdependently with expansions in networks of synthetic reactions. Proteins involved in core bioenergetics and translation share many of the most ancient folds, and catalyze related and sometimes overlapping functions of activation and ligation. In particular, thioester chemistry catalyzed widely across these families (Jakubowski 2017) is also known to link most compounds in the universal metabolic core into a single connected network (Goldford et al., 2017). Increasing evidence from prebiotic chemistry experiments (Sahai et al., 2016), analysis of natural systems (Menez et al., 2018, Sforna et al., 2018), and study of the architecture of metabolic networks seems to suggest that these early biochemical processes were once catalyzed by mineral surfaces and small molecules other than (and prior to) protein and nucleotide polymers. These types of studies open a window to very early stages of biochemical evolution, and could help to clarify the conditions leading to the emergence of self-sustaining metabolic processes from planetary chemistry, and the balance of contingency and universality in the rise of macromolecules.

In parallel, relations are being found among motifs in core folds at the sub-domain level (Lupas et al., 2017, Cheng et al., 2014), which together with fold progression frozen into the ribosome (Kovacs et al., 2017) and diversification in the aaRS (Fournier et al.,. 2011), promise to reveal rules of fold dynamics that would have governed the modes of catalysis available in each evolutionary era, and the properties they would have required from amino acids synthesized in that era. A more rule-based evolutionary dynamic for early folds would not only constrain our interpretation of imperfectly-recorded history; it also would explain how core structures and functions selected at first for enhancement of energy transduction and organosynthesis could then have served as a foundation for ongoing evolutionary diversification of special capabilities (Ranaan et al., 2020).

Pre-Darwinian evolution and the search for physico-chemical principles that underlie the emergence of life

Is life a phenomenon that may be expected to emerge generally within systems of sufficient complexity, distance from thermodynamic equilibrium, or other physico-chemical properties? Is it possible to conceive theoretically and/or assemble in the laboratory a system able to display emergence of self-reproduction and open-ended evolution? How would such a *system depend on boundary conditions related to the physico-chemical state of a planet, such as availability of a diverse chemistry and energy gradients? Can a rudimentary pre-Darwinian evolutionary process arise even in the absence of macromolecules? How would such an evolutionary process appear to us, and how could it be described mathematically?*

These questions are not new in the study of early life, and have been the subject of research at the crossroads of biology, chemistry and physics. Theoretical and experimental work to address them constitutes an important complement to research aimed at understanding life-as-we-know-it. These avenues of research will likely grow in coming years, especially as barriers across disciplines are lowered, and opportunities for collaborations between physicists, chemists and biologists become more widespread. Advances in nonequilibrium thermodynamics and statistical physics, and the formalization of biological knowledge will jointly provide a chance for transformative quantitative studies of the emergence of life. In addition to providing a deeper understanding of how "living matter" differs from inanimate matter in terms of its physical properties, this line of research may provide a fundamental characterization of living systems as a category anywhere in the Universe, broadening the scope of signs to look for beyond specific molecular players.

The origin of oxygen producing photosynthesis

When did oxygenic photosynthesis evolve? How is atmospheric oxygenation and geochemical oxidation related in timing and mechanism to the emergence of this metabolism? How important is oxygen to evolving and maintaining a habitable biosphere with a detectable biosignature?

One of the most enigmatic and controversial events in early evolution is the invention of oxygen producing (oxygenic) photosynthesis. While the timing of the "great oxygenation event" (GOE) at 2.3-2.4 billion years ago is well-established (Luo et al., 2016), the timing for the origin of oxygenic photosynthesis remains poorly constrained, with much indirect geochemical evidence of pre-GOE oxygenation remaining intriguing, but controversial (Frei et al., 2009). Even less constrained is the history of the biosphere and biogeochemical fluxes with respect to the history of photosynthesis, including anoxygenic and oxygenic variants. The two photosystems, while homologous, represent two different types using different electron acceptors, and each type is found as a single photosystem in different bacterial groups. It is clearly established the genes encoding photosynthesis have been transferred between different bacterial phyla, and that single-photosystem anoxygenic photosynthesis must have preceded two-photosystem oxygenic photosynthesis (Sanchez-Baracaldo & Cardona, 2020); however, debate continues on how the two photosystems that work in series to lift electrons from the redox potential of water to the level of NADP came together in a single electron transport chain [Hamilton, 2019]. Similarly, the origins of the oxygen evolving complex remain cryptic. How much primary productivity, and how large of a biosphere, could anoxygenic photosynthesis support on the early earth before the origin of oxygenic photosynthesis? What electron donors were available, and in what amounts? How early did oxygenic photosynthesis evolve, and how much time elapsed before oxygen levels finally irreversibly rose in the atmosphere? What processes could potentially forestall this accumulation? How and when did aerobic respiration evolve, and did these oxygen consuming metabolisms delay atmospheric oxygenation as a potential biosignature mask?

The increased availability of protein structure data combined with tools to search and classify relationships between similar structures has the potential to unravel some of the remaining mysteries surrounding the origin of oxygen producing photosynthesis: From which precursors did the water splitting complex originate? Did the duplications that resulted in the dimers that form each of the two reaction centers occur independently, or was this more ancient event, whose mapping in molecular phylogenies is complicated by gene conversion events? Additionally, in the next decade, continued improvements in sequence-based molecular clocks and divergence time estimates will improve our timeline for the divergence and diversification of cyanobacteria and anoxygenic phototrophic lineages. Molecular clocks applied to the horizontally transferred components of photosynthesis, including photosystems, pigment biosynthesis genes, and carbon fixation machinery will all directly constrain the timing of these major evolutionary innovations, as well as inform interpretations of the geochemical and microbial fossil records, and testing of ecological and biogeochemical hypotheses for the Early Earth.

Learning about planetary metabolism from the study of how microbial communities contribute to biogeochemical cycles

Through large metagenomic sequencing efforts worldwide, we are now starting to appreciate the enormous diversity and complexity of the microbial communities that populate most locations on our planet today. In addition to leading to the discovery of new branches of the tree of life, and new genes - often with unknown functions - these efforts are starting to provide insight into how microbial communities assemble and change as a function of environmental conditions. On top of the metabolic networks present in individual organisms, microbial communities seem to engage in metabolic competition and exchange across species, effectively operating as ecosystem-level networks that respond (and contribute) to environmental changes at multiple scales. For example, division of labor seems to emerge spontaneously in microbial ecosystems, providing possible clues for evolutionary strategies that can result in diversification and specialization of roles. These ecosystem-level microbial metabolic processes can help refine models of global biogeochemical cycles, and are an integral part of current efforts to understand the past and future history of our biosphere. Research on microbial communities has therefore the potential to provide a different perspective on how life could emerge and be maintained at an ecosystem-level.

Conclusions Understanding early evolutionary history and the major evolutionary transitions that shaped our biosphere have been important goals of NASA's Astrobiology and Exobiology Programs. The top down approaches must continue to play indispensable roles in origin of life research as well as the design of life detection strategies, which together constitute the majority of astrobiology research overall. The next decade will likely see major advances due to a number of recent advances and innovations described in this document. Early evolution research in the United State must be supported for the benefit of future NASA missions.

References

Cleaves II HJ, et al. (2012) Chem. Soc. Rev. 41:5502--5525.
Branscomb E & Russell MJ. (2013) Biochim. Biophys. Acta 1827:62--78.
Fournier GP & Alm EJ. (2015) J. Mol. Evol. 80:171--185.
Fournier GP, et al. (2011) Orig Life Evol Biosph. 41(6):621--32.
Frei R, et al. (2009) Nature. 461(7261):250-3.
Cheng H, et al. (2014) PLOS Computational Biology 10(12): e1003926.
Goldford JE, et al. (2017) Cell 168:1126--1134
Hamilton TL. (2019) Free Radic Biol Med. 140:233-249.
Jakubowski H. (2017) Life 7:6.
Kluyver AJ, Donker HJL (1926) Chem Zelle Gewebe, 13:134–190
Kovacs NA, et al. (2017) Mol. Biol. Evol. 34:1252--1260.
Luo G, , et al. (2016) Sci Adv. 2(5):e1600134.
Lupas AN & Vikram A. (2017) J. Struct. Biol. 198:74--81.
Menez B, et al. (2018) Nature 564:59--63
Ranaan H, et al. (2020) Proc. Nat. Acad. Sci. USA 117:7193--7199
Sahai N. et al. (2016) Elements 12, 389-394
Sanchez-Baracaldo P & Cardona T. (2020) New Phytol. 225(4):1440-1446.
Sforna MC, et al. (2018) Nature Comm. 9:5049
Woese CR (1965) Proc Natl Acad Sci USA. 54:1546–1552