ASTROBIOLOGY
Volume 00, Number 00, 2026
Mary Ann Liebert, A Part of Sage

DOI: 10.1177/15311074261452809

# Rethinking the Last Universal Common Ancestor of Life: Network Convergence and the Root of the Tree

Jessica C. Bowman,[1,2] Nigel Goldenfeld,[3] Karyn L. Rogers,[4,5] Anton S. Petrov,[1,2] and Loren Dean Williams[1,2]

**Abstract**

The tree of life is rooted at the "origin of life." One model holds that core biochemistry, which includes the genetic code, the ribosome, biopolymer backbones, and amino acid and nucleotide monomer alphabets, was inherited vertically from a single origin of life. In this model, core biochemistry is a frozen accident that reflects prebiotic chemistry. In an alternative model explored here, life arose across diverse planetary environments and generated diverse biochemistries that competed and cooperated. These biochemistries converged through selection driven by the "network effect." The network effect conferred greater fitness on participants in increasingly dominant biochemistries: the more extensive the adoption of a biochemistry, the greater the benefits for systems using it. In this model, the evolution of core biochemistry was driven, in part, by compatibility, integration, and coordination. The last universal common ancestor (LUCA) in this model represents a diffuse tipping process—where biochemical convergence reached critical mass. LUCA is a process of convergence rather than a specific organism or collection of organisms. At the tipping point, the biosphere committed to the transition from competing biochemical platforms to a universal standard. After the tipping point, biological innovation exploded, with fixed core biochemistry. This model makes testable predictions: core biochemistry should show evidence of evolutionary optimization rather than frozen accidents; core biochemistry should show molecular entanglement that reflects incremental coevolution; and biosynthetic pathways should differ from prebiotic chemistry. These predictions appear to be supported by observations. Key Words: Network effect—Communal evolution—Origin of life. Astrobiology 00, 000–000.

## 1. The Tree of Life

The tree of life (TOL) (Moody et al., 2024; Hug et al., 2016; Doolittle and Brunet, 2016) provides a unifying framework for organizing and interpreting biology. It represents evolutionary relationships through common ancestry and divergence. The TOL integrates an extraordinary range of forms, functions, and strategies; it relates lions to butterflies and iron-reducing bacteria to thermophilic archaea. It reflects evolutionary innovations across metabolism (Marcet-Houben et al., 2007), genomic organization (Merkl and Wiezer, 2009), cellular architecture (Siefert and Fox, 1998), body plans (Willmore, 2012), morphology (Wiens, 2004), and reproductive strategy (Gomez-Mestre et al., 2012).

The earliest branches of the TOL (Fig. 1) diverge at the last universal common ancestor (LUCA) (Kyrpides et al., 1999). LUCA represents mature biology—cells with biopolymers, ribosomes, a canonical genetic code, membranes, and metabolic networks. The root of the TOL, the antecedent to LUCA, gave rise to these systems and processes. The root, which starts with chemistry and ends with LUCA, represents the emergence and early evolution of life. Here we describe the foundational assumptions, rationales, and uncertainties of models for the root of the TOL. Root models carry implications not only for life's emergence but for defining LUCA and the character of biospheres. Assumptions about the root of the TOL shape interpretations of the processes that gave rise to life and the environments that supported the early evolution of life on Earth. These interpretations about the nature of early life influence the search for life in the Universe.

## 2. Core Biochemistry

Evolutionary theory predicts diversity; biochemistry reveals universality. The vast diversity of biological forms rests on a highly conserved biochemical core (Jacob, 1977). The TOL, from LUCA to the present, from Bacteria to

[1] NASA Center for Integration of the Origins of Life (iCOOL), Georgia Institute of Technology, Atlanta, Atlanta, Georgia, USA.
[2] School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia, USA.
[3] Department of Physics, University of California San Diego, La Jolla, California, USA.
[4] Department of Earth and Environmental Sciences, Rensselaer Polytechnic Institute, Troy, New York, USA.
[5] Rensselaer Astrobiology Research and Education Center (RARE), Rensselaer Polytechnic Institute, Troy, New York, USA.

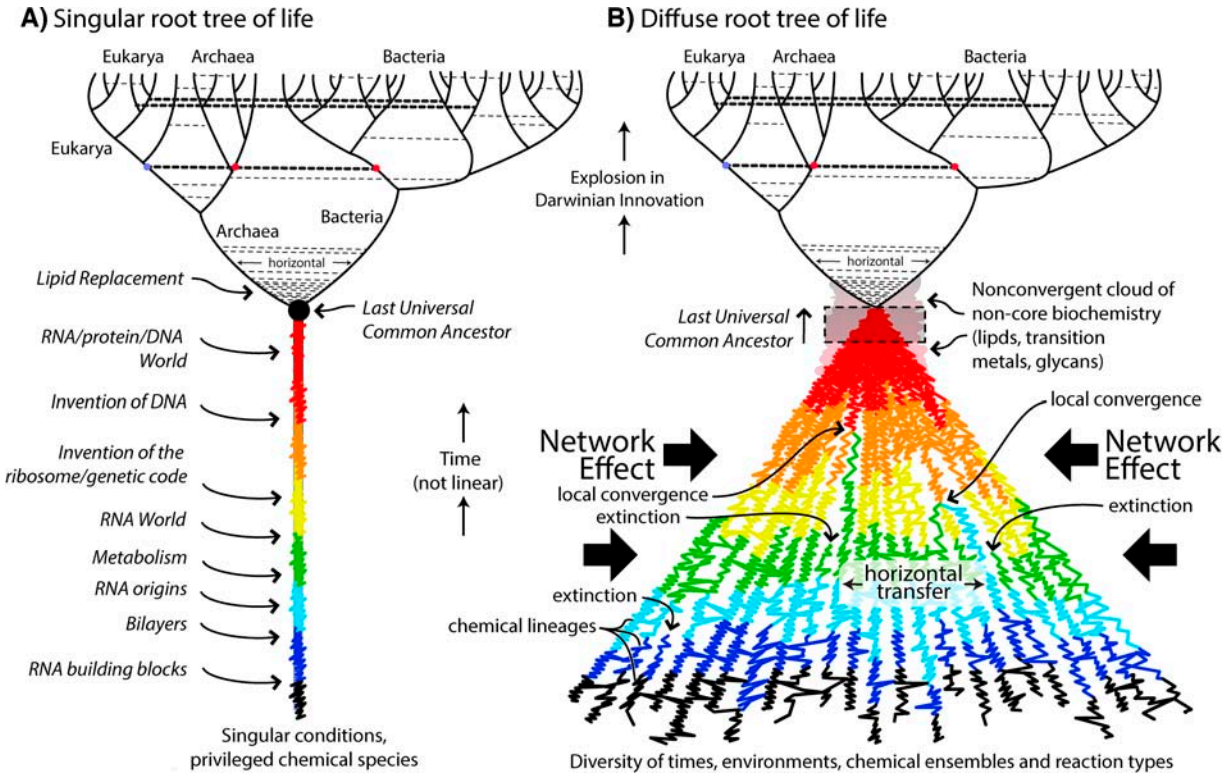


**FIG. 1.** Two models of the root of the tree of life. **(A)** The conventional singular root model annotated by the RNA World hypothesis. **(B)** A network convergence model with many interactive roots forced to convergence by the network effect. This model initiates with diverse chemical species and systems distributed across diverse environments. The influence of the network effect on diversity is indicated by thick arrows. Here the phases begin with chemical lineages near the bottom and progress to networked autocatalytic cycles near the top. Convergence, horizontal transfer, and extinction occur continuously, in all phases. The last universal common ancestor (LUCA) is a process during which the system tips from multiple competing biochemical platforms to a single universal core biochemistry. Membrane systems, transition metal compositions, and glycans maintained a cloud of diversity. Horizontal gene transfer is indicated by thin horizontal dashed lines. Pre-LUCA horizontal transfer is indicated by intersecting lines. Endosymbiosis, including eukaryogenesis and plastid acquisition, is indicated by thick dashed lines.

Archaea to Eukarya, conserves core structures, molecules, and processes. All cellular life contains informational polymers with specific backbone chemistries and chiralities—RNA with phosphodiester-linked ribose, DNA with phosphodiester-linked deoxyribose, and proteins with polypeptide backbones; a restricted set of informational polymer building blocks—nucleobases, ribose and deoxyribose, and L-amino acids; energy currency and cofactors—ATP, GTP, $NAD^+$, FAD, and coenzyme A; the translation system—ribosomes, tRNAs, mRNAs, and the canonical genetic code; essential metals that include sodium, potassium, and magnesium; and water as both a solvent and as a ubiquitous reaction constituent (reactant and product) (Frenkel-Pinter et al., 2021; Nelson et al., 2021; Bernier et al., 2018). Nowhere in Earth's biosphere do we find ribosomes composed of biopolymers other than RNA and protein, ribosomes that synthesize anything other than proteins, or proteins synthesized by anything other than the condensation of amino acids (Woese, 2000).

The universal biochemical core captures only part of life's chemistry. Membrane compositions across the TOL are highly variable; bacteria and archaea use fundamentally distinct lipids (Siliakus et al., 2017). Even within a single species such as humans, tissues and organelles display striking lipid variability (Harayama and Riezman, 2018). Glycans show extensive diversity in linkages and modifications across bacteria, archaea, and eukarya (Gagneux et al., 2022). The noncanonical amino acid pyrrolysine is found in four archaeal and five bacterial genera (Guo et al., 2022; Fournier et al., 2011; Gaston et al., 2011). Similarly, transition metal usage is not universal—some organisms lack iron (Posey and Gherardini, 2000), while others lack zinc-dependent enzymes (Zerkle et al., 2005). Core biochemistry persists within a cloud of chemical diversity.

### 3. A Singular Root

The universality of core biochemistry is commonly interpreted as evidence that the TOL has a single root—that life on Earth arose from a single origin (Moody et al., 2024; Harrison et al., 2023; Theobald, 2010; Koonin and Novozhilov, 2009; Crick, 1968). Under this model, a single prebiotic chemical system crossed the threshold to biology (Fig. 1A), and all living systems inherited their common biochemical core from this singular origin through vertical inheritance. As Darwin stated, "all organic beings which have ever lived on this earth have descended from some one primordial form" (Darwin, 1959). This singular origin subsequently

diversified at LUCA and spread through Earth's lithosphere, hydrosphere, and atmosphere while retaining its ancestral biochemical core.

### 3.1. Boutique environments and unlikely events

A singular root implies that amid the vastness of Earth's 4.5-billion-year history with all of its different environments in space and time, only one prebiotic system crossed the threshold into biology and survived. This exceptional rarity has been attributed to the requirement for unlikely chains of events that include specific meteorite impacts, volcanic eruptions, and hydrothermal activity; extended sequences of complex organic reactions; purification and transport of fragile molecules; evaporative basins concentrating and organizing reactants; and episodic compartmentalization, scaffolding, and templating, along with fortuitous and well-timed shifts in temperature, pH, redox potential, and salinity.

### 3.2. Frozen accidents and chemical vestiges

A single root of the TOL implies that extant biochemistry carries interpretable information about life's deepest origins. If all life descends from a biochemical system whose core features became locked in place early, then modern biochemistry provides a window into life's beginnings. In this model, the molecules of asteroid Bennu (Furukawa et al., 2026; Glavin et al., 2025) contain information about the earliest biochemistry. Biology's reliance on electrochemical gradients reflects continuity with the geochemical proton gradients that predated life (Harrison et al., 2023). The informational and catalytic properties of RNA suggest a transitional system in which one polymer served dual roles of heredity and catalysis (Benner et al., 2020).

In this view, biochemistry is the product of specific prebiotic chemistries that date from the early solar system and/or Hadean environments, with the implication that biochemistry retains vestiges of prebiotic chemistries. Biochemistry thus functions as a kind of molecular fossil that reports on chemical systems that operated before the onset of evolution. A single root implies that the foundations of biochemistry—the genetic code, the amino acid alphabet, the nucleotide alphabet, and aspects of metabolism—are "frozen accidents," embedded from temporally remote chemical environments.

### 3.3. Universal biochemistry reflects fundamental constraints

A single root of the TOL frames our expectations for life in the Universe by suggesting that core biochemistry reflects fundamental constraints. While chemistry allows for vast combinatorial diversity, only a narrow chemical subset appears compatible with the functional requirements of life. Under this view, physical and chemical constraints channeled life's emergence into a narrow biochemical space. Similar principles should govern life's origin elsewhere in the Universe. Features of Earth's biochemistry—such as the use of nucleic acids for information storage, proteins for catalysis, and electrochemical gradients for energy—would reflect general solutions to universal challenges. This view suggests that life on other Earth-like worlds could converge on broadly similar biochemical architectures.

## 4. A Diffuse Root

An alternative to the single root model was proposed by Vetsigian, Woese, and Goldenfeld (VWG) (Vetsigian et al., 2006). VWG showed how a diffuse system could evolve in the era that preceded LUCA. VWG rationalized their model on three salient features of the translation system:

i Optimality—the genetic code is highly robust to errors because functionally similar amino acids are clustered in the code (Koonin and Novozhilov, 2017; Butler et al., 2009; Freeland et al., 2000; Haig and Hurst, 1991; Freeland and Hurst, 1998; Woese, 1965).
ii Rapidity of evolution—the translation machinery emerged within 300–700 million years (Moody et al., 2024; Weiss et al., 2018).
iii Uniqueness—the code is universal and has remained essentially unchanged for 3.8 billion years of biological evolution (Woese, 2000).

Extending the VWG framework, Matange, Marland, Frenkel-Pinter, and Williams (MMFW) (Matange et al., 2025) demonstrated that the backbones of core biopolymers are the products of evolution in the era that preceded LUCA. MMFW identified five hallmarks of biopolymer evolution:

iv Polyfunctionality—core biopolymer backbones form a variety of elaborate assemblies with diverse functions.
v Function-switching—core biopolymer backbones change structure and function in response to subtle environmental or chemical perturbations.
vi Molecular recognition—core biopolymers recognize themselves and others through complementary molecular interactions.
vii Recalcitrance—core biopolymers modulate their own kinetics and thermodynamics of chemical transformation.
viii Emergence—the chemical and physical properties of monomers change significantly upon polymerization.

### 4.1. Aminoacyl-tRNA synthetases

Additional evidence for the emergence of core biochemistry via pre-LUCA evolution is seen in the phylogeny and interactions of aminoacyl-tRNA synthetases (AARSs), which charge tRNAs with cognate amino acids (Gomez and Ibba, 2020). AARSs fall into two structurally unrelated classes (Class I and Class II) that recognize the opposite faces of tRNA. Phylogenetic analyses of AARSs indicate extensive pre-LUCA evolution. The AARSs descended from multiple distinct ancestors (Fournier et al., 2011; O'Donoghue and Luthey-Schulten, 2003; Woese et al., 2000; Wolf et al., 1999) and could have originated multiple times. Evolutionary logic indicates that the number of surviving AARS classes (two) must be far smaller than the number of ancestral lineages, which implies extensive extinction and evolutionary filtering. AARSs therefore emerged in concert with a culling of a vast prebiotic repertoire of building blocks (Glavin et al., 2025; Wogan et al., 2020) through chemical evolution during the emergence of biopolymers, genes, and translation.

### 4.2. The ribosome

The ribosome is composed of two subunits that emerged and evolved independently. The functional interface that couples the decoding subunit (small subunit) and the synthetic subunit (large subunit) of the ribosome is a relatively late acquisition during the evolution of the ribosome (Petrov et al., 2015; Petrov et al., 2014). Each subunit underwent extensive independent growth before its association as part of a functional translation system. As with the AARSs, the number of surviving ribosomes (one) is anticipated to be far smaller than the number of ancestral ribosomes, implying extensive extinction and evolutionary selection in chemical environments.

### 4.3. Prebiological evolution

VWG developed a dynamical systems model of code emergence through collective networked evolution. This model shows how statistical ensembles of building blocks can produce statistical ensembles of proto-proteins and how precise functional genes and translation machinery and an optimized code can emerge from imprecise translation. In this model, lineage is communal. Without horizontal transfer among lineages, the model evolves slowly and becomes trapped in nonoptimal, nonunique codes. Horizontal transfer refers to acquisition from contemporaneous lineages, whereas vertical transfer refers to inheritance from parent to offspring within a lineage (Soucy et al., 2015). When horizontal transfer was incorporated (Fig. 1B), simulations rapidly and repeatedly converged to a single optimized genetic code, which thus explains the salient features of translation (i–iii, above).

Why does horizontal transfer succeed where vertical transfer alone fails? Horizontal transfer is nonlocal in fitness space and can make large jumps that enable exploration of distant regions of the fitness landscape in parallel across diverse environments. By contrast, vertical transfer is local in fitness space and promotes trapping in local optima. In network convergence (Fig. 1B), horizontal exchange of biochemical components—including different polymer backbones and monomer alphabets—explored distant regions of the fitness landscape.

### 4.4. Pre-LUCA lineages and the LUCA transition

The lines that lead up to LUCA in Figure 1B represent chemical lineages. Chemical lineage is chemical kinship through transformation. In primitive forms, toward the bottom of the figure, chemical kinship is chemical relatedness through transformation—thioesters, esters, and amides are chemical kin related by reactivity descent, each derivable from its predecessor by acyl transfer. In its elaborated form, approaching LUCA, chemical lineage involves successions of autocatalytic cycles in which descendants are related to ancestors by transformation of reactants, products, and catalytic mechanisms. Ancestral autocatalytic cycles interact and adapt to produce daughter autocatalytic cycles. The figure traces chemical lineages from primitive chemical kinship in the earliest phases to autocatalytic descent in the later phases.

The combined data (Matange et al., 2025; Vetsigian et al., 2006) point to the coevolution of core biochemistry, including the translation system, with complexity. The biopolymer backbones, side chain alphabets, and translational precision annealed, along with the genotype–phenotype distinction, allowing network-level mechanisms to efficiently enable horizontal transfer.

The evolutionary progression requires (a) diverse early pre-LUCA translation with diverse proto-biopolymers, (b) late pre-LUCA convergence of these systems toward biochemical universals, and (c) post-LUCA dominance of vertical gene transfer, driving sequence-level diversification. Critically, as MMFW noted, effective horizontal gene transfer (HGT) post-LUCA requires not only a universal genetic code but also universal biopolymer backbones and side chain alphabets. HGT and other sharing processes constrain the biochemical core.

A networked system that arises with diverse biochemistries, driven to biochemical universality by selection (Fig. 1B), is a predictive model consistent with available data. In this framework, biopolymer backbones, monomer alphabets, and the genetic code arose by a network-driven evolutionary process that converged at LUCA. This conclusion echoes earlier suggestions by Woese and Fox (1977) and by Jacob (1977) that evolution underwent fundamental transitions during the emergence of life.

## 5. Networking

Here, we propose that the extant biosphere is a network unified by a biochemical core. The network is linked across scales, from molecules to cells to ecosystems. Ecosystems presuppose core biochemistry; without it, trophic interactions, biogeochemical cycling, and coevolution would not be possible. Core biochemistry links organisms through shared chemical elements and redox couples (Sardans and Penuelas, 2024). At the organism level, core biochemistry permits organisms to consume one another and exchange metabolites through trophic and syntrophic relationships. It enables pathogenesis, symbiosis, and endosymbiosis (Margulis and Sagan, 2008). At the cellular level, it underpins HGT, expression of foreign proteins, and mobility of genetic elements such as plasmids and viruses. For multicellular organisms, it makes developmental coordination possible; shared regulatory mechanisms enable diverse cell types to coordinate and communicate.

Networking does not preclude competition; it is a precondition. Organisms compete for the same resources, prey on one another, and coevolve because they share a core biochemistry.

### 5.1. A networked path to core biochemistry

Metcalfe's "law" states that the value of a network grows with its adoption (Metcalfe, 2013). The network effect explains why a single core biochemistry would emerge from an initially diverse biochemical landscape. Once one biochemistry achieved sufficient adoption, alternative biochemistries would become increasingly isolated—unable to develop relationships with dominant systems or share metabolites, genetic innovations, and other resources—which would create runaway selection for biochemical universality.

Biochemical systems that evolved in more common and widespread environments were more likely to achieve

dominance than those that arose in rare or restricted conditions. Environmental ubiquity drove extensive adoption, which in turn created positive feedback loops that accelerated convergence on a single, universal biochemistry. Once network lock-in occurred, even superior biochemical alternatives would have been competitively disadvantaged, unable to interact with the dominant biochemistry. These isolated biochemistries would have been driven extinct, regardless of any intrinsic advantages.

Systems subject to the network effect are common and familiar (Goldenfeld, 2014; Metcalfe, 2013; Metcalfe, 1995). Although computational analogies to biology are fraught, the internet offers a striking and illustrative example of the network effect (Greenstein, 2015). In the early days of computing, many networking protocols competed—DECnet, NCP, IPX/SPX, AppleTalk, SNA, and others. Over time, this diversity converged to a single protocol. Universal adoption of TCP/IP enabled interoperability between different types of hardware and software, which conferred returns on all participants. The advantages of the universal standard are profound; they span the technical, economic, and social domains. This convergence did not occur because TCP/IP was inherently superior to all other solutions but because universality conferred a systemic advantage, because TCP/IP was good enough, and because of contingent political factors (Townes, 2012). In the early 1980s, the US Department of Defense mandated TCP/IP adoption. It cut off all noncompliant systems from ARPANET—the networking protocol equivalent of an environmental driving factor in biochemical evolution. In the 1990s, networking crossed a tipping point to a single protocol (Townes, 2012). From the time of this "LUCA" of networking, all non-TCP/IP networking protocols were destined for extinction. This time was followed by intense innovation and the spread of the internet, which demonstrated the catalyzing power of the network effect.

### 5.2. The tipping point

LUCA represents a tipping point (Gladwell, 2006), or more correctly, a protracted tipping process. This tipping process marks a broad threshold where convergence reached critical density and universal core biochemistry began displacing competing platforms. The threshold was not instantaneous but was distributed across time, chemical networks, and ecological niches, with no single moment that marked the crossover. Beyond a threshold, the integrated biosphere exchanged energy and information, molecules and metals, and biological evolution accelerated.

### 5.3. What Is LUCA?

In the network convergence model, LUCA is the final processes of biochemical convergence. LUCA is not a specific organism or collection of organisms (Fig. 1B). At or near LUCA, lineages with divergent biochemistries were outcompeted and went extinct. LUCA was followed by intense biological evolutionary innovation (with fixed core biochemistry), which remains ongoing, nearly 4 billion years later.

## 6. Predictions of Networking

TOL root models make predictions about biology. Under a single root model, core biochemistry—including the genetic code, biopolymer backbones, and monomer alphabets—reflects inheritance from a single origin, with universal features reflecting prebiotic chemistry frozen in place by early constraints. By contrast, the network convergence model predicts that these same features are highly evolved, arising from evolutionary optimization and convergence. Indeed, the genetic code, biopolymer backbones, and monomer alphabets are products of evolution (Matange et al., 2025; Mayer-Bacon and Freeland, 2021; Vetsigian et al., 2006; Philip and Freeland, 2011; Freeland and Hurst, 1998) that support network convergence over a single root.

### 6.1. Entanglement

Network convergence predicts deep entanglement: reciprocal functional and chemical dependencies between different molecular classes. Entangled molecules cannot be synthesized or sustained independently. Entanglement is a hallmark of evolution. It manifests as webs of interdependence in which function and identity emerge not in isolation but through mutual reliance. Biological molecules form codependent networks of synthesis and function, where no component can exist or operate without the others (Lanier et al., 2017). For example, proteins and RNA are entangled: proteins are synthesized by RNA in ribosomes, and RNAs are synthesized by protein polymerases. Under network convergence, such entanglement emerged gradually, *en route* to universality. Over time, selection by network convergence both reinforced advantageous molecular partnerships and eliminated incompatible alternatives; it progressively integrated the network into a tightly entangled system at LUCA. A single root requires that mutually dependent systems arose independently, without prior refinement.

### 6.2. The non-networked cloud

Network convergence predicts that features that do not contribute to cooperation or resource sharing would escape convergence and retain heterogeneity—a pattern most evident in lipids, transition metals, and protein translation factors. The escape from convergence is indicated by the LUCA cloud in Figure 1B. Humans possess a cloud of approximately 100,000 distinct lipids (Conroy et al., 2024; Harayama and Riezman, 2018), compared with around $\sim$500 amino acids (20 core proteinogenic amino acids) (Genchi, 2017). The composition and use of transition metals as ligands and cofactors are variable. The essential translation initiation factors bIF3 of bacteria and a/eIF1 of archaea and eukaryotes participate similarly in start codon selection but do not share an ancestor (Schmitt et al., 2020; Maduzia et al., 2010). This pattern is precisely what the network convergence model predicts: universality in features under selection for compatibility and heterogeneity in features without such constraints. Under a single root, there is no clear basis for expecting this differential pattern. Network convergence predicts biochemical and code universality within any given biosphere but diversity between biospheres. This prediction is testable if future NASA missions discover and characterize life on other worlds.

### 6.3. Accretion

Finally, network convergence predicts that pre-LUCA molecular systems, including the ancestors of the ribosome and RNase P, accreted RNA ancestral elements from a common pool over chemical evolutionary history. In a network convergence scenario, diverse chemical lineages and assemblies would develop and exchange functional modules. Successful modules would be incorporated into emerging assemblies through horizontal transfer. This process would leave distinctive molecular fingerprints—discrete structural substructures, each folding competent and layered via accretion.

This prediction is observed. Structural analysis reveals stepwise accretion of folding-competent RNA elements into both the ribosome (Petrov et al., 2015; Petrov et al., 2014) and RNase P (Petrov et al., 2026). Modules exchanged between various assemblies (Petrov et al., 2026). Mosaic structures are built from horizontally transferred components. This outcome is the expected consequence of horizontal transfer processes operating in a network.

### 6.4. Diffuse origins

Network convergence does not require coincident unlikely events or unique chemistries in specific temporal sequences in boutique environments. The global requirements may have been as simple as a planet with oceans and land and organic molecules, a planet that rotates on its axis at a distance from its star that induces water cycles between liquid and vapor phases. Diverse environments—each containing organic and inorganic inventories—would have sufficed.

In this model, molecular ensembles contained diverse linkage chemistries, including thiol ester, ester, peptide, and phosphodiester bonds (Haas et al., 2024; Frenkel-Pinter et al., 2022), a diverse side chain alphabet (Glavin et al., 2025; Roche et al., 2023; Schmitt-Kopplin et al., 2010), and a broad variety of sugars (Furukawa et al., 2026; Wang and Yu, 2024; Eschenmoser, 1997). Ensembles of molecules cooperated and competed in advance of, or alongside, the surviving building blocks (Hud et al., 2013). Interactions within and between ensembles gave rise to mutually reinforcing relationships such as autocatalytic networks that engaged in selection via increasingly biological dynamics (Williamson, 2024; Xavier et al., 2020). Transitions of early chemical lineages to biopolymers were facilitated by autocatalytic selection and closure. Autocatalysis has persisted in core biochemistry (Könnyű et al., 2024; Barenholz et al., 2017).

## 7. Summary

The distinction between a singular root and a diffuse root that anneals via network convergence is fundamental. Under a singular root model, universal biochemistry reflects vertical inheritance from a single origin. Under network convergence, universal biochemistry reflects convergence from diffuse origins through evolutionary processes. These two models can be distinguished by their predictions; the observations presented here appear to support network convergence.

## Acknowledgments

The authors thank Wehbi Sawsan, Joanna Masel, Jennifer Glass, Raphael Rosenzweig, Ram Krishnamurthy, Michael Lynch, and David Baum for helpful discussions. N.G. thanks Carl Woese and Kalin Vetsigian for their collaboration on the VWG article and thanks Gary Olsen, Norm Pace, and Eugene Koonin for helpful discussions. This work was supported by the National Aeronautics Space Agency Grants 80NSSC24K0344, 80NSSC23K1564, and 80NSSC19M0069.

## References

Barenholz U, Davidi D, Reznik E, et al. Design principles of autocatalytic cycles constrain enzyme kinetics and force low substrate saturation at flux branch points. *Elife* 2017;6: e20667.

Benner SA, Bell EA, Biondi E, et al. When did life likely emerge on Earth in an RNA-first process? *ChemSystemsChem* 2020;2(2):e1900035.

Bernier CR, Petrov AS, Kovacs NA, et al. Translation: The universal structural core of life. *Mol Biol Evol* 2018;35(8): 2065–2076.

Butler T, Goldenfeld N, Mathew D, et al. Extreme genetic code optimality from a molecular dynamics calculation of amino acid polar requirement. *Phys Rev E Stat Nonlin Soft Matter Phys* 2009;79(6 Pt 1):060901.

Conroy MJ, Andrews RM, Andrews S, et al. LIPID MAPS: Update to databases and tools for the lipidomics community. *Nucleic Acids Res* 2024;52(D1):D1677–D1682.

Crick FH. The origin of the genetic code. *J Mol Biol* 1968; 38(3):367–379.

Darwin C. (1959) *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London: John Murray.

Doolittle WF, Brunet TD. What is the tree of life? *PLoS Genet* 2016;12(4):e1005912.

Eschenmoser A. Towards a chemical etiology of nucleic acid structure. *Orig Life Evol Biosph* 1997;27(5–6):535–553.

Fournier GP, Andam CP, Alm EJ, et al. Molecular evolution of aminoacyl tRNA synthetase proteins in the early history of life. *Orig Life Evol Biosph* 2011;41(6):621–632.

Freeland SJ, Hurst LD. The genetic code is one in a million. *J Mol Evol* 1998;47(3):238–248.

Freeland SJ, Knight RD, Landweber LF, et al. Early fixation of an optimal genetic code. *Mol Biol Evol* 2000;17(4):511–518.

Frenkel-Pinter M, Bouza M, Fernández FM, et al. Thioesters provide a plausible prebiotic path to proto-peptides. *Nat Commun* 2022;13(1):2569–2568.

Frenkel-Pinter M, Rajaei V, Glass JB, et al. Water and life: The medium is the message. *J Mol Evol* 2021;89(1–2):2–11.

Furukawa Y, Sunami S, Takano Y, et al. Bio-essential sugars in samples from asteroid Bennu. *Nat Geosci* 2026;19(1):19–24.

Gagneux P, Panin V, Hennet T, et al. (2022) Evolution of glycan diversity. In Varki A, Cummings RD, Esko JD(eds.) *Essentials of Glycobiology*. (4th edition) Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, chapter 20.

Gaston MA, Jiang R, Krzycki JA. Functional context, biosynthesis, and genetic encoding of pyrrolysine. *Curr Opin Microbiol* 2011;14(3):342–349.

Genchi G. An overview on D-amino acids. *Amino Acids* 2017; 49(9):1521–1533.

Gladwell M. (2006) *The Tipping Point: How Little Things Can Make a Big Difference*. New York: Little, Brown.

Glavin DP, Dworkin JP, Alexander CMOD, et al. Abundant ammonia and nitrogen-rich soluble organic matter in samples from asteroid (101955) Bennu. *Nat Astron* 2025;9(2):199–210.

Goldenfeld N. Looking in the right direction: Carl Woese and evolutionary biology. *RNA Biol* 2014;11(3):248–253.

Gomez MAR, Ibba M. Aminoacyl-tRNA synthetases. *RNA* 2020;26(8):910–936.

Gomez-Mestre I, Pyron RA, Wiens JJ. Phylogenetic analyses reveal unexpected patterns in the evolution of reproductive modes in frogs. *Evolution* 2012;66(12):3687–3700.

Greenstein S. (2015) *How the Internet Became Commercial: Innovation, Privatization, and the Birth of a New Network*. Princeton, NJ: Princeton University Press.

Guo L-T, Amikura K, Jiang H-K, et al. Ancestral archaea expanded the genetic code with pyrrolysine. *J Biol Chem* 2022;298(11):102521.

Haas S, Sinclair KP, Catling DC. Biogeochemical explanations for the world's most phosphate-rich lake, an origin-of-life analog. *Commun Earth Environ* 2024;5(1):28.

Haig D, Hurst LD. A quantitative measure of error minimization in the genetic code. *J Mol Evol* 1991;33(5):412–417.

Harayama T, Riezman H. Understanding the diversity of membrane lipid composition. *Nat Rev Mol Cell Biol* 2018;19(5):281–296.

Harrison SA, Rammu H, Liu F, et al. Life as a guide to its own origins. *Annu Rev Ecol Evol Syst* 2023;54(1):327–350.

Hud NV, Cafferty BJ, Krishnamurthy R, et al. The origin of RNA and "my grandfather's axe". *Chemistry & Biology* 2013;20(4):466–474.

Hug LA, Baker BJ, Anantharaman K, et al. A new view of the tree of life. *Nat Microbiol* 2016;1(5):16048.

Jacob F. Evolution and tinkering. *Science* 1977;196(4295):1161–1166.

Könnyű B, Szathmáry E, Czárán T, et al. Kinetics and coexistence of autocatalytic reaction cycles. *Sci Rep* 2024;14(1):18441.

Koonin EV, Novozhilov AS. Origin and evolution of the genetic code: The universal enigma. *IUBMB Life* 2009;61(2):99–111.

Koonin EV, Novozhilov AS. Origin and evolution of the universal genetic code. *Annu Rev Genet* 2017;51(1):45–62.

Kyrpides N, Overbeek R, Ouzounis C. Universal protein families and the functional content of the last universal common ancestor. *J Mol Evol* 1999;49(4):413–423.

Lanier KA, Petrov AS, Williams LD. The central symbiosis of molecular biology: Molecules in mutualism. *J Mol Evol* 2017;85(1–2):8–13.

Maduzia LL, Moreau A, Poullet N, et al. The role of eIF1 in translation initiation codon selection in Caenorhabditis elegans. *Genetics* 2010;186(4):1187–1196.

Marcet-Houben M, Puigbo P, Romeu A, et al. Towards reconstructing a metabolic tree of life. *Bioinformation* 2007;2(4):135–144.

Margulis L, Sagan D. (2008) *Acquiring Genomes: A Theory of The Origin of Species*. New York: Basic Books.

Matange K, Marland E, Frenkel-Pinter M, et al. Biological polymers: Evolution, function, and significance. *Acc Chem Res* 2025;58(5):659–672.

Mayer-Bacon C, Freeland SJ. A broader context for understanding amino acid alphabet optimality. *J Theor Biol* 2021;520:110661.

Merkl R, Wiezer A. GO4genome: A prokaryotic phylogeny based on genome organization. *J Mol Evol* 2009;68(5):550–562.

Metcalfe B. Metcalfe's law after 40 years of ethernet. *Computer (Long Beach Calif)* 2013;46(12):26–31.

Metcalfe RM. Metcalfe's law: A network becomes more valuable as it reaches more users. *Infoworld* 1995;17(40):53–54.

Moody ERR, Álvarez-Carretero S, Mahendrarajah TA, et al. The nature of the last universal common ancestor and its impact on the early Earth system. *Nat Ecol Evol* 2024;8(9):1654–1666.

Nelson DL, Lehninger AL, Cox MM. (2021) *Lehninger Principles of Biochemistry*. (8th edition) Austin, TX: Macmillan Learning.

O'Donoghue P, Luthey-Schulten Z. On the evolution of structure in aminoacyl-tRNA synthetases. *Microbiology and Molecular Biology Reviews* 2003;67(4):550–573.

Petrov AS, Alvarez-Carrenoa C, Williams LD, et al. Coevolution of RNase P and the ribosome. *Proc Natl Acad Sci U S A* 2026;123(10):e2518495123.

Petrov AS, Bernier CR, Hsiao C, et al. Evolution of the ribosome at atomic resolution. *Proc Natl Acad Sci U S A* 2014;111(28):10251–10256.

Petrov AS, Gulen B, Norris AM, et al. History of the ribosome and the origin of translation. *Proc Natl Acad Sci U S A* 2015;112(50):15396–15401.

Philip GK, Freeland SJ. Did evolution select a nonrandom "alphabet" of amino acids? *Astrobiology* 2011;11(3):235–240.

Posey JE, Gherardini FC. Lack of a role for iron in the Lyme disease pathogen. *Science* 2000;288(5471):1651–1653.

Roche TP, Nedumpurath PJ, Karunakaran SC, et al. One-pot formation of pairing proto-RNA nucleotides and their supramolecular assemblies. *Life (Basel)* 2023;13(11):2200.

Sardans J, Penuelas J. Atomic ecology: Coupling atoms and ecology. *Sci Bull (Beijing)* 2024;69(1):11–16.

Schmitt E, Coureux P-D, Kazan R, et al. Recent advances in archaeal translation initiation. *Front Microbiol* 2020;11:584152.

Schmitt-Kopplin P, Gabelica Z, Gougeon RD, et al. High molecular diversity of extraterrestrial organic matter in Murchison meteorite revealed 40 years after its fall. *Proc Natl Acad Sci U S A* 2010;107(7):2763–2768.

Siefert JL, Fox GE. Phylogenetic mapping of bacterial morphology. *Microbiology (N Y)* 1998;144:2803–2808.

Siliakus MF, van der Oost J, Kengen SW. Adaptations of archaeal and bacterial membranes to variations in temperature, pH and pressure. *Extremophiles* 2017;21(4):651–670.

Soucy SM, Huang J, Gogarten JP. Horizontal gene transfer: Building the web of life. *Nat Rev Genet* 2015;16(8):472–482.

Theobald DL. A formal test of the theory of universal common ancestry. *Nature* 2010;465(7295):219–222.

Townes M. The spread of TCP/IP: How the Internet became the Internet. *Millennium* 2012;41(1):43–64.

Vetsigian K, Woese C, Goldenfeld N. Collective evolution and the genetic code. *Proc Natl Acad Sci U S A* 2006;103(28):10696–10701.

Wang J, Yu H. Threose nucleic acid as a primitive genetic polymer and a contemporary molecular tool. *Bioorg Chem* 2024;143:107049.

Weiss MC, Preiner M, Xavier JC, et al. The last universal common ancestor between ancient Earth chemistry and the onset of genetics. *PLoS Genet* 2018;14(8):e1007518.

Wiens J. The role of morphological data in phylogeny reconstruction. *Syst Biol* 2004;53(4):653–661.

Williamson MP. Autocatalytic selection as a driver for the origin of life. *Life (Basel)* 2024;14(5):590.
Willmore KE. The body plan concept and its centrality in evo-devo. *Evo Edu Outreach* 2012;5(2):219–230.
Woese CR. On the evolution of the genetic code. *Proc Natl Acad Sci U S A* 1965;54(6):1546–1552.
Woese CR. Interpreting the universal phylogenetic tree. *Proc Natl Acad Sci U S A* 2000;97(15):8392–8396.
Woese CR, Fox GE. The concept of cellular evolution. *J Mol Evol* 1977;10(1):1–6.
Woese CR, Olsen GJ, Ibba M, et al. Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. *Microbiol Mol Biol Rev* 2000;64(1):202–236.
Wogan N, Catling DC, Zahnle KJ. (2020) Molecules for the origin of life from impact-generated atmospheres on early Earth. AGU Fall Meeting Abstracts P053-0003.
Wolf YI, Aravind L, Grishin NV, et al. Evolution of aminoacyl-tRNA synthetases—analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. *Genome Res* 1999;9(8): 689–710.
Xavier JC, Hordijk W, Kauffman S, et al. Autocatalytic chemical networks at the origin of metabolism. *Proc Biol Sci* 2020; 287(1922):20192377.
Zerkle AL, House CH, Brantley SL. Biogeochemical signatures through time as inferred from whole microbial genomes. *Am J Sci* 2005;305(6–8):467–502.

Address correspondence to:
*Loren Dean Williams*
*Georgia Tech*
*School of Chemistry and Biochemistry*
*901 Atlantic Drive NW*
*Atlanta, Georgia*
*USA*

*E-mail:* ldw@gatech.edu

Submitted January 27, 2026
Accepted May 5, 2026
Associate Editor: Don A. Cowan