



The Origin(s) of Cell(s): Pre-Darwinian Evolution from FUCAs to LUCA

To Carl Woese (1928–2012), for his Conceptual Breakthrough of Cellular Evolution

Shiping Tang^{1,2}

Received: 12 September 2020 / Accepted: 29 May 2021 / Published online: 25 June 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

The coming of the Last Universal Cellular Ancestor (LUCA) was the singular watershed event in the making of the biotic world. If the coming of LUCA marked the crossing of the “Darwinian Threshold”, then pre-LUCA evolution must have been Pre-Darwinian and at least partly non-Darwinian. But how did Pre-Darwinian evolution before LUCA actually operate? I broaden our understanding of the central mechanism of biological evolution (i.e., variation-selection-inheritance) and then extend this broadened understanding to its natural starting point: the origin(s) of the First Universal Cellular Ancestors (FUCAs) before LUCA. My hypothesis centers upon vesicles’ making-and-remaking as variation and competition as selection. More specifically, I argue that vesicles’ acquisition and merger, via breaking-and-repacking, proto-endocytosis, proto-endosymbiosis, and other similar processes had been a central force of both variation and selection in the pre-Darwinian epoch. These new perspectives shed important new light upon the origin of FUCAs and their subsequent evolution into LUCA.

Keywords FUCAs · LUCA · Vesicle · Horizontal biomolecule transfer (HBMT) · Horizontal gene transfer (HGT) · Pre-Darwinian evolution

Abbreviations

HBMT Horizontal biomolecules transfer
HGT Horizontal gene transfer
HTAV Hydrothermal alkaline vent
TAGTF Terrestrial anoxic geothermal fields

Introduction

The coming of the Last Universal Cellular Ancestor (LUCA) was the singular watershed event in the making of the biotic world. If the coming of LUCA marked the crossing of the “Darwinian Threshold” (Woese 2002), then it follows that pre-LUCA evolution must have been Pre-Darwinian and hence at least partly non-Darwinian. But how did

Pre-Darwinian evolution before LUCA actually operate (cf. Tessera 2018)?

The central mechanism of biological evolution, variation-selection-inheritance, is one of the most universal mechanisms known. Much of our understanding of variation-selection-inheritance, however, has been dominated by the Neo-Darwinian Modern Synthesis with a rather narrow understanding of what constitutes variation, selection, and inheritance (or more precisely, persistence and retention, for the discussion here). This unduly narrow understanding may have been a key cause behind our failure to adequately explain some critical puzzles in biological evolution, especially the origin(s) of the first cell(s).

I thus broaden our understanding of variation-selection-inheritance, by bringing together but also critically extending insights from earlier and recent contributions.¹ I then

Handling editor: Aaron Goldman.

✉ Shiping Tang
twukong@fudan.edu.cn; twukong@yahoo.com

¹ Yunnan University, Kunming, China

² Fudan University, Shanghai, China

¹ E.g., Oparin 1953 [1938]; Fox 1973; Woese and Fox 1977; Gánti 1997; Szathmáry and Maynard Smith 1997; Norris and Raine 1998; Woese 1998, 2000, 2002, 2004; Cavalier-Smith 2001; Szostak et al. 2001; Luisi 2006, 2016; Chen and Walde 2010; Fry 2011; Pross 2011; Egel 2012, 2017; Bouchard 2014; Doolittle 2014; Koonin 2014a, b; O’Malley 2014; Baum 2015; Damer and Deamer 2015, 2020; Spitzer et al. 2015; Smith and Morowitz 2016; Pascal and Pross 2016; Cornish-Bowden and Cárdenas 2017; Garson 2017; Lanier and

extend this broadened understanding to its “natural” starting point: the origin of the First Universal Cellular Ancestors (FUCAs).²

Our broadened perspective of variation-selection-inheritance is NOT a rejection of Darwinian evolution. Rather, it is a necessary extension of it to the Pre-Darwinian epoch. Indeed, without broadening and then extending variation-selection-inheritance, it will be difficult to arrive at a reliable understanding about the origin of LUCA (Woese 1998, 2002; Vetsigian et al. 2006; cf. de Duve 2005b), because variation-selection-inheritance must have operated quite differently in the Pre-Darwinian epoch. As Woese (2002, p. 8472) put it: “We cannot expect to explain cellular evolution if we stay locked into the classical Darwinian mode of thinking.” Hence, in a similar spirit, “the present model [also] strives to release the fetters [that] classical Darwinian thinking imposes on the concept of cellular evolution.” (Woese 2002, p. 8746).

Taking the origin(s) of cell(s) as a jigsaw puzzle (e.g., Brack 1998; Higgs and Lehman 2015; Spitzer 2017), I shall focus on the areas in which consensuses have been lacking while assuming that areas with (near) consensuses are more or less settled.³ More concretely, I take the path from prebiotic chemical synthesis to LUCA has the following four areas with (near) consensuses despite differences on specifics.

First, prebiotic chemical synthesis of the building blocks of macro biomolecules (e.g., amino acids, nucleotides, lipids) must have preceded biological evolution, and it had been made possible by the environment of the Hadean Eon (Cantine and Fournier 2018).

Second, compartmentalization of biomolecules is essential because it restricts diffusion, increases stability, and allows for molecular crowding, which in turn facilitates interactions, reactions, and hence the coevolution of biomacromolecules (Segré et al. 2001; Spitzer and Poolman 2009, 2016; Ichihashi and Yomo 2014; Saha et al. 2014; Spitzer et al. 2015; Higgs 2016).

Third, following Woese (1998, 2002), progenotes (or FUCAs here) must have preceded genotes (or LUCA here) (see also Koonin 2014a, b).

Fourth, LUCA was already a fairly “modern” cell with a sophisticated membrane system, a (nearly) fully functioning metabolism system, a nearly complete translation apparatus (with the standard genetic code being a core part of it), and perhaps a RNA/DNA hybrid-centric replication system, among other things (Koonin 2014a, b; Fournier et al. 2015; Gogarten and Deamer 2016; Cornish-Bowden and Cárdenas 2017; For details, see also Section I below).

In contrast, we have little consensus regarding two critical questions: (1) how did FUCAs come to exist in the first place? (2) How did FUCAs evolve into LUCA exactly? It is on these two questions that I shall focus here.

The origin(s) of FUCAs and the coming of LUCA may well be one of those puzzles that we simply cannot stop thinking about but can never come up with a definitive answer. I do not pretend to provide a definitive answer either. Rather, my purpose is to show that taking a different starting point does allow us to integrate more existing data and evidence, resolve some key controversies, and point to more fruitful future inquiries.

A few clarifications on terms are now in order.

First of all, I focus on the origin(s) of cell rather than the origin(s) of life. It is to Carl Woese’s penetrating mind that we now accept not only that the origin(s) of cell is different from the origin(s) of life but also that the former must have been an equally, if not more, decisive step in the making of the biotic world (esp. Woese and Fox 1977; Woese 1998, 2002; Harold 2014; Koonin 2014a, b). Moreover, while both “what is life?” and “what is a (proto-)cell?” are hard questions, the latter is a bit easier because it is less philosophical (Luisi 2006, p. 17). As a mostly empirical task, our search for the origin(s) of cell is better off by staying away from potentially distracting philosophical muddles (e.g., Bedau and Cleland 2010).

Second, FUCAs here roughly corresponds to what Woese called “progenote” whereas LUCA to “genotes” (Woese 1998, 2002; see also Woese and Fox 1977; Woese 1982; Fox et al. 1982). Moreover, although LUCA has been conventionally taken to be the “Last Universal *Common* Ancestor”, it is now generally accepted that LUCA must have been a fairly complete cell (Koonin 2014a, b; Gogarten and Deamer 2016; Egel 2017; Spitzer 2017). In this sense, “the Universal *Cellular* Ancestor” is more proper than “the Universal *Common* Ancestor” because the former eliminates any ambiguity that LUCA must have been cellular. Also, by using FUCAs in plural whereas LUCA in singular, I convey the message that FUCAs had been a commune of different (proto-)

Footnote 1 (continued)

Williams 2017; Spitzer 2017; Toman and Flegr 2017; Doolittle and Inkpen 2018; Deamer 2019; Deamer et al. 2019; Lopez and Fiore 2019; Kunnev 2020. I apologize for not citing all the relevant literature: it is simply too voluminous. For overviews of our journey since Oparin, see Miller et al. 1997; Lazcano 2010; Deamer 2019.

² The notation of FUCA is similar to the First Eukaryotic Common Ancestors (FECAs) before the Last Eukaryotic Common Ancestor (LECA). Just as FECAs were most likely a pool of organisms (Eme et al. 2017), FUCAs were also most likely a pool of (proto-)organisms. See Section I below.

³ I also skip some rather minor quibbles (e.g., whether LUCA had introns).

cellular lineages whereas LUCA was more likely a single cell that came to produce all the organisms on this planet.⁴

Third, I use persistence for non-cellular entities but survival for (proto-)cellular entities. Similarly, following Luisi (2006, chap. 7, 2016, chap. 10), I use replication only for genetic replication but reproduction for vesicles and proto-cells that may grow and then divide, with or without genetic replication (see also Szathmáry and Maynard Smith 1997).

Pre-Darwinian Evolution Before LUCA: Principal Premises

If Woese (2002, p. 8746) was correct in insisting that we need to break out of the classical (Neo-)Darwinian framework to achieve a more adequate understanding of the origin(s) of cell(s), especially the path from FUCAs to LUCA, we may well need a new framework with radical departure points. This section advances such a framework, however tentatively.

Before going further, however, we need to differentiate FUCAs from LUCA: without such an explicit differentiation, some confusion is inevitable. For instance, Woese and Fox (1977) and Woese (1982, 1998, 2000, 2002) did not explicitly differentiate FUCAs and LUCA but instead grouped both under the “universal (communal) ancestor(s)”. Moreover, Woese insisted that all three domains (i.e., Bacteria, Archaea, and Eucarya) had diverged directly from the “universal (communal) ancestor(s)” (see also Egel 2017). Recent advances, however, have casted serious doubt on the three domains thesis (Williams et al. 2013; Koonin 2014a, pp. 197–199).

Similarly, Koonin (2009) earlier argued that LUCA was a pool of non-cellular virus-like genetic elements within inorganic compartments (see also Koonin and Martin 2005), and that the two primary domains (i.e., Bacteria and Archaea) then came to exist by independently escaping from these compartments as (proto-)cells. After grasping the tension between his “primordial virus world scenario” and the fact that LUCA might have been a (fairly complex) cellular entity with several hundred genes (and a fairly complete translation system) in place from his own earlier work (Koonin 2003; Wolf and Koonin 2007; see also Harris et al. 2003; Charlebois and Doolittle 2004; Ranea et al. 2006), Koonin (2014a, pp. 201–202) reconciled his two positions by rejecting Woese’s position and clearly differentiating progenotes (i.e., FUCAs) from LUCA (i.e., genote), insisting that whereas

progenotes might or might not have been cellular, LUCA as a genote most likely had been cellular.

Here, I explicitly differentiate “universal ancestors” into two phases: FUCAs and LUCA. FUCAs were proto-cells and what Woese and Fox (1977), Woese (1982, 1998, 2002) called “progenotes” or simply “universal (communal) ancestors”. In contrast, LUCA was what Woese (1998, 2002) called genote. More critically, LUCA was already a cell, and a fairly complex and perhaps (quasi-)modern one (Koonin 2014a, b). As becomes clear in section IV below, this differentiation allows us to clearly delineate two big pictures regarding the timing of the emergence of the two primary domains: did the two primary domains emerge after and from LUCA or did they emerge directly and independently from FUCAs?

I now outline five key premises. Building upon these premises but also drawing from evidence and perspectives scattered in the literature, we can then generate new and interesting hypotheses for understanding the origin(s) of FUCAs and LUCA (next section).

- (1) *Persistence as survival came long before genetic replication, certainly before division as reproduction (with or without genetic replication). Before replicators and reproducers, there must be survivors, to paraphrase Szathmáry and Maynard Smith (1997).* An entity, be it a compound, a complex, or a vesicle, has to exist and then persist within the (pre-)biotic system before it can become part of life, especially when it cannot metabolize or replicate. This law that persistence comes before (cellular or not) metabolism, replication, and reproduction holds most forcefully in the Pre-Darwinian epoch (Pascal and Pross 2016; Egel 2017; Toman and Flegr 2017). Also, persistence, metabolism, and replication were not coupled for much of the early prebiotic evolution: their coupling had been a product of Pre-Darwinian evolution.
- (2) *Before FUCAs, variation had been a primary means toward persistence.* From the very beginning of bioorganic evolution that eventually led to FUCAs and then LUCA, evolution was mostly about gaining more molecular and hence functional diversities so that protocells could survive in more diverse environment with a more potent arsenal (Pross 2013). Moreover, variations back then were not generated by genetic mutation (which did not exist for a long time) alone, but by two additional processes: (1) prebiotic chemical synthesis, polymerization, and stereochemical mutualism, both outside and inside of vesicles; (2) absorption via breaking-and-repacking, acquisition or engulfing via proto-endocytosis, merger or fusion via proto-endosymbiosis, and other similar processes.

⁴ Hence, my position critically differs from Prosdocimi et al. (2019), who also used the abbreviation FUCA for “the First Universal Common Ancestor”. Prosdocimi et al. insists that *FUCA is not a cell but a RNA-dependent peptide-synthesis machinery or peptidyl transferase center*. Also, they merely hypothesized that FUCA existed but provided no hypothesis for its origin.

Table 1 Four Pre-Darwinian selection mechanisms before FUCAs

Mechanisms	Objects being selected upon	Properties being selected
Most chemical	Molecules	Availability, stability, solubility, polymerization, and interaction
Chemical and physical	Bioorganic molecules and complexes	Interactions with other bioorganic molecules, forming complexes, and generating emergent physical and chemical properties (e.g., permeable membrane)
Vesicular	Vesicles	Absorption, acquisition, and merger
Proto-cellular	Protocells (before FUCAs)	Persistence (as survival), absorption, growth, and division, first without and then with primitive metabolism and genetic replication, and then tighter coupling of metabolism, growth, replication, and division

(3) *Natural selection can operate without genetic replication or even metabolism (at least not cellular metabolism), as long as different molecules, complexes, and vesicles have differential rate of persistence and reproduction within a system* (cf. de Duve 2005a; b, pp. 17–21, 154–55). Because Darwinian selection itself had been a product of the Pre-Darwinian epoch (Woese 2002; Koonin 2014a, b), an epoch of Pre-Darwinian and hence non-Darwinian selection must have operated during the Pre-Darwinian epoch and before the coming of Darwinian selection: the former had produced the latter. Four major non-Darwinian selection mechanisms, which most likely had appeared in the following order, had worked together in the process leading to FUCAs (see Table 1 for a summary).

- (a) *The first Pre-Darwinian selection mechanism is mostly chemical.* It operates upon molecules and selects not only their chemical properties as monomers but also their capacities for forming polymers and complexes. Here, key yardsticks of “fitness” include availability (i.e., steady supply from abiotic synthesis or meteoric bombardment), stability, solubility, polymerization, and stereochemical mutualism for forming larger complexes (Pross 2013; Pascal and Pross 2016; Lanier et al. 2017; Toman and Flegr 2017; Vitas and Dobovišek 2018).
- (b) *The second Pre-Darwinian selection mechanism is both chemical and physical.* It selects the different capacities of different bioorganic molecules and complexes to interact with each other, and in turn, whether their interactions confer new (or emergent) life-facilitating properties, structural and functional (Egel 2012; Pross 2013). Among possible interactions, two were perhaps most central: (1) amino acids, alpha-helix forming peptides, and (poly-)nucleotides that can not only interact with and stabilize vesicles but also make vesicles selectively permeable (e.g., Lear et al.

1988; Black et al. 2013; Cornell et al. 2019); (2) peptides and RNAs that can not only interact with each other but also lead to new or enhanced properties (e.g., more efficient and reliable in terms of synthesizing RNAs and peptides) via their interactions (e.g., Tagami et al. 2017; Frenkel-Pinter et al. 2020; Longo et al. 2020).

- (c) *The third Pre-Darwinian selection mechanism selects the different capacities of different vesicles (1) to absorb biomolecules and components via simple absorption and breaking-and-packing, (2) to engulf (or acquire) via proto-endocytosis and to merge (or fuse) via proto-endosymbiosis or similar processes.* Vesicles with superior capacities in both absorption and merger/acquisition will enjoy advantages over those with less effective capacities, in terms of persistence, variation, and evolvability (Oparin 1953; Fox 1973; Margulis 1981, 1991; Egel 2017). A cool-and-hot cycle and a wet-and-dry (or moisturization-and-evaporation) cycle might have driven both processes (Mansy and Szostak 2008; Damer and Deamer 2015, 2020; Higgs 2016).
- (d) *The fourth Pre-Darwinian selection mechanism operates upon vesicles that now approach protocells. Among these now fairly stable vesicles, those that can not only absorb, acquire, and merge but also can produce primitive metabolism and replication and then divide (or reproduce) will hold critical selection advantage over those that cannot.* Here, the key yardsticks of “fitness” included persistence (as survival), absorption, growth, and division, first without and then with primitive metabolism and genetic replication (e.g., Gánti 1997; Cavalier-Smith 2001; Szostak et al. 2001; Schrum et al. 2010; Luisi 2016, part IV; Takagi et al. 2020).
- (4) *Vesicles are evidently compartments of retention. More critically, however, vesicles’ absorption, acquisition,*

and fusion via breaking-and-repacking, proto-endocytosis, proto-endosymbiosis and other similar processes are both processes of variation and processes of selection. These processes of acquisition and merger had therefore been a central and powerful force in the pre-Darwinian evolution before LUCA, long before eukaryogenesis (e.g., Sagan 1967; Margulis 1981, 1991; O'Malley 2014; cf. de Duve 2005b, chap. 17).

- (a) Absorption, acquisition, and merger are processes of variation because they produce different compartmentalization and hence different crowding, combination, and coevolution of biomolecules within vesicles. Absorption, acquisition. And merger entail extensive “horizontal biomolecule transfer” (HBMT) rather than merely “horizontal gene transfer” (HGT): HBMT thus subsumes HGT. Indeed, only with HBMT, could have pre-Darwinian evolution drawn from “global inventions” (Woese 1998, 2002), and only with HBMT could have pre-Darwinian evolution overcome the seemingly insurmountable hurdle of bringing “the overwhelming amount of novelty needed to bring modern cells into existence” (Woese 2004, p. 182). HBMT was therefore the more pivotal and pervasive process than HGT, at least in the Pre-Darwinian epoch. In fact, Woese’s emphasizing of HGT during the evolution from FUCAs (i.e., progenotes) to LUCA (i.e., genote) in his millennial series is valid only if he meant HBMT with HGT (see also Vetsigian et al. 2006).⁵
- (b) Absorption, acquisition, and merger are also processes of selection because via these processes, some molecules will be retained and integrated within vesicles while some will be excluded from vesicles, and some vesicles will no longer exist.

- (5) *Reproduction and replication being tightly coupled had evolved from reproduction and replication not being linked at all and then to reproduction and replication*

⁵ Indeed, Vetsigian et al. (2006, p. 10,697) came extremely close to such a position: “The central conjecture in our model is that innovation-sharing, which involves horizontal transfer of genes and perhaps other complex elements among the evolving entities [*a dynamic far more rampant and pervasive than our current perception of horizontal gene transfer (HGT)*], is required to bring the evolving translation apparatus, its code, and by implication the cell itself to their current condition.” (Italics added.) Moreover, starting with a quite different perspective, Woese and Fox (1977, p. 5) had actually argued that “endosymbiosis should probably be considered an aboriginal (...) [and not-so-rare] trait” rather than “an acquired trait” or “a relatively recent and rare occurrence”. For this author, it is a mystery why Woese’s penetrating mind had not joined these two themes together.

being only loosely coupled. Only protocells with some kind of coupling of replication and reproduction via division became FUCAs, which then eventually evolved into LUCA. To paraphrase Szathmáry and Maynard Smith (1997) again, the tight coupling of reproduction and replication had also been a critical transition toward FUCAs as protocells.

Three criteria should be applied to any new thesis regarding the origin(s) of cell(s). First, it should provide a more coherent and consistent ordering of existing data and integrates new or previously underappreciated data with fewer ad hoc hypotheses than existing ones. Second, it should resolve more controversies than existing ones, again with fewer ad hoc hypotheses. Finally and perhaps most intuitively, a thesis about the origin(s) of cell(s) is the more credible one if it is less miraculous than other competing theses (Fry 2011; Lanier and Williams 2017).⁶

The next two sections aim to show that our new framework provides us with a foundation for deriving new and interesting hypotheses that are better supported by evidence.

Key Hypotheses

Building upon the premises above but also drawing from existing hypotheses and insights scattered in the existing literature, I now lay out key hypotheses regarding the Pre-Darwinian evolution before LUCA. Because I have differentiated FUCAs from LUCA earlier, I divide the hypotheses into two parts: before and after FUCAs. Figure 1 schematically summarizes the whole process from biopolymers to FUCAs and then LUCA.

Evolution before FUCAs

Abiotic synthesis of bioorganic molecules was the first step in the origin of life. Once bioorganic molecules came to exist, first as monomers (e.g., amino acids, nucleotides, fatty acids, and later on, phospholipids) and then as polymers (e.g., short peptides, small RNAs), they came under the force of natural selection even though replication did not operate back then. During this stage, there were two key selection yardsticks. The first is thermochemical stability or survivability within the system. The second is solvability and a minimum level of availability that allows a minimum level of concentration for monomers to be assembled into polymers and more complex hetero-biomolecules. Both stability and

⁶ For example, the thesis offered here is less miraculous than existing competing theses out there (e.g., multiple escapes and multiple invasions; Cavalier-Smith 2001; Forterre 2006; Lane et al. 2010).

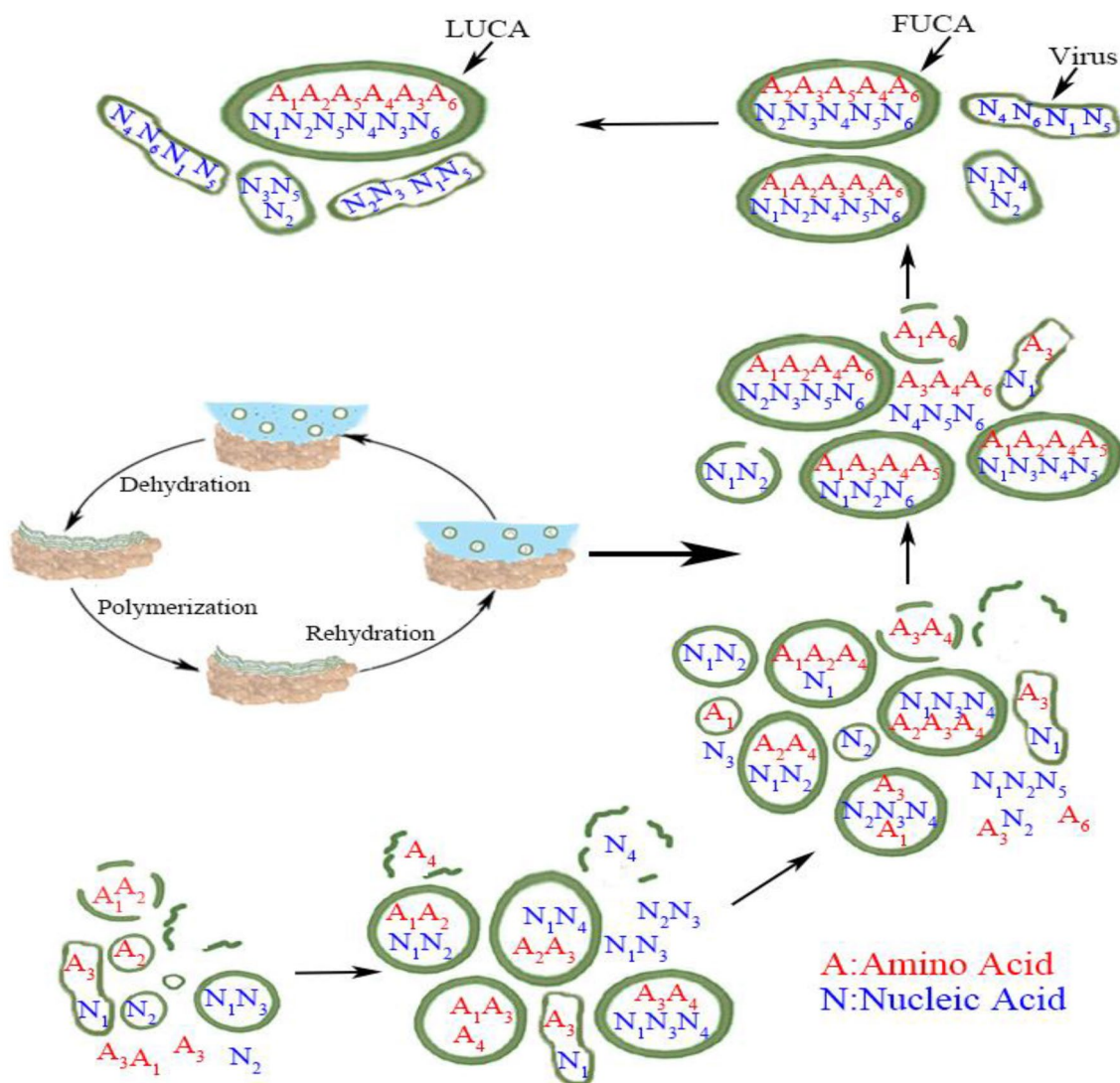


Fig. 1 From FUCAs to LUCA. Numbers in subscript denote different amino acids and nucleic acids. The exact matching between amino acid and nucleic acid within LUCA, in a metaphorical sense, implies that the standard genetic code had evolved most completely by the time of LUCA. The less than exact matching amino acid and nucleic acid within FUCAs and vesicles before FUCAs denotes the evolutionary path of the standard genetic code from a rudimentary form to a mature form in LUCA. Protocells or vesicles are in closed circles whereas

broken vesicles are in broken circles. Viruses are in elongated or other non-circular shapes. The three to one ratio of virus versus cell at the stage of LUCA is to imply the fact that virus may be the most abundant biological entity in the biosphere. A wet-and-dry cycle (on the left of the diagram) or a cool-and-hot cycle might have played a key role in driving the packing-breaking-and-repacking cycle and facilitating acquisition and merger of vesicles. The wet-and-dry cycle part within the figure is adapted from Damer and Deamer (2015) with permission

availability partly depend on the relative easiness of synthesis from simple precursors and protection from UV light.

For simple bioorganic molecules to be assembled into more complex hetero-biomolecules, they have to be stereochemically compatible with each other: in other words, there must be “molecular mutualism” (Lanier et al. 2017; Vitas and Dobovišek 2018). Key examples of such molecular mutualism include (1) that only some amino acids or peptides can interact with nucleotides and simple RNAs, non-covalently (e.g., Tagami et al. 2017; Frenkel-Pinter et al. 2020; Longo et al. 2020) and (2) that only some peptides

can form α -helices and then insert themselves into lipid membranes to make lipid membranes more permeable (e.g., Lear et al. 1988; Pohorille et al. 2003; Pohorille and Deamer 2009). Moreover, once some kind of molecular mutualism is fixed, it may become difficult to change or unravel. A key implication of molecular mutualism is that simplicity does not always mean better. Because only certain configurations are compatible with certain assembling strategies for bringing different molecules together, those molecules that can interact, bind, or fit with each other properly, rather than those that are merely simpler, may be selected.

Amphiphiles form vesicles in certain conditions. A vesicle succeeds in persisting in the system if it can retain its basic structure, float within a solution, absorb ingredients from its environment (e.g., via proto-endocytosis), and merge with other vesicles. Most likely, such vesicles also has the capacity of “dividing” without either strict reproduction or genetic replication. Rather, they divide via pinching or budding due to enlargement of size by (1) absorbing more lipids, peptides, (poly-)nucleotides, and other bioorganic molecules, (2) merging with and engulfing other vesicles [hence, each vesicle is also a target of (proto-) endosymbiosis by other vesicles], and (3) synthesizing new polymers within (Mansy et al. 2008; Zhu and Szostak 2009; Budin and Szostak 2011; Budin et al. 2012; Kurihara et al. 2015; Armstrong et al. 2018). During this stage, persistence and division have not been coupled with active reproduction, genetic replication or even sophisticated metabolism.

Even if RNA alone is capable of both replication and metabolism (Joyce 2002; Orgel 2004; Robertson and Joyce 2012; Horning and Joyce 2017), RNA might have come to interact with amino acids and peptides quite early on, and the primitive translation apparatus (and the genetic code) had originated from this interaction and then coevolution of amino acids/peptides with RNAs (Wong 1975, 1981; Wolf and Koonin 2007; Fox et al. 2012; Francis 2011, 2013, 2015; Sengupta and Higgs 2015; Koonin and Novozhilov 2017). During this stage of coevolution, precision in RNA replication (and proto-translation) is not necessarily an advantage (Vetsigian et al. 2006). Rather, during this stage of coevolution, the key was to make more RNAs and peptides without too much precision so that the structural diversity and hence the functional diversity of RNAs and peptides could increase more rapidly. With more diverse structures and functions, RNAs can then support the production of more diverse peptides with different properties, and these peptides in turn interact with RNAs more diversely to generate more emergent properties. This mutually reinforcing increase in structure and function of both RNAs and peptides laid a key foundation for the evolution of more complex ribonucleoprotein (RNP) world, the standard genetic code, and eventually a more versatile metabolism system.

For a period of time, the evolution of peptide-lipid membrane and the evolution of RNA-peptide (as the proto-translation machinery) might have proceeded independently from each other. The two processes might even have operated in different locations such as different terrestrial hydrothermal ponds or fields (Mulkiđjanian et al. 2009, 2012; Damer and Deamer 2015, 2020; Koonin 2014b, pp. 35–36). *Eventually, however, these two processes had to come together, and the moment in which these two processes merged was the first decisive step from replicators to reproducers that paved the way toward the first protocells or FUCAs* (Szathmáry and Maynard Smith 1997; Segré et al. 2001; Schrum et al. 2010; Higgs and Lehman 2015; Pressman et al. 2015; Luisi 2016;

Joyce and Szostak 2018). The fusing of the two processes was perhaps achieved by a peptide-lipid vesicle absorbing several RNA-peptide complexes (as proto-endocytosis), mediated via the interaction between RNA and lipids or peptide on the vesicle’s surface. Some of the products from this fusion became vesicles with (proto-)replicators inside, and some of these vesicles eventually became FUCAs.

Hence, FUCAs came to exist not only via packing-breaking-and-repacking of vesicles (e.g., Damer and Deamer 2015, 2020), but also via proto-endocytosis and proto-endosymbiosis by vesicles, thus drawing useful ingredients or components from “global inventions”. *Put it differently, FUCAs did not come to exist via de novo and in situ evolution within individual protocells alone: this will essentially mean that every FUCA must evolve almost entirely independently, and such a possibility would have been a miracle*. Rather, FUCAs came to exist via drawing and fusing innovations from many precursors. It is through HBMT that is underpinned by acquisition and merger rather than HGT alone that FUCAs eventually came to possess a proto-machinery of survival and a proto-machinery of replication within the same protocell (cf. Woese 1998, 2002; Vetsigian et al. 2006).

FUCAs continued to absorb RNAs, peptides, other biomolecules, vesicles, and possibly other fellow FUCAs from its external environment. As a result, coevolution of biomolecules within FUCAs accelerated (Segré et al. 2001; Wilson et al. 2014; Saha et al. 2018). Within a stable microenvironment provided by FUCAs’ membrane, other more fragile and elaborate proteins began to exist and operate, perhaps with the help of proto-chaperones (either RNA or peptide/protein). Within some FUCAs, metabolism eventually came to support the synthesis of more diverse and elaborate proteins and RNAs that could reinforce metabolism within a better regulated membrane system. Together, these two processes constitute a positive feedback loop. Eventually, within some FUCAs protocells, metabolism, survival, division, and replication of genetic material came to be gradually coupled with each other. The coupling was loose, however: FUCAs were protenotes that had yet to cross “the Darwinian Threshold”.

From FUCAs to LUCA

Once FUCAs came to possess both a proto-machinery of survival (roughly, metabolism supported by proteins within a membrane) and a proto-machinery of replication (now supported by both peptides/proteins and RNAs), survival and replication began to co-evolve with each other. Because both machineries require some kind of metabolism machinery (including bioenergetics), metabolism came to join survival and replication within the coevolution process (Takagi et al. 2020). This coevolutionary process laid the foundation for all subsequent evolutionary processes, and became only possible within protocells with a regulated membrane rather

than directly from the “naked” RNA world (Segré et al. 2001; Pohorille and Deamer 2009; Lombard et al. 2012; Koonin 2014a, b; Lopez and Fiore 2019). Fairly accurate replication of genetic materials (either RNA or DNA) supported by a replication apparatus could have only evolved within protocells such as FUCAs.

Different FUCAs not only competed against each other for various ingredients but also divided and survived differently within the system, as a form of pre-Darwinian selection process (Cheng and Luisi 2003; Chen et al. 2004; Wei and Pohorille 2011, 2013; Zhu et al. 2012; Adamala and Szostak 2013; Kurihara et al. 2015; Armstrong et al. 2018). Along the way, FUCAs continued to absorb useful ingredients and integrate them into more complex, versatile, and effective macromolecules, including more complex proteins and RNAs. During this phase, FUCAs might have also continued to absorb other (sub-)cellular components from other vesicles and integrate them into more tightly regulated cellular components. Hence, rampant extinction of (proto-)cellular lineages occurred during this period (Fournier et al. 2015). In this phase, a tight coupling of survival and replication might not have held any selective advantage. Indeed, the opposite might have been true: being more promiscuous and having more flexibility provides a protocell with significant advantage for survival (Szostak 2011; Koga 2014).

Within the commune of FUCAs, each FUCA protocell competed against each other. After a period during which survival, division, and replication co-evolved with each other, some of the FUCAs eventually became protocells in which division and replication are more tightly coupled and smoothly regulated. Eventually, a few lucky FUCAs with the right and tight coupling of metabolism, translation machinery, division with genetic replication, and energy efficiency will dominate the system, and these lucky few FUCAs merged into a single lineage or only one lineage of FUCAs survived: this lone surviving lineage became LUCA. *Because LUCA possessed a tight coupling of cell division with genetic replication, it was a genote that had crossed “the Darwinian Threshold”* (see Fig. 2 below). A mostly fully functioning translation system with the full standard genetic code had also “crystallized” in LUCA.

Due to the promiscuous origin of FUCAs and hence LUCA, LUCA most likely had been a “totipotent” cell that could have survived in quite different environments. More specifically, the lipid membrane of LUCA had been heterochiral rather than homochiral (Koga 2011, 2012, 2014; Lombard et al. 2012; cf. Coleman et al. 2019). Only later on when the original population of LUCA was split into two sub-populations,⁷ did they diverge into Bacteria and Archaea.

⁷ Alternatively, if there had been more than two sub-populations initially, eventually only two had survived.

By FUCAs, only some replicons had been selected to occupy the central genetic apparatus of the first cells whereas many were left out because only some replicons were helpful for the protocells whereas many others were not or even harmful.

Eventually, some of the left-out replicons came to persist as viruses: genetic parasites had been the inevitable result of a selection process (Koonin 2016). Some, if not most, positive single-stranded RNA viruses [(+)ssRNA viruses] might have originated from those RNA molecules and RNA-peptide complexes that were not integrated into FUCA’s proto-genome due to their disruptive properties (Koonin and Dolja 2013).⁸

Once genetic parasites came to exist, they began to play a critical role in driving the evolution of their hosts and themselves (Koonin and Dolja 2013; Koonin 2016).⁹ Without a (primitive) defense system against genetic invasion by mobile genetic elements, primitive cells would be extremely vulnerable to genetic invasion. In contrast, acquiring an even primitive defense system against genetic invasion inevitably reduces the rate of HGT.

The arms race between HGT and defense against genetic invasions by genetic parasites via HGT therefore most likely began only with FUCAs the earliest. The acquiring of a defense system against genetic invasion also marks the coming of HGT as a reduced form of HBMT. Thus, only in FUCAs did the narrower HGT replace the broader HBMT as the more critical force in driving evolution, although HBMT continued to operate, most dramatically in eukaryogenesis (Sagan 1967; Margulis 1981). Moreover, only in FUCAs did HGT gradually become more harmful (Jain et al. 1999). In contrast, because protocells before FUCAs accomplished HBMT via absorption and merger/acquisition, they were less likely to have evolved a defense system against HBMT early on.

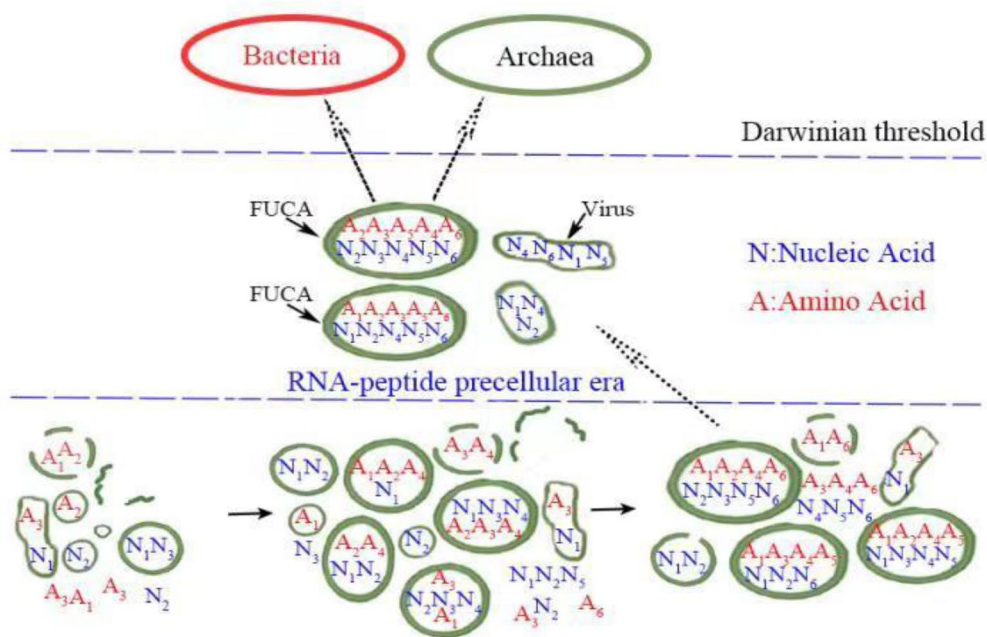
LUCA already possessed a quite sophisticated defense system against mobile genetic elements, which eventually evolved into the core defense systems against mobile genetic elements in the two primary domains, including the pAgo system, the CRISPR-Cas system, and the toxin-antitoxin (TA) system (Koonin 2017; Koonin and Makarova 2017; Koonin et al. 2020).

Due to some kind of geological ruptures or accidents (e.g., overflowing of a hydrothermal field or a “warm little pond”), the original LUCA population was split into two

⁹ Unfortunately, because viruses and other replicons tend to evolve rather rapidly, their evolution may inform us little about the evolution of FUCAs and LUCA (Moreira et al. 2009; cf. Koonin and Dolja 2013).

⁸ This possibility best explains the great diversity of this group of virus, and why (+)ssRNA viruses remain the largest group of viruses (de Farias et al. 2017).

Thesis 1: The two primary domains had escaped independently from the communal ancestors or FUCAs (Woese 1998; 2002; Koonin 2009; see also Koonin and Martin 2005).



Thesis 2: The two primary domains had bifurcated from LUCA, which had evolved from FUCAs as proto-cells.

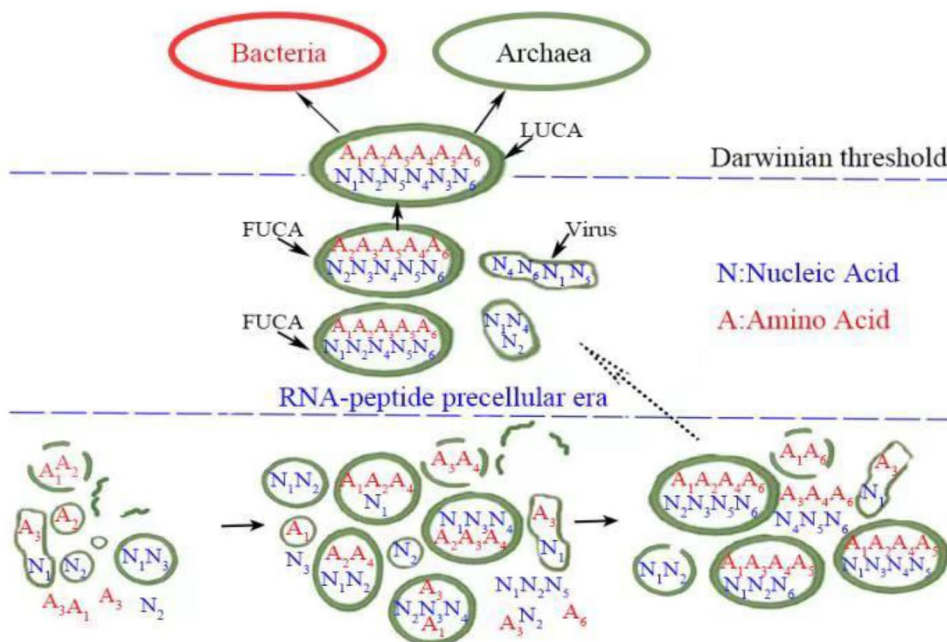


Fig. 2 Two Theses from the RNA-peptide world to the two primary domains

sub-populations in two different niches. This event started the making of Bacteria and Archaea as the two primary domains, including their differences in lipid membrane (i.e., from heterochiral to homochiral), DNA to RNA transcription, and DNA replication, etc. (Koga 2011, 2012, 2014; Lombard et al. 2012; López-García and Moreira 2015; Williams et al. 2013; Spang et al. 2015; Dacks et al. 2016; Lombard 2016; Eme et al. 2017).

Most likely, the genomes of FUCAs contained only short RNA molecules rather than a single long RNA molecule. The transition from RNA to DNA via a RNA–DNA hybrid system might have begun between FUCAs and LUCA or only after the coming of LUCA. LUCA, however, was split into two sub-populations before the transition from a RNA–DNA hybrid system to a DNA-only system had been complete. As a result, a DNA-only replication machinery had evolved twice independently, in the two primary domains after their divergence (Leipe et al. 1999). The whole transition from a RNA-only to a DNA-only via a RNA–DNA hybrid system therefore had gone through two phases: a phase of RNA–DNA hybrid and a phase of full transition to a DNA-only system. The former lasted from FUCAs to LUCA and beyond, even though some components of a primitive DNA replication machinery might have been already in place within FUCAs and before LUCA. The latter started after the divergence of LUCA into Bacteria and Archaea and continued well after.

Evidence: Chemical and Biological-Functional

Because FUCAs were a community of protocells and the evolutionary process after FUCAs with rampant HBMT had erased or at least obscured most of the early genetic footprints (Woese 2002; Fournier et al. 2015), empirical support for any theory regarding the origin(s) of cell(s) must be mostly chemical and biological-functional rather than genetic.¹⁰ This section presents mostly chemical and biological-functional evidence that supports our core hypotheses, in addition to some genetic evidence. The next section highlights that our theory resolves some key controversies regarding the origin(s) of cell(s), thus providing another source of support.

- (1) The universality of the standard genetic code, as part of the core translation apparatus, is indisputable. The uni-

versality of the standard genetic code is best explained by the coming of amino acid/peptide-RNA interaction (and then the coevolution of amino acid/peptide-RNA interaction and the proto-translation system) very early on, long before the coming of DNA replication and DNA to RNA transcription. The possibility that the standard genetic code came to exist via initial chemical mutualism between amino acid/peptide with RNA and then the coevolution of peptide/protein and RNA is now generally accepted (Wong 1975, 1981; Wolf and Koonin 2007; Yarus et al. 2009; Yarus 2017; Francis 2011, 2013, 2015; Petrov et al. 2015; Sengupta and Higgs 2015; Kovacs et al. 2017; Saad 2018; Kim et al. 2019).¹¹ This possibility is reinforced by the fact that small and simple peptides derived from the core part of ribosome proteins can enhance the catalytic activities of RNA polymerase ribozyme, with the smallest size being merely ten amino acids (Tagami et al. 2017; Frenkel-Pinter et al. 2020).¹² The fact that many ancient proteins and protein domains are connected to nucleotides and RNAs, perhaps initially as peptide-RNA cofactors, also suggests an early rather than a late RNA-peptide world (White 1976; Szathmáry and Maynard Smith 1997, pp. 563–568; Ma et al. 2008; Goldman and Kacar 2021).

- (2) Apparently, the translation apparatus is the most universal among the three parts in the information process system of modern organisms (i.e., replication of DNA, transcription from DNA to RNA, and translation from mRNA to protein). This fact suggests that a fairly complete translation system (most likely including the whole standard genetic code, a proto-ribosome, mRNA, tRNA, and most aminoacyl-tRNA synthetases) must have evolved by LUCA (Wolf and Koonin 2007; see also Wolf et al. 1999; Woese 2000). This fact strongly points to not only the possibility that vesicles with per-

¹² The fact that replicons with protein capsids (i.e., viruses) are more widespread than replicons without protein capsids (i.e., viroids) also suggest that the former is more advantageous than the latter. Also, viroids have so far only been found in higher plants whereas viruses have been found in all three domains. This fact suggests that viroids most likely have a recent origin.

¹¹ The evolution of the standard (or universal) genetic code is a research area all by itself. Here, suffice to say that our thesis is compatible with the emerging consensus that three mechanisms must have jointly played key role in the evolution of the code and they reinforce each other (see Koonin and Novozhilov 2017 for review). The three key mechanisms are: stereochemistry, coevolution of the code and then translation plus proteins, and error minimization. Note, however, all three mechanisms implicitly center on the evolution of the code within a single cell population and this possibility is highly unlikely (see below). Hence, HBMT as a mechanism that can draw from “global innovation” [and subsumes HGT (e.g., Vetsigian et al. 2006)] may be absolutely necessary.

¹⁰ By biological-functional, we mean that the same biological function can be performed by different molecules, with or without significant structural or sequence similarities. For example, an alpha-helix can be formed by different stretches of peptide.

meable membrane came to exist quite early on but also the possibility that these vesicles were able to merge with and acquire each other via proto-endosymbiosis and proto-endocytosis.

- (a) *Simply put, it would have been extremely difficult for the translation apparatus to evolve out of a single protocell or vesicle with HGT alone.* Woese (1982, p. 12) put it forcefully very early on: “the translation apparatus is too large, too complex to have arisen in a single evolutionary process.” (See also Woese and Fox 1977, pp. 2–3; Fox 2010; Woese 1998, 2001, 2002; Vetsigian et al. 2006; Koonin 2014a.)
- (b) A majority of the one hundred or so universally conserved genes were within the translation apparatus (Koonin 2003; Ranea et al. 2006; Puigbò et al. 2009). This fact suggests not only the possibility that a RNA-amino acid/peptide world came to exist very early on but also the possibility that HBMT via vesicles’ merger-and-acquisition had been a key force behind the emergence of a fairly complex translation apparatus in LUCA. The evolution of the complex translation apparatus requires input from “global inventions” that can only be provided by HBMT via vesicles’ merger-and-acquisition. With HBMT, vesicles within a pool of vesicles or protocells could have easily drawn from “global inventions”, not only in genetic materials but also in metabolism and other ingredients, within the community of FUCAs as progenotes (Pohorille and Deamer 2009; Mulki-djanian et al. 2009; Lombard et al. 2012; Koonin 2014a, b). Obviously, HBMT assumes vesicles with proto-membrane that can retain molecules within whereas HGT does not.
- (c) RNAs and amphiphiles not only can interact with each other, their interaction also confers new properties to each other (e.g., Black et al. 2013). To begin with, lipids can assist polymerization of nucleotides into RNA-like molecules in simulated terrestrial geothermal environment (Olasagasti and Rajamani 2019). Very importantly, RNAs (as replicators) encapsulated by vesicles tend to be more stable (Saha et al. 2018; Shah et al. 2019). Finally, positive-strand RNA viruses can readily work with membranes during replication and infection by “hijacking” or recruiting membrane lipids (Miller and Krijnse-Locker 2008; van der Schaar et al. 2016; Altan-Bonnet 2017). All these facts suggest that the fusion of RNA-peptide and the lipid-peptide vesicle is quite likely.
- (3) The fact that Bacteria and Archaea share membrane-associated proteins that are key components of bioenergetics, respiratory chains, secretion/membrane-targeting, protein export and membrane-insertion pathways strongly points to the possibility that vesicles with permeable membrane came to exist early rather than late (Lombard et al. 2012). In fact, a recent study by Harris and Goldman (2021) has shown that a system of signal recognition particle (SRP) and a Sec translocation channel were already in place by the time of LUCA. In particular, the ancestor of Fth and FtsY proteins as the two principal components of the SRP system most likely existed before LUCA, or late FUCAs here, as suggested by the fact that late amino acids have been quite rare within both Fth and FtsY proteins. Even more remarkably, both Fth and FtsY proteins have been extremely conserved across the three primary domains, sequence-wise and structurally. Together, these facts further points to the possibility that vesicles were able to merge with and acquire each other via proto-endosymbiosis and proto-endocytosis. Again, it would have been extremely difficult for a single protocell or vesicle to evolve all these complex and not always interlinked machineries, within and in situ, even with HGT. *Rather, they require inputs from “global inventions” via HBMT brought about by merge and acquisition.*
- (4) LUCA most likely had been a quite complex organism with several hundred genes and hence a “totipotent” cell that is can survive with quite different environments (Koonin 2003; Harris et al. 2003; Charlebois and Doolittle 2004; Ranea et al. 2006; El Baidouri et al. 2020; Krupovic et al. 2020; cf. Weiss et al. 2016). This fact is more consistent with the possibility that LUCA came from many FUCAs via acquisition and merger than with the possibility that LUCA came to evolve from a single FUCA lineage all by itself or by HGT alone. In other words, FUCAs and then LUCA had a promiscuous origin. In fact, Woese’s (1998, 2002) notion of “global invention” is just another term for a promiscuous origin of FUCAs and then LUCA.
- (5) The possibility that HBMT via proto-endosymbiosis and proto-endocytosis has been a key mechanism in the evolution of FUCAs and LUCA is supported by additional evidence.
- (a) Many artificial vesicles indeed can grow and divide without replication, by simply absorbing either ingredients or other (mini-)vesicles (Kurihara et al. 2015; Saha and Chen 2015; see also Pohorille and Deamer 2009). They can also undergo structural changes under different conditions (e.g., different pH, different concentration, a wet-and-dry cycle, a hot-and-cool cycle,

- redox chemistry), thus facilitating acquisition and merger (e.g., Cheng and Luisi 2003; Chen et al. 2004; Chen and Walde 2010; Zhu and Szostak 2009; Zhu et al. 2012; Budin and Szostak 2011; Budin et al. 2012; Damer and Deamer 2015; Qiao et al. 2017; see also Oparin 1953 on coacervates).
- (b) In natural settings, “terrestrial anoxic geothermal fields (TAGTFs)” or Darwin’s “warm little pond(s)” can drive “wet-and-dry” and other cycles, which can in turn drive vesicles through the cycle of breaking up old vesicles and then reforming new vesicles by changing the concentration of ions and other ingredients or the overall physical microenvironment within a “warm little pond” (Mulkiđjanian et al. 2012; Damer and Deamer 2015, 2020). This process of packing, breaking, and then repacking of biomolecules partly overcomes the physical barrier against HBMT imposed by (partially impermeable) vesicles and essentially works as a physical and chemical process of recombination and reconfiguration. The fact that quite a few paths (e.g., pH, concentration, wet-and-dry, hot-and-cool, or even redox) can propel this process of vesicular recombination and that different vesicles containing different biomolecules have different capacities of growth (via absorption and in-taking) and division under different conditions strongly suggests that such pathways might have been powerful forces of variation and selection in the evolution of FUCAs.¹³
- (c) The notion that mitochondria and chloroplast and other plastid-like structures (e.g., hydrogenosomes, and mitosomes) had evolved as relics of endosymbiosis is now widely accepted (Sagan 1967; Margulis 1981, 1991; Zimorski et al. 2014; Ku et al. 2015; *J. Theo. Biol.*, special issue 2017). This fact suggests that (proto-)endosymbiosis and (proto-)endocytosis might have been quite ancient and frequent in pre-cellular evolution, as Woese and Fox (1977, p. 5) had argued long ago.¹⁴
- (d) In situ and de novo evolution, aided by HGT alone, would not have been a viable route for the making of a “totipotent” LUCA because HGT via replicons could not have brought together a large genome that can sustain a complex life as LUCA or even FUCAs. With the exception of megaviruses,¹⁵ most viruses and other replicons (e.g., plasmids) have a small genome. Although one can argue that these replicons came to their small genome size via loss of genes, the possibility that early replicons really were just small genetic fragments is far more plausible (Leipe et al. 1999). Thus, Woese’s (1998, 2000, 2002, 2004; see also Vetsigian et al. 2006) vision that HGT was a critical force in the evolution of FUCAs to LUCA does not have a viable mechanism other than HBMT via vesicles’ acquisition and merger. For LUCA to possess several hundred genes made of short RNA molecules, HBMT via vesicles’ acquisition and merger might have been the only viable option. Thus, once we replace HGT with HBMT, which subsumes the former, we can resolve most, if not all, of the contradictions and inconsistencies in Woese’s (1998; 2002) treatises on the origin of the cell (see also Koonin 2014a). Indeed, Woese (2002, 8744) came close to admit this possibility: “*Were that organization (of a progenote as a protocell) simple and modular enough, all of the componentry of a cell could potentially be horizontally displaced over time.*” Quite evidently, such a possibility is entirely compatible with HBMT as the central mechanism leading to the coming of FUCAs and LUCA, but not so compatible with HGT alone.
- (6) Because a protocell is a survival machine first and a reproduction machine second, molecules that play essential roles in the survival of a protocell (as an organism) must have also originated early on. Besides the bilayered lipid membrane itself as a protective apparatus, some other protective mechanisms must have also evolved very early on. Intuitively, any organism or protocell with some kind of machinery that can cope with stressful environmental changes should possess significant advantages over those without. Hence, FUCAs and LUCA must possess a stress response system quite early on.
- (a) Such a machinery is most likely in the form a stress response mechanism and apparatus, or “heat shock response (HSR)” as known today.

¹³ This possibility therefore supports the “TAGTF” thesis while going against the “hydrothermal alkaline vents (HTAV)” thesis. See below.

¹⁴ In this sense, eukaryogenesis was a unique event only because it produced a eukaryotic cell but not because endosymbiosis had only operated this once (Woese and Fox 1977; Booth and Doolittle 2015; cf. Sagan 1967; Margulis 1981; Lane and Martin 2015).

¹⁵ Megaviruses, as double-stranded DNA viruses, must have originated after the origin of DNA replication. Hence, megaviruses do not support the thesis that early replicons or protocells had possessed big genomes.

Very likely, HSR was an early invention as requisites for life on the edge, for the sake of survival. Major components within HSR, such as Hsp100, Hsp90, Hsp70, Hsp60, and other small heat shock proteins (HSPs), are highly conserved across the three domains (Richter et al. 2010). Therefore, HSPs might have been some of the earliest proteins to be firmly integrated into FUCAs and then retained by LUCA, and this possibility best explain why so many HSPs are highly conserved across all three domains.

- (b) Membranes formed by amphiphiles alone are often too impermeable to ions and other bioorganic molecules. One potential solution for this challenge is to insert peptides that can form α -helices into the membranes. Also, lipid membranes are stabilized and regulated by peptides and this mutualism between membrane lipids and peptides or proteins is a key part of the foundation of survival. Hence, a critical aspect of the coevolution of membrane and membrane peptides/proteins was about protecting the membrane (Pohorille et al. 2003; Mulikidjanian et al. 2009; Lombard et al. 2012).
- (c) Several observations suggest that a small heat shock protein, Hsp12, may be a molecular fossil of such a machinery of coping with stress (Richter et al. 2010). In its soluble form, yeast Hsp12 is unfolded or unstructured). Yet, when Hsp12 interacts with lipids, it becomes α -helical. Also, Hsp12's four α -helices are critical for its proper function. Hsp12, when binding membranes, stabilizes membranes by decreasing membrane fluidity. Remarkably, 50% of Hsp12's sequence is made of five amino acids: Ala, Asp, Glu, Gly, and Lys. Among the five amino acids, four of them (i.e., Ala, Asp, Glu, and Gly) belong to the “first amino acids” (Francis 2013; Sengupta and Higgs 2015). Both Ala and Gly tend to form transmembrane α -helices via Ala/Gly-X-X-X-Ala/Gly whereas Asp can stabilize α -helices by capping α -helix TM domain, and Ala/Gly-X-X-X-Ala/Gly sequences also facilitate dimerization (Francis 2013).¹⁶
- (d) Somewhat more speculatively, many bacterial small proteins (fifth amino acids or shorter) have been founded to interact with cell membrane but lack enzymatic capacities (reviewed by Storz et al. 2014). If genes for small proteins are reservoirs for

genes for larger proteins (Carvunis et al. 2012), some small peptides might have been selected for their capacities of interacting with and stabilizing proto-membranes rather than their enzymatic capacities.

Controversies Resolved, Partly and Possibly

If our thesis is valid, then it should help resolve some key controversies about the origins of cell(s) (for earlier reviews, see Woese 1998, 2002; Doolittle 2000; Fry 2011; Egel 2012; 2017; Koonin 2014a, b; Spitzer 2017; Cantine and Fournier 2018; Damer and Deamer 2020). This section addresses several key controversies that our theory may help resolve, while also insisting that some debates may never be confidently resolved and hence less useful.

The RNA-Peptide World Versus the Vesicle-Peptide World

The question which comes first, the RNA world (and then a RNA-peptide world) or the peptide-lipid membrane (with or without some rudimentary metabolism), is one of those unfruitful controversies because it can never be firmly resolved (Egel 2009; Fry 2011). Yet, regardless whether these two worlds or pathways had originated simultaneously or sequentially, vesicles as compartmentalized minispaces must have come to exist early on and these vesicles had made many subsequent evolutionary processes possible (Pohorille and Deamer 2009; Lombard et al. 2012; Koonin 2014a; Cantine and Fournier 2018). *Most critically, the merging of these two worlds or pathways must have been the more decisive step because it made possible the eventual coupling of survival, metabolism, and replication.* Indeed, in experimental studies, we are now witnessing an integration of the RNA world and the membrane/vesicle world (Pressman et al. 2015; Joyce and Szostak 2018).

The thesis put forward here bridges and overcomes key difficulties within hypotheses centered upon replication/replicator alone and hypotheses centered upon peptide plus vesicle/membrane alone. The key difficulty for hypotheses centered upon replication/replicator alone is how a replicator can become a reproducer. In contrast, the key difficulty for hypotheses centered upon peptide plus vesicle/membrane alone is how a not-so-evolvable reproducer (i.e., vesicles that can grow and divide) without replicators inside can become an evolvable reproducer with coupled replication and reproduction (Thomas and Rana 2007). Our thesis overcomes the

¹⁶ So far, no Hsp12 homologs have been discovered beyond fungi, but this may be due to the fact that membrane proteins are less conserved than water-soluble proteins (Sojo et al. 2016). More speculatively, polyglycine, made of Glycine, might have played a similar role in stabilizing protocell membranes (Kim et al. 2019).

two challenges by suggesting that replicator and vesicles (as weak reproducers) came to evolve together via a continuous process of merge and acquisition until LUCA.

Our thesis also resolves a key self-contradiction within Woese (1998, 2002) and Koonin (2009, 2014a, b; see also Koonin and Martin 2005). Because (complex) membrane proteins can only come to exist with some form of proto-membrane and yet complex membrane protein can only come after a fairly complex translation system, there is a classic chicken-and-egg difficulty within the evolution of (proto-)cell membrane and membrane proteins (Koonin 2014b, pp. 32–35).¹⁷ Our thesis resolves this difficulty by maintaining that acquisition and merger of vesicles (as precursors to protocells) and then protocells (e.g., FUCAs) has played a central role in overcoming this seemingly intractable difficulty.

Gain-of-Structure/Function versus Loss-of-Structure/Function before LUCA

During the process leading to FUCAs and LUCA, evolution had been mostly about gain in structure and function so that FUCAs and LUCA could survive in diverse environments with a more potent arsenal. Within this period, metabolic and synthetic innovation was equally critical, if not more critical than genetic ones (Ranea et al. 2006). More likely than not, any significant genetic and functional streamlining or simplification came after LUCA, when “evolutionary temperature” had cooled down considerably (Woese 1998). As de Duve (2005b, p. 163, fn. 3) put it pithily, “There can be no reduction without prior ‘complexification.’”

Two additional facts suggest that any significant streamlining must have come after LUCA. First, the standard genetic code had evolved in two phases, a phase with early amino acids and a phase with late amino acids (Wong 1975, 1981; Francis 2011, 2013, 2015; Sengupta and Higgs 2015; Koonin and Novozhilov 2017). *Moreover, both phases must have been completed before LUCA.* Second, LUCA might have possessed only about one hundred protein domains that performed multiple functions (Ranea et al. 2006).

The possibility that streamlining via loss of functions and genes came after LUCA goes against the thesis that LUCA was a proto-eukaryotic cell and the thesis that Bacteria, Archaea, and Eukarya came to exist via multiple genetic reductions or escapes (e.g., Glansdorff 2002; Glansdorff et al. 2008; Egel 2017; see also Forterre and Philippe 1999;

Philippe and Forterre 1999; Woese 1998, 2002). Such a thesis does not explain how the complexity came to exist in the first place (De Duve 2005b, chaps. 13–15). Also, recent advances strongly suggest a Tree of Life with Bacteria and Archaea as the two primary domains, and Eukarya arose from Archaea (Lombard et al. 2012; Williams et al. 2013; Koonin and Yutin 2014; López-García and Moreira 2015; Spang et al. 2015; Dacks et al. 2016; Eme et al. 2017). By all likelihood, the hypothesis that LUCA was a proto-eukaryotic cell can now be ruled out (cf. Staley and Fuerst 2017).

From Precellular to (Proto-)Cellular: Two Theses

Extending the arguments above, we can now resolve two possible theses regarding the evolution from the precellular era to the (proto-)cellular era as depicted in Fig. 2 (see also Koonin 2014b, p. 31). The first thesis (upper half of Fig. 2) has been advanced by Woese and Fox (1977), Woese et al. (1978), Woese et al. (1978), Woese (1982, 1998, 2000, 2002), Kandler (1994a, b, 1998; Koonin and Martin (2005), Koonin (2009, 2014a, b), and Egel (2017) in slightly different forms. This thesis holds that the two primary cellular lineages or domains (i.e., Bacteria and Archaea) had evolved directly from a commune of non-cellular entities (perhaps virus-like) and that the Darwinian Threshold was crossed only when the two primary domains emerged or escaped from the commune state. This thesis thus does not differentiate FUCAs and LUCA. Instead, they are grouped under the same heading, the universal common ancestor, which is communal and non-cellular.

The second thesis (lower half of Fig. 2), as explicitly put forward here, contends that the two primary domains had bifurcated from a single LUCA population (see also Williams et al. 2013; Koonin 2014b; Gogarten and Deamer 2016; Cornish-Bowden and Cárdenas 2017). Moreover, LUCA (the genote) had evolved from a commune of FUCAs (or progenotes), and the Darwinian Threshold was already crossed by LUCA. Furthermore, both FUCAs and LUCA were already (proto-)cellular: the key difference between them has been that the former had only loose whereas the latter had tight coupling of metabolism, replication, translation, and reproduction (i.e., division). In Woese’s terms, LUCA was almost fully “crystallized” whereas FUCAs had yet to crystallize.¹⁸

Our thesis explicitly rejects the thesis that Bacteria, Archaea, and eventually Eukarya had emerged independently and sequentially by breaking free or escaping from a pool of either prokaryotic or even proto-eukaryotic precells

¹⁷ To resolve this classic chicken-and-egg difficulty, Koonin wisely switched from the thesis of maritime origin within inorganic chambers (Koonin and Martin 2005; Koonin 2009) to a thesis of terrestrial origin (Mulkiđjanian et al. 2009, 2012; see also Koonin 2014a, b). Yet, Koonin (2014b) has retained the possibility of multiple escape from the communal ancestors (see also Koonin and Dolja 2013). See below.

¹⁸ Hence, by 1998, Woese had changed from his original conception of progenotes that did not differentiate progenotes and genotes to a more valid conception that does differentiate the two.

(e.g., Woese 1982; Kandler 1994a, b, 1998; Wächtershäuser 2003; Egel 2017). In short, the precells-then-escape hypothesis not only goes against the two primary domains thesis but too implicitly banks upon miracles.

The Origin(s) of LUCA: Terrestrial & Heterotrophic vs. Maritime & Autotrophic

Due to the imperatives of survival, vesicles or protocells could not afford to be too picky: some degree of heterogeneity (via promiscuity) might have been not only helpful but actually indispensable (Mansy and Szostak 2008; Szostak 2011). The notion that FUCAs came together through HBMT based on proto-endosymbiosis and proto-endocytosis that bring together different vesicles containing different components suggests a heterotrophic origin of FUCAs and LUCA. Autotrophic life was only achieved after a long period of heterotrophic evolution (Oparin 1953 [1938]; Fry 2011; Damer and Deamer 2015, 2020; Egel 2017).

The leading autotrophic hypothesis that LUCA had evolved a nearly complete arsenal within the chambers of HTAV and then escaped as a fully autotrophic precell has numerous difficulties (Mulkiđjanian et al. 2012; Gogarten and Deamer 2016; Jackson 2016, 2017; cf. Martin and Russell 2003, 2007; Koonin and Martin 2005; Martin et al. 2008; Lane et al. 2010; Weiss et al. 2016; Egel 2017). *Indeed, by insisting that all the good things for LUCA must have evolved in a single location, the HTAV hypothesis approaches a theory banking on miracles.*

Our theory is more consistent with the possibility that FUCAs came to exist in “terrestrial anoxic geothermal fields (TAGTFs)” or Darwin’s “warm little pond” rather than in HTAVs (Mulkiđjanian et al. 2012; Damer and Deamer 2015, 2020). Most critically, TAGTFs allow the wet-and-dry cycle (perhaps cool-and-hot cycle too), which in turns drive vesicles through the cycle of breaking-and-repacking by changing the concentration of ions and other ingredients or the overall physical microenvironment within the “warm little pond” (Deamer and Barchfeld 1982; Zhu and Szostak 2009; Budin et al. 2012; Damer and Deamer 2015, 2020; Da Silva et al. 2015; Milshteyn et al. 2018; see also Pearce et al. 2017 for additional support). In contrast, HTAVs do not allow for such a cycle of packing-breaking-and-repacking.

If FUCAs and LUCA had indeed originated in TAGTFs, then an overflow or inundation induced by a heavy rainfall could have easily washed some LUCA from a “warm little pond” to another niche thus started the making of the two primary domains after LUCA. In contrast, LUCA from HTAVs demand multiple escapes from HTAVs to make the two primary domains. Apparently, the former scenario is less demanding than the latter.

Life After LUCA: Two Additional Controversies

Our thesis may also help resolve two additional controversies.

Our thesis emphasizes the role of acquisition and merger by vesicles and then protocells in driving the evolution of FUCAs to LUCA. With extensive acquisition and merger, FUCAs and LUCA most likely had proto-membranes that were heterochiral with both isoprenoid-based and fatty acid-based phospholipids (Peretó et al. 2004; Lombard et al. 2012; Lombard 2016).

The fact that heterochiral membrane is actually more stable than homochiral membranes in higher temperature is also more consistent with the possibility that FUCAs and the LUCA were “totipotent” cells that can survive in diverse and stressful environments (Shimada and Yamagishi 2011; Jain et al. 2014; Caforio et al. 2018; cf. Wächtershäuser 2003, 2007). A recent discovery that bacteria of the Fibrobacteres–Chlorobi–Bacteroidetes (FCB) group superphylum encode an archaeal lipid pathway in addition to a bacterial lipid pathway in natural settings and the encoded enzymes can be fully expressed in *E. coli* adds even more powerful evidence that heterochiral membrane could have existed in natural settings (Villanueva et al. 2020). Homochiral membranes came only after the divergence of Bacteria and Archaea from LUCA (Koga 2011, 2012, 2014), possibly driven by the coevolution of membrane and membrane proteins in different environments (Williams et al. 2013; Sojo 2019).

Our interpretation strongly questions the possibility that cellular DNA replication came from invasion by DNA viruses rather than FUCAs/LUCA with some kind of rudimentary DNA replication machinery (Forterre 2006).¹⁹ Forterre’s thesis of viral invasion does not explain how those different viruses had evolved without a proper cellular host in the first place.

More likely, the transition from RNA to DNA was accomplished by reverse transcription and this transition was not completed before LUCA diverged to Bacteria and Archaea. A fully functioning DNA replication machinery had evolved twice, once in Bacteria and once in Archaea (Leipe et al. 1999).

Ways Forward

Although we may never be able to (re-)make a whole cell de novo by starting with primitive building blocks of life (cf. Ichihashi et al. 2013), our new thesis does point to several directions for experimentally testing its key hypotheses.

¹⁹ Indeed, multiple escapes and multiple invasions are like two sides of the same miracle.

More concretely, in light of earlier experiments that have demonstrated various properties of vesicles, four types of experiment will be of particular interest (for earlier discussions, see Chen and Walde 2010; Schrum et al. 2010; Deamer et al. 2019; Damer and Deamer 2015, 2020; Lopez and Fiore 2019). If we can show these four possibilities experimentally, our thesis regarding the origin(s) of cell should be considered well supported.

First, in conditions that are somewhat similar to primitive environment, different vesicles made of different amphiphiles can absorb and especially engulf not only small molecules (e.g., amino acids, nucleotides, peptides, short oligonucleotides/RNAs or alike, metals, and other ingredients) but also RNA-peptide complexes as Woese's (2002) "supramolecular aggregates" with different efficiencies. If confirmed, these experiments will strongly support not only the possibility that proto-endocytosis by vesicles had been a key mechanism but also the possibility that the merging of the lipid membrane world and the RNA-peptide world had been a key step toward FUCAs according to the thesis advanced here (Fig. 1) and many others (e.g., Schrum et al. 2010; Damer and Deamer 2015, 2020).

Second, different vesicles with different in-takes of ingredients not only have different chemical, physiological, and bio-energetic properties, but more importantly, compete against each other via absorption, acquisition, and merger hence have different rate of persistence (as survival) within a system. This possibility, if demonstrated, should constitute a very decisive set of evidence for our thesis that centers on acquisition and merger by progenotes as HBMT. Ample experimental evidence exists that vesicles compete against each other for ingredient up-taking and that their growth and division in turn depend on such in-taking and subsequent reactions within (e.g., Chen et al. 2004; Zhu and Szostak 2009; Budin and Szostak 2011; Kurihara et al. 2015). Yet, the possibility that different vesicles based on different ingredients and with different encapsulated biomolecules compete against each other via acquisition and merger has not been directly tested, to the best of my knowledge: most existing experiments on vesicle merger-and-acquisition have either utilized only one type of vesicles or tested different types of vesicles separately.

Third, heterogeneous vesicles from acquisition and merger may be more stable and even "reproductive" than homogenous ones. This possibility, if demonstrated, will further buttress the thesis that heterogeneity had indeed been more advantageous in pre-FUCAs evolution, together with existing evidence that heterogeneous vesicles can be more stable than homogenous vesicles (Mansy and Szostak 2008; Chen and Walde 2010; Szostak 2011).

Fourth, recombined vesicles via acquisition and merger lead to chemical interactions that in turn lead to new physical-chemical and (proto-)physiological properties: such

new properties preview the coming of new metabolic pathways, and perhaps eventually, a more regulated coupling of metabolism and reproduction (see also Lopez and Fiore 2019). This possibility, admittedly more difficult to demonstrate, should powerfully buttress our thesis if demonstrated, because it directly tests the possibility that artificial vesicles via absorption, acquisition, and merger can eventually lead to some kind of coupling of metabolism and reproduction.

In addition, with computer simulation (for reviews of earlier attempts, see Lancet and Shenhav 2009; Klein et al. 2017), it may be shown that vesicles can eventually evolve into protocells with a tight coupling of RNAs and peptides as a form of proto standard genetic code, and eventually an efficient lineage can come to total domination within the system as the LUCA (Fig. 1). More concretely, within a population of vesicles, aided by a steady supply of amino acids and nucleotides, a rudimentary genetic code can evolve. Initially, such a rudimentary codon system may be a form of probabilistic (weak to firm) association between certain amino acids (e.g., Arginine) and certain short oligonucleotides (Szathmáry and Maynard Smith 1997; Cavalier-Smith 2001; Yarus et al. 2009; Yarus 2017). Gradually, however, these initially probabilistic associations become more stable and more frequent due to lock-in effects, which further accelerate the crystallization of the codon system (Vetsigian et al. 2006).²⁰

Finally, some new directions for sequence and structural phylogenetic analysis can also be conceived. In particular, analyses of components within the heat shock response apparatus, small proteins (including small HSPs) that are responsible for protecting the membrane (e.g., Hsp12), ubiquitin-like proteins (UBLs), and ubiquitin-related modifiers (URMs) will be of special interest. The evolutionary implications these components and proteins for the origins of FUCAs and LUCA have not been adequately explored.

Ubiquitin (UB) is universally conserved in eukaryotes while UBLs and URM are highly conserved in Bacteria and Archaea (Hochstrasser 2009; Richter et al. 2010). Moreover, key similarities between eukaryotic UB and bacterial UBLs (e.g., THiS and MoaD) are mostly structural: the key domain is the "small but versatile" beta-grasp fold that is common to UB, UBLs, and URM (Burroughs et al. 2012a, b). UB, UBLs, and URM play critical roles in regulating some of the vital functions of life, including responding to stress such as oxidative, hypoxic, osmotoxic, genotoxic, and heat. These functions are vital for survival but not for replication. UB and UBL (in the Urm1-Uba4 system, which is much closer to THiS and THiF) might have been molecular fossils from the more ancient sulfur-transfer pathway (Hochstrasser 2009, p. 425; van der Veen and Ploeh 2012, pp. 342–3). Metal

²⁰ Our testing of this possibility with computer simulation will be reported separately.

and sulfur-proteins might have been some of the first sets of proteins that were recruited or assembled into the first protocell. This fact is consistent with the hypothesis that life mostly originated from a hydrothermal environment that is rich in metal and sulfur. Moreover, the fact that beta-grasp fold of UBLs was a RNA-binding domain with connections to RNA metabolism also suggests that its origin has been ancient, most likely before LUCA (Burroughs et al. 2012a, b).²¹ All these facts point to the possibility that UBLs might have been an ancient protein family that play roles in both metabolism centered upon sulfur and RNA-binding/translation before LUCA.

If heat shock response apparatus, UBLs, and UBMs had been early inventions that were essential for the survival of FUCAs and LUCA, sequence and structural phylogenomic analysis of them may shed important new light upon the evolution of FUCAs and LUCA. For instance, whereas Bacteria and most Archaea do not encode the ubiquitin system, at least one particular archaea group, *Caldiarchaeum subterraneum* does and it does not appear to be a result of HGT (Nunoura et al. 2011; Koonin and Yutin 2014). This suggests that ubiquitin evolved from UBLs only in some archaea lineages that later on became ancestors of eukaryotes.

Concluding Remarks

This article advances a new thesis regarding the origin(s) of FUCAs and LUCA as the first cell(s) by broadening our understanding of what constitutes variation, selection, and retention in the pre-Darwinian evolution before LUCA. Most critically, I argue that vesicles' acquisition and merger via (proto-)endosymbiosis and (proto-)endocytosis has been a powerful force for both variations and selection, and hence a critical mechanism leading to the origin(s) of FUCAs and then LUCA. Moreover, the impact of this mechanism is not limited to genetic, but also structural, functional, and metabolic due to HBMT's more extensive impact than HGT. Our thesis is not only supported by quite extensive evidence but also resolves some key controversies regarding the origin(s) of FUCAs and then LUCA. Our thesis also points to new directions for further inquiries.

Acknowledgements For insightful comments and suggestions, I thank Bruce Damer, David Deamer, Aaron Goldman, Eugene Koonin, four anonymous reviewers, and the editor. Jin He and Tian Yue have provided outstanding research assistance. Special thanks go to Chang-an Liu for drawing the two figures.

Author Contribution ST developed the theory and wrote the paper.

Funding The search has not received any funding.

²¹ Indeed, UBLs play critical roles in pre-mRNA splicing in eukaryotes (Chanarat and Mishra 2018).

Declarations

Conflict of interest The author declares that he has no competing interests.

References

- Adamala K, Szostak JW (2013) Competition between model protocells driven by an encapsulated catalyst. *Nat Chem* 5:495–501
- Altan-Bonnet N (2017) Lipid tales of viral replication and transmission. *Trends Cell Biol* 27:201–213
- Armstrong DL, Lancet D, Zidovetzki R (2018) Replication of simulated prebiotic amphiphilic vesicles in a finite environment exhibits complex behavior that includes high progeny variability and competition. *Astrobiology* 18:419–430
- Baum DA (2015) Selection and the origin of cells. *Bioscience* 65:678–684
- Bedau MA, Cleland CE (2010) The nature of life: classical and contemporary perspectives from philosophy and science. Cambridge University Press, Cambridge
- Black RA, Blosser MC, Stottrup BL, Tavakley R, Deamer DW, Keller SL (2013) Nucleobases bind to and stabilize aggregates of a prebiotic amphiphile, providing a viable mechanism for the emergence of protocells. *Proc Natl Acad Sci USA* 110:13272–13276
- Booth A, Doolittle WF (2015) Eukaryogenesis, how special really? *Proc Natl Acad Sci USA* 112:10278–10285
- Bouchard F (2014) Ecosystem evolution is about variation and persistence, not populations and reproduction. *Biol Theory* 9:382–391
- Brack A (1998) The molecular origins of life: assembling pieces of the puzzle. Cambridge University Press, Cambridge
- Budin I, Szostak JW (2011) Physical effects underlying the transition from primitive to modern cell membranes. *Proc Natl Acad Sci USA* 108:5249–5254
- Budin I, Debnath A, Szostak JW (2012) Concentration-driven growth of model protocell membranes. *J Am Chem Soc* 134:20812–20819
- Burroughs AM, Iyer LM, Aravind L (2012a) The natural history of ubiquitin and ubiquitin-related domains. *Front Biosci* 17:1433–1460
- Burroughs AM, Iyer LM, Aravind L (2012b) Structure and evolution of ubiquitin and ubiquitin-related domains. *Methods Mol Biol* 832:15–63
- Caforio A, Siliakus MF, Exterkate M, Jain S, Jumde VR, Andringa RLH, Kengen SWM, Minnaard AJ, Driessen AJM, van der Oost J (2018) Converting *Escherichia coli* into an archaeobacterium with a hybrid heterochiral membrane. *Proc Natl Acad Sci USA* 115:3704–3709
- Cantine MD, Fournier GP (2018) Environmental adaptation from the origin of life to the last universal common ancestor. *Orig Life Evol Biosph* 48:35–54
- Carvunis AR, Rolland T, Wapinski I, Calderwood MA, Yildirim MA, Simonis N, Charlotiaux B, Hidalgo CA, Barbette J, Santhanam B, Brar GA, Weissman JS, Regev A, Thierry-Mieg N, Cusick ME, Vidal M (2012) Proto-genes and de novo gene birth. *Nature* 487:370–374
- Cavalier-Smith T (2001) Obcells as proto-organisms: membrane heredity, lithophosphorylation, and the origins of the genetic code, the first cells, and photosynthesis. *J Mol Evol* 53:555–595
- Chanarat S, Mishra SK (2018) Emerging roles of ubiquitin-like proteins in pre-mRNA splicing. *Trends Biochem Sci* 43:896–907

- Charlebois RL, Doolittle WF (2004) Computing prokaryotic gene ubiquity: rescuing the core from extinction. *Genome Res* 14:2469–2477
- Chen IA, Walde P (2010) From self-assembled vesicles to protocells. *Cold Spring Harb Perspect Biol* 2:a002170
- Chen IA, Roberts RW, Szostak JW (2004) The emergence of competition between model protocells. *Science* 305:1474–1476
- Cheng Z, Luisi PL (2003) Coexistence and mutual competition of vesicles with different size distributions. *J Phys Chem B* 107:10940–10945
- Coleman GA, Pancost RD, Williams TA (2019) Investigating the origins of membrane phospholipid biosynthesis genes using outgroup-free rooting. *Genome Biol Evol* 11:883–898
- Cornell CE, Black RA, Xue M, Litz HE, Ramsay A, Gordon M, Mileant A, Cohen ZR, Williams JA, Lee KK, Drobny GP, Keller SL (2019) Prebiotic amino acids bind to and stabilize prebiotic fatty acid membranes. *Proc Natl Acad Sci USA* 116:17239–17244
- Cornish-Bowden A, Cárdenas ML (2017) Life before LUCA. *J Theor Biol* 434:68–74
- Da Silva L, Maurel MC, Deamer D (2015) Salt-promoted synthesis of RNA-like molecules in simulated hydrothermal conditions. *J Mol Evol* 80:86–97
- Dacks JB, Field MC, Buick R, Eme L, Gribaldo S, Roger AJ, Brochier-Armanet C, Devos DP (2016) The changing view of eukaryogenesis-fossils, cells, lineages and how they all come together. *J Cell Sci* 129:3695–3703
- Damer B, Deamer D (2015) Coupled phases and combinatorial selection in fluctuating hydrothermal pools: a scenario to guide experimental approaches to the origin of cellular life. *Life* 5:872–887
- Damer B, Deamer D (2020) The hot spring hypothesis for an origin of life. *Astrobiology* 20:429–452
- de Duve C (2005a) The onset of natural selection. *Nature* 433:581–582
- de Duve C (2005b) Singularities: landmarks on the pathways of life. Cambridge University Press, Cambridge
- de Farias ST, dos Santos JAP, Regro TG, Jose MV (2017) Origin and evolution of RNA-dependent RNA polymerase. *Front Genet*. <https://doi.org/10.3389/fgene.2017.00125>
- Deamer DW (2019) Assembling life. Oxford University Press, Oxford
- Deamer DW, Barchfeld GL (1982) Encapsulation of macromolecules by lipid vesicles under simulated prebiotic conditions. *J Mol Evol* 18:203–206
- Deamer D, Damer B, Kompanichenko V (2019) Hydrothermal chemistry and the origin of cellular life. *Astrobiology* 19:1523–1537
- Doolittle WF (2000) The nature of the universal ancestor and the evolution of the proteome. *Curr Opin Struct Biol* 10:355–358
- Doolittle WF (2014) Natural selection through survival alone, and the possibility of Gaia. *Biol Philos* 29:415–423
- Doolittle WF, Inkpen SA (2018) Processes and patterns of interaction as units of selection: an introduction to ITSNTS thinking. *Proc Natl Acad Sci USA* 115:4006–4014
- Egel R (2009) Peptide-dominated membranes preceding the genetic takeover by RNA: latest thinking on a classic controversy. *BioEssays* 31:1100–1109
- Egel R (2012) Life's order, complexity, organization, and its thermodynamic-holistic imperatives. *Life* 2:323–363
- Egel R (2017) 'Parabiogenic evolution': from stochasticity in geochemical and subsequent processes to genes, genomes and modular cells. Preprints 2017100153. <https://doi.org/10.20944/preprints201710.0153.v1>
- El Baidouri F, Venditti C, Suzuki S, Meade A, Humpries S (2020) Phenotypic reconstruction of the last universal common ancestor reveals a complex cell. *bioRxiv*. <https://doi.org/10.1101/2020.08.20.260398>
- Eme L, Spang A, Lombard J, Stairs CW, Ettema TJG (2017) Archaea and the origin of eukaryotes. *Nat Rev Microbiol* 15:711–723
- Forterre P (2006) Three RNA cells for ribosomal lineages and three DNA viruses to replicate their genomes: a hypothesis for the origin of cellular domain. *Proc Natl Acad Sci USA* 103:3669–3674
- Forterre P, Philippe H (1999) Where is the root of the universal tree of life? *BioEssays* 21:871–879
- Fournier GP, Andam CP, Gogarten JP (2015) Ancient horizontal gene transfer and the last common ancestors. *BMC Evol Biol* 15:70
- Fox SW (1973) Molecular evolution to the first cells. *Pure Appl Chem* 34:641
- Fox GE (2010) Origin and evolution of the ribosome. *Cold Spring Harb Perspect Biol* 2:a003483. <https://doi.org/10.1101/cshperspect.a003483>
- Fox GE, Luehrsen KR, Woese CR (1982) Archaeobacterial 5S Ribosomal RNA. *Zbl Bakt Hyg. i. Abr Orig C* 3:330–345
- Fox GE, Tran Q, Yonath A (2012) An exit cavity was crucial to the polymerase activity of the early ribosome. *Astrobiology* 12:57–60
- Francis BR (2011) An alternative to the RNA world hypothesis. *Trends Evol Biol* 3:e2
- Francis BR (2013) Evolution of the genetic code by incorporation of amino acids that improved or changed protein function. *J Mol Evol* 77:134–158
- Francis BR (2015) The hypothesis that the genetic code originated in coupled synthesis of proteins and the evolutionary predecessors of nucleic acids in primitive cells. *Life* 5:467–505
- Frenkel-Pinter M, Haynes JW, Mohyeldin AM, Martin C, Sargon AB, Petrov AS, Krishnamurthy R, Hud NV, Williams LD, Leman LJ (2020) Mutually stabilizing interactions between proto-peptides and RNA. *Nat Commun* 11:3137
- Fry I (2011) The role of natural selection in the origin of life. *Orig Life Evol Biosph* 41:3–16
- Gánti T (1997) Biogenesis itself. *J Theor Biol* 187:583–593
- Garson J (2017) A generalized selected effects theory of function. *Phil Sci* 84:523–543
- Glansdorff N (2002) About the last common ancestor, the universal life-tree, and lateral gene transfer: a reappraisal. *Mol Microbiol* 38:177–185
- Glansdorff N, Xu Y, Labedan B (2008) The last universal common ancestor: emergence, constitution and genetic legacy of an elusive forerunner. *Biol Direct* 3:29
- Gogarten JP, Deamer D (2016) Is LUCA a thermophilic progenote? *Nat Microbiol* 1:16229
- Goldman AD, Kacar B (2021) Cofactors are remnants of life's origin and early evolution. *J Mol Evol* 89:127–133
- Harold FM (2014) In search of cell history: the evolution of life's building blocks. University of Chicago Press, Chicago
- Harris AJ, Goldman AD (2021) The very early evolution of protein translocation across membranes. *PLoS Comput Biol*. <https://doi.org/10.1371/journal.pcbi.1008623>
- Harris JK, Kelley ST, Spiegelman GB, Pace NR (2003) The genetic core of the universal ancestor. *Genome Res* 13:407–412
- Higgs PG (2016) The effect of limited diffusion and wet-dry cycling on reversible polymerization reactions: implications for prebiotic synthesis of nucleic acids. *Life* 6:24
- Higgs PG, Lehman N (2015) The RNA world: molecular cooperation at the origins of life. *Nat Rev Genet* 16:7–17
- Hochstrasser M (2009) Origin and function of ubiquitin-like proteins. *Nature* 458:422–429
- Horning DP, Joyce GF (2017) Amplification of RNA by an RNA polymerase ribozyme. *Proc Natl Acad Sci USA* 113:9786–9791
- Ichihashi N, Yomo T (2014) Positive roles of compartmentalization in internal reactions. *Curr Opin Chem Biol* 22:12–17
- Ichihashi N, Usui K, Kazuta Y, Sunami T, Matsuura T, Yomo T (2013) Darwinian evolution in a translation-coupled RNA replication system within a cell-like compartment. *Nat Commun* 4:2494. <https://doi.org/10.1038/ncomms3494>

- Jackson JB (2016) Natural pH gradients in hydrothermal alkali vents were unlikely to have played a role in the origin of life. *J Mol Evol* 83:1–11
- Jackson JB (2017) Ancient living organisms escaping from, or imprisoned in, the vents? *Life* 7:36
- Jain R, Rivera MC, Lake JA (1999) Horizontal gene transfer among genomes: the complexity hypothesis. *Proc Natl Acad Sci USA* 96:3801–3806
- Jain S, Caforio A, Driessen AJM (2014) Biosynthesis of archaeal membrane ether lipids. *Front Microbiol* 5:641
- Joyce GF (2002) The antiquity of RNA-based evolution. *Nature* 418:214–221
- Joyce GF, Szostak JW (2018) Protocells and RNA self-replication. *Cold Spring Harb Perspect Biol* 10:a034801
- Kandler O (1994a) The early diversification of life. In: Bengtson S (ed) *Early life on earth: Nobel Symposium no. 84*. Columbia University Press, New York, pp 152–160
- Kandler O (1994b) Cell wall biochemistry in Archaea and its phylogenetic implications. *J Biol Phys* 20:165–169
- Kandler O (1998) The early diversification of life and the origin of the three domains: a proposal. In: Wiegel J, Adams MWW (eds) *Thermophiles: the keys to molecular evolution and the origin of life*. Taylor & Francis, London, pp 19–28
- Kim Y, Opron K, Burton ZF (2019) A tRNA- and anticodon-centric view of the evolution of aminoacyl-tRNA synthetases, tRNAomes, and the genetic code. *Life* 9:37. <https://doi.org/10.3390/life9020037>
- Klein A, Bock M, Alt W (2017) Simple mechanisms of early life-simulation model on the origin of semi-cells. *Biosystems* 151:34–42
- Koga Y (2011) Early evolution of membrane lipids: how did the lipid divide occur? *J Mol Evol* 72:274–282
- Koga Y (2012) Thermal adaptation of the archaeal and bacterial lipid membranes. *Archaea* 2012:789652
- Koga Y (2014) From promiscuity to the lipid divide: on the evolution of distinct membranes in Archaea and Bacteria. *J Mol Evol* 78:234–242
- Koonin EV (2003) Comparative genomics, minimal gene-sets and the last universal common ancestor. *Nat Rev Microbiol* 1:127–136
- Koonin EV (2009) On the origin of cells and viruses: primordial virus world scenario. *Ann N Y Acad Sci* 1178:47–64
- Koonin EV (2014a) Carl Woese's vision of cellular evolution and the domains of life. *RNA Biol* 11:197–204
- Koonin EV (2014b) The origin of cellular life. *Antonie Van Leeuwenhoek* 106:27–41
- Koonin EV (2016) Viruses and mobile elements as drivers of evolutionary transitions. *Philos Trans R Soc Lond B Biol Sci* 371:20150442
- Koonin EV (2017) Evolution of RNA- and DNA-guided antiviral defense systems in prokaryotes and eukaryotes: common ancestry vs convergence. *Biol Direct* 12:5
- Koonin EV, Martin W (2005) On the origin of genomes and cells within inorganic compartments. *Trends Genet* 21:647–654
- Koonin EV, Dolja VV (2013) A virocentric perspective on the evolution of life. *Curr Opin Virol* 3:546–557
- Koonin EV, Yutin N (2014) The dispersed archaeal eukaryome and the complex archaeal ancestor of eukaryotes. *Cold Spring Harb Perspect Biol* 6:a016188
- Koonin EV, Makarova KS (2017) Mobile genetic elements and evolution of CRISPR-Cas systems: all the way there and back. *Genome Biol Evol* 9(10):2812–2825
- Koonin EV, Novozhilov AS (2017) Origin and evolution of the universal genetic code. *Annu Rev Genet* 27:45–62
- Koonin EV, Makarova KS, Wolf YI, Krupovic M (2020) Evolutionary entanglement of mobile genetic elements and host defence systems: guns for hire. *Nat Rev Genet* 21:119–131
- Kovacs NA, Petrov AS, Lanier KA, Williams LD (2017) Frozen in time: the history of proteins. *Mol Biol Evol* 34:1252–1260
- Krupovic M, Dolja VV, Koonin EV (2020) The LUCA and its complex virome. *Nat Rev Microbiol* 18:661–670
- Ku C, Nelson-Sathi S, Roettger M, Sousa FL, Lockhart PJ, Bryant D, Hazkani-Covo E, McInerney JO, Landan G, Martin WF (2015) Endosymbiotic origins and differential loss of eukaryotic genes. *Nature* 254:427–432
- Kunnev D (2020) Origin of life: the point of no return. *Life* 10:269. <https://doi.org/10.3390/life10110269>
- Kurihara K, Okura Y, Matsuo M, Toyota T, Suzuki K, Sugawara T (2015) A recursive vesicle-based model protocell with a primitive model cell cycle. *Nat Commun* 6:8352
- Lancet D, Shenhav B (2009) Compositional lipid protocells: reproduction without polynucleotides. In: Rasmussen S, Bedau MA, Chen L, Deamer D, Krakauer DC, Packard NH, Stadler PF (eds) *Protocells: bridging nonliving and living matter*. MIT, Cambridge, pp 233–252
- Lane N, Martin WF (2015) Eukaryotes really are special, and mitochondria are why. *Proc Natl Acad Sci USA* 112:E4823
- Lane N, Allen JF, Martin W (2010) How did LUCA make a living? Chemiosmosis in the origin of life. *BioEssays* 32:271–280
- Lanier KA, Williams LD (2017) The origin of life: models and data. *J Mol Evol* 84:85–92
- Lanier KA, Petrov AS, Williams LD (2017) The central symbiosis of molecular biology: molecules in mutualism. *J Mol Evol* 85:8–13
- Lazcano A (2010) Historical development of origins research. *Cold Spring Harb Perspect Biol* 2:a002089
- Lear JD, Wasserman ZR, DeGrado WF (1988) Synthetic amphiphilic peptide models for protein ion channels. *Science* 240:1177–1181
- Leipe DD, Aravind L, Koonin EV (1999) Did DNA replication evolve twice independently? *Nucleic Acids Res* 27:3389–3401
- Lombard J (2016) Early evolution of polyisoprenol biosynthesis and the origin of cell walls. *PeerJ* 4:e2626
- Lombard J, López-García P, Moreira D (2012) The early evolution of lipid membranes and the three domains of life. *Nat Rev Microbiol* 10:507–515
- Longo LM, Despotović D, Weil-Ktorza O, Walker MJ, Jabłońska J, Fridmann-Sirkis Y, Varani G, Metanis N, Tawfik DS (2020) Primordial emergence of a nucleic acid-binding protein via phase separation and statistical ornithine-to-arginine conversion. *Proc Natl Acad Sci USA* 117:15731–15739
- López-García P, Moreira D (2015) Open questions on the origin of eukaryotes. *Trends Ecol Evol* 30:697–708
- Lopez A, Fiore M (2019) Investigating prebiotic protocells for a comprehensive understanding of the origins of life: a prebiotic systems chemistry perspective. *Life* 9(2):49. <https://doi.org/10.3390/life9020049>
- Luisi PL (2006) *The emergence of life*. Cambridge University Press, Cambridge
- Luisi PL (2016) *The emergence of life*, 2nd edn. Cambridge University Press, Cambridge
- Ma BG, Chen L, Ji HF, Chen ZH, Yang FR, Wang L, Qu G, Jiang YY, Ji C, Zhang HY (2008) Characters of very ancient proteins. *Biochem Biophys Res Commun* 366:607–611
- Mansy SS, Szostak JW (2008) Thermostability of model protocell membranes. *Proc Natl Acad Sci USA* 105:13351–13355
- Mansy SS, Schrum JP, Krishnamurthy M, Tobe S, Treco DA, Szostak JW (2008) Replication of a genetic polymer inside of a model protocell. *Nature* 454:122–125
- Margulis L (1981) Symbiosis in cell evolution: life and its environment on the early earth. W. H. Freeman, San Francisco
- Margulis L (1991) Symbiogenesis and symbiogenesis. In: Margulis L, Fester R (eds) *Symbiosis as a source of evolutionary innovation: speciation and morphogenesis*. MIT Press, Cambridge, pp 1–14

- Martin W, Russell MJ (2003) On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells. *Philos Trans R Soc Lond B Biol Sci* 358:59–85
- Martin W, Russell MJ (2007) On the origin of biochemistry at an alkaline hydrothermal vent. *Philos Trans R Soc Lond B Biol Sci* 362:1887–1926
- Martin W, Baross J, Kelley D, Russell MJ (2008) Hydrothermal vents and the origin of life. *Nat Rev Microbiol* 6:805–814
- Miller S, Krijnse-Locker J (2008) Modification of intracellular membrane structures for virus replication. *Nat Rev Microbiol* 6:363–374
- Miller SL, Schopf JW, Lazcano A (1997) Oparin's "origin of Life": sixty years later. *J Mol Evol* 44:351–353
- Milshcheyn D, Damer B, Havig J, Deamer D (2018) Amphiphilic compounds assemble into membranous vesicles in hydrothermal hot spring water but not in sea water. *Life* 8:11
- Moreira D, Lombard J, López-García P (2009) Ten reasons to exclude viruses from the tree of life. *Nat Rev Microbiol* 7:306–311
- Mulkidjanian AY, Galperin MY, Koonin EV (2009) Coevolution of primordial membranes and membrane proteins. *Trends Biochem Sci* 34:206–215
- Mulkidjanian AY, Bychkov AY, Dibrova DV, Galperin MY, Koonin EV (2012) Origin of first cells at terrestrial, anoxic geothermal fields. *Proc Natl Acad Sci USA* 109:E821–E830
- Norris V, Raine DJ (1998) A fission-fusion origin for life. *Orig Life Evol Biosph* 28:523–537
- Nunoura K, Takaki Y, Kakuta J, Nishi S, Sugahara J, Kazama H, Chee GJ, Hattori M, Kanai A, Atomi H, Takai K, Takami H (2011) Insights into the evolution of *Archaea* and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. *Nucleic Acids Res* 39:3204–3223
- Olasagasti F, Rajamani S (2019) Lipid-assisted polymerization of nucleotides. *Life*. <https://doi.org/10.3390/life9040083>
- O'Malley MA (2014) Endosymbiosis and its implications for evolutionary theory. *Proc Natl Acad Sci USA* 112:10270–10277
- Oparin AI (1953 [1938]) *The origin of life*, 2nd edn. Dover, New York [First published by MacMillan.]
- Orgel LE (2004) Prebiotic chemistry and the origin of the RNA world. *Crit Rev Biochem Mol Biol* 39:99–123
- Pascal R, Pross A (2016) The logic of life. *Orig Life Evol Biosph* 46:507–513
- Pearce BKD, Pudritz RE, Semenov DA, Henning TK (2017) Origin of the RNA world: the fate of nucleobases in warm little ponds. *Proc Natl Acad Sci USA* 114:11327–11332
- Peretó J, López-García P, Moreira D (2004) Ancestral lipid biosynthesis and early membrane evolution. *Trends Biochem Sci* 29:469–477
- Petrov AS, Gulen B, Norris AM, Kovacs NA, Bernier CR, Lanier KA, Fox GE, Harvey SC, Wartell RM, Hud NV, Williams LD (2015) History of the ribosome and the origin of translation. *Proc Natl Acad Sci USA* 112:15396–15401
- Philippe H, Forterre P (1999) The rooting of the universal tree of life is not reliable. *J Mol Evol* 49:509–523
- Pohorille A, Deamer D (2009) Self-assembly and function of primitive cell membranes. *Res Microbiol* 160:449–456
- Pohorille A, Wilson MA, Chipot C (2003) Membrane peptides and their role in protobiological evolution. *Orig Life Evol Biosph* 33:173–197
- Pressman A, Blanco C, Chen IA (2015) The RNA world as a model system to study the origin of life. *Curr Biol Rev* 25:R953–R963
- Prosdocimi F, José MV, de Farias ST (2019) The first universal common ancestor (FUCA) as the earliest ancestor of LUCA's (last UCA) lineage. In: Pontarotti P (ed) *Evolution, origin of life, concepts and methods*. Springer, Berlin, pp 43–54
- Pross A (2011) Toward a general theory of evolution: extending Darwinian theory to inanimate matter. *J Syst Chem* 2:1
- Pross A (2013) The evolutionary origin of biological function and complexity. *J Mol Evol* 76:185–191
- Puigbò P, Wolf YI, Koonin EV (2009) Search for a 'tree of life' in the thicket of the phylogenetic forest. *J Biol* 8:59
- Qiao H, Hu N, Bai J, Ren L, Liu Q, Fang L, Wang Z (2017) Encapsulation of nucleic acids into giant unilamellar vesicles by freeze-thaw: a way protocells may form. *Orig Life Evol Biosph* 47:499–510
- Ranea JA, Sillero A, Thornton JM, Orengo CA (2006) Protein superfamily evolution and the last universal common ancestor. *J Mol Evol* 63:513–525
- Richter K, Haslbeck M, Buchner J (2010) The heat shock response: life on the verge of death. *Mol Cell* 40:253–266
- Robertson MP, Joyce GF (2012) The origins of the RNA world. *Cold Spring Harb Perspect Biol* 4:a003608
- Saad NY (2018) A ribonucleopeptide world at the origin of life. *J Syst Evol* 56:1–13
- Sagan L (1967) On the origin of mitosing cells. *J Theoret Biol* 14:225–274
- Saha R, Chen IA (2015) Origin of life: protocells red in tooth and claw. *Curr Biol* 25:R1175–R1177
- Saha R, Pohorille A, Chen IA (2014) Molecular crowding and early evolution. *Orig Life Evol Biosph* 44:319–324
- Saha R, Verbanic S, Chen IA (2018) Lipid vesicles chaperone an encapsulated RNA aptamer. *Nat Commun* 9:2313
- Schrump JP, Zhu TF, Szostak JW (2010) The origins of cellular life. *Cold Spring Harb Perspect Biol* 2:a002212
- Segré D, Ben-Eli D, Deamer DW, Lancet D (2001) The lipid world. *Orig Life Evol Biosph* 31:119–145
- Sengupta S, Higgs PG (2015) Pathways of genetic code evolution in ancient and modern organisms. *J Mol Evol* 80:229–243
- Shah V, de Bouter J, Pauli Q, Tupper AS, Higgs PG (2019) Survival of RNA replicators is much easier in protocells than in surface based, spatial systems. *Life* 9:65
- Shimada H, Yamagishi A (2011) Stability of heterochiral hybrid membrane made of bacterial sn-G3P lipids and archaeal sn-G1P lipids. *Biochemistry* 50:4114–4120
- Smith E, Morowitz HJ (2016) *The origin and nature of life on earth*. Cambridge University Press, Cambridge
- Sojo V (2019) Why the lipid divide? Membrane proteins as drivers of the split between the lipids of the three domains of life. *BioEssays* 2019(41):1800251
- Sojo V, Dessimoz C, Pomiankowski A, Lane N (2016) Membrane proteins are dramatically less conserved than water-soluble proteins across the tree of life. *Mol Biol Evol* 33:2874–2884
- Spang A, Saw JH, Jørgensen SL, Zaremba-Niedzwiedzka K, Martijn J, Lind AE, van Eijk R, Schleper C, Guy L, Ettema TJG (2015) Complex Archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* 521:173–179
- Spitzer J (2017) Emergence of life on earth: a physicochemical jigsaw puzzle. *J Mol Evol* 84:1–7
- Spitzer J, Poolman B (2009) The role of biomacromolecular crowding, ionic strength and physicochemical gradients in the complexities of life's emergence. *Microbiol Mol Biol Rev* 73:371–388
- Spitzer J, Poolman B (2016) Complex molecular mixtures under cycling gradients as basis for life's origins. *bioRxiv*. <https://doi.org/10.1101/050740>
- Spitzer J, Pielak GJ, Poolman B (2015) Emergence of life: physical chemistry changes the paradigm. *Biol Direct* 10:33
- Staley JT, Fuerst JA (2017) Ancient, highly conserved protein from a LUCA with complex cell biology provide evidence in support of the nuclear compartment commonality (NuCom) hypothesis. *Res Microbiol* 168:395–412

- Storz G, Wolf YI, Ramamurthi KS (2014) Small proteins can no longer be ignored. *Annu Rev Biochem* 83:753–777
- Szathmáry E, Maynard Smith J (1997) From replicators to reproducers: the first major transitions leading to life. *J Theor Biol* 187:555–571
- Szostak JW (2011) An optimal degree of physical and chemical heterogeneity for the origin of life? *Philos Trans R Soc B Biol Sci* 366:2894–2901
- Szostak JW, Bartel DP, Luisi PL (2001) Synthesizing life. *Nature* 409:387–390
- Tagami S, Attwater J, Holliger P (2017) Simple peptides derived from the ribosomal core potentiate RNA polymerase ribozyme function. *Nat Chem* 9:325–332
- Takagi YA, Nguyen DH, Wexler TB, Goldman AD (2020) The Coevolution of cellularity and metabolism following the origin of life. *J Mol Evol* 88:698–617
- Tessera M (2018) Is pre-Darwinian evolution possible? *Biol Direct* 13:18
- Thomas JA, Rana FR (2007) The influence of environmental conditions, lipid composition, and phase behavior on the origin of cell membranes. *Orig Life Evol Biosph* 37:267–285
- Toman J, Flegr J (2017) Stability-based sorting: the forgotten process behind (not only) biological evolution. *J Theor Biol* 435:29–41
- van der Schaar HM, Dorobantu CM, Albulescu L, Strating JRPM, van Kuppeveld FJM (2016) Fat(al) attraction: picornaviruses usurp lipid transfer at membrane contact sites to create replication organelles. *Trends Microbiol* 24:535–546
- van der Veen AG, Ploeh HL (2012) Ubiquitin-like proteins. *Annu Rev Biochem* 81:323–357
- Villanueva L, von Meijenfeldt FAB, Westbye AB, Yadav S, Hopmans EC, Dutilh BE, Sinninghe Damsté JS (2020) Bridging the membrane lipid divide: Bacteria of the FCB group superphylum have the potential to synthesize archaeal ether lipids. *ISME J*. <https://doi.org/10.1038/s41396-020-00772-2>
- Vetsigian K, Woese C, Goldenfeld N (2006) Collective evolution and the genetic code. *Proc Natl Acad Sci USA* 103:10696–10701
- Vitas M, Dobovišek A (2018) In the beginning was a mutualism-on the origin of translation. *Orig Life Evol Biosph* 48:223–243
- Wächtershäuser G (2003) From pre-cells to Eukarya—a tale of two lipids. *Mol Microbiol* 47:13–22
- Wächtershäuser G (2007) On the chemistry and evolution of the pioneer organism. *Chem Biodivers* 4:584–602
- Wei C, Pohorille A (2011) Permeation of nucleosides through lipid bilayers. *J Phys Chem B* 115:3681–3688
- Wei C, Pohorille A (2013) Permeation of aldoentoses and nucleosides through fatty acid and phospholipid membranes: implication to the origins of life. *Astrobiology* 13:177–188
- Weiss MC, Sousa FL, Mrnjavac N, Neukirchen S, Roettger M, Nelson-Sathi S, Martin MF (2016) The physiology and habitat of the last universal common ancestor. *Nat Microbiol* 1:16116
- White HB (1976) Coenzymes as fossils of an earlier metabolic state. *J Mol Evol* 7:101–104
- Williams TA, Foster PG, Cox CJ, Embley TM (2013) An archaeal origin of eukaryotes supports only two primary domains of life. *Nature* 504:231–236
- Wilson MA, Wei C, Pohorille A (2014) Towards coevolution of membrane proteins and metabolism. *Orig Life Evol Biosph* 44:357–361
- Woese CR (1982) Archaeobacteria and cellular origins: an overview. *Zbl Bakt Hygi Abr Orig C3*:1–17
- Woese CR (1998) The universal ancestor. *Proc Natl Acad Sci USA* 95:6854–6859
- Woese CR (2000) Interpreting the universal phylogenetic tree. *Proc Natl Acad Sci USA* 97:8392–8396
- Woese CR (2001) Translation: in retrospect and prospect. *RNA* 7:1055–1067
- Woese CR (2002) On the evolution of cells. *Proc Natl Acad Sci USA* 99:8742–8747
- Woese CR (2004) A new biology for a new century. *Microbiol Mol Biol Rev* 68:173–186
- Woese CR, Fox GE (1977) The concept of cellular evolution. *J Mol Evol* 10:1–6
- Woese CR, Magrum LJ, Fox GE (1978) Archaeobacteria. *J Mol Evol* 11:245–252
- Wolf YI, Koonin EV (2007) On the origin of the translation system and the genetic code in the RNA world by means of natural selection, exaptation, and subfunctionalization. *Biol Direct* 2:14. <https://doi.org/10.1186/1745-6150-2-14>
- Wolf YI, Aravind L, Grishin NV, Koonin EV (1999) Evolution of aminoacyl-tRNA synthetases—analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. *Genome Res* 9(8):689–710
- Wong JT (1975) A coevolution theory of the genetic code. *Proc Natl Acad Sci USA* 72:1909–1912
- Wong JT (1981) Coevolution of genetic code and amino acid biosynthesis. *Trends Biochem Sci* 6:33–36
- Yarus M, Widmann JJ, Knight R (2009) RNA–amino acid binding: a stereochemical era for the genetic code. *J Mol Evol* 69:406–429
- Yarus M (2017) The genetic code and RNA-amino acid affinities. *Life* 7:13
- Zhu TF, Szostak JW (2009) Coupled growth and division of model protocell membranes. *J Am Chem Soc* 131:5705–5713
- Zhu TF, Adamala K, Zhang N, Szostak JW (2012) Photochemically driven redox chemistry induces protocell membrane pearling and division. *Proc Natl Acad Sci USA* 25:9828–9832
- Zimorski V, Ku C, Martin WF, Gould SB (2014) Endosymbiotic theory for organelle origins. *Curr Opin Microbiol* 22:38–48