

The Origin(s) of Cell: Some Radical New Hypotheses

Shiping Tang

(Fudan University, Shanghai, China, twukong@fudan.edu.cn)

Abstract

The central mechanism of biological evolution, variation-selection-inheritance (VSI), is one of the most universal mechanisms known. Much of our understanding of VSI, however, has been dominated by the Neo-Darwinian Modern Synthesis with a rather narrow understanding of what constitutes variation, selection, and inheritance. This unduly narrow understanding of VSI might have been a key cause behind our failure to adequately explain some critical puzzles in biological evolution, from the origin of the first cell to the origin of the eukaryotes, the puzzling biology of metabolism, apoptosis, aging, and cancer in metazoan.

I broaden our understanding of VSI, in a spirit that is somewhat similar to several recent contributions and then extend this broadened view of VSI to its “natural” starting point: the origin of the First Universal Cell Ancestor (FUCA). I advance three principal arguments. First, survival comes before replication. Before the coming of reproducer and replicator, there must be survivors, to paraphrase Szathmary and Maynard Smith (1997). Second, natural selection, especially the non-Darwinian kind, can operate without replication or even metabolism, as long as different molecules, complexes, and vesicles have differential survival rate within a system. Third, merger and acquisition, via breaking-and-re-encapsulation, endocytosis, endosymbiosis, and processes similar to them, had been a far more powerful force of variation and selection in the pre-Darwinian period of evolution that led to LUCA and long before eukaryogenesis. Endosymbiosis therefore had been a far more foundational force than even Lynn Margulis and many of her supporters have appreciated. Our thesis thus goes beyond Woese’s emphasis of horizontal gene transfer (HGT) and actually subsumes HGT with Margulis’ emphasis of endosymbiosis. Combing these three new perspectives with other perspectives and evidence sheds important new light upon the origin of FUCA, the singular water-shedding moment in the evolution of life.

Keywords: LUCA; FUCA; Horizontal Biomolecule Transfer; Horizontal Gene Transfer

Contents

Introduction.....	2
Seven Premises.....	4
Core Hypotheses	7
Evolution before FUCA	7
From FUCA to LUCA.....	10
Evidences: Chemical and Biological-Functional	11
Controversies Resolved, Partly and Possibly	18
Concluding Remarks: Unsolved Mysteries and Ways Forward	22

References.....	24
-----------------	----

Introduction

The central mechanism of biological evolution, variation-selection-inheritance (VSI), is one of the most universal mechanisms known. Much of our understanding of VSI, however, has been dominated by the Neo-Darwinian Modern Synthesis with a rather narrow understanding of what constitutes variation, selection, and inheritance. This unduly narrow understanding of VSI might have been a key cause behind our failure to adequately explain some critical puzzles in biological evolution, from the origin of the first cell to the origin of the eukaryotes, the puzzling biology of metabolism, apoptosis, aging, and cancer in metazoan.

This article broadens our understanding of VSI, in a spirit that is somewhat similar to several recent contributions (e.g., Bouchard 2014; Doolittle 2014; Doolittle and Inkpen 2018; Garson 2017; O'Malley 2014; Toman and Flger 2017).¹ I then extend this broadened view of VSI to its “natural” starting point: the origin of the First Universal Cell Ancestor (FUCA).² Indeed, without broadening and then extending VSI, it may be difficult to arrive at a reliable understanding about the origin of FUCA and many key puzzles thereafter (cf. de Duve 2005a; 2005b; Fry 2011). The key reason is that the VSI operates quite differently in the pre-biotic and pre-cellular world (Koonin 2014a; Egel 2017), if the coming of the Last Universal Cell Ancestor (LUCA) marks the crossing of the “Darwinian Threshold” (Woese 2002).³ This broadened perspective of VSI therefore is NOT a rejection of Darwinian evolution. Rather, it is a necessary extension of it to the phase of pre-cellular and pre-Darwinian evolution.

I advance three principal arguments.

First, survival comes before replication. Before the coming of reproducer and replicator,

¹ The extension advanced here differs critically from other extensions on two fronts. First, they are more theoretical and philosophical than empirical. Second, with the exception of O'Malley (2014), they seek to account for evolution at the community or even ecosystem level. In contrast, the extension advanced here is squarely about the evolution of individual entities, from molecules, to complexes and protocells.

² This notation is similar to the First Eukaryotic Common Ancestor (FECA) before the Last Eukaryotic Common Ancestor (LECA). Just as both FECA and LECA were most likely a pool of organisms (Eme et al 2017), FUCA and LUCA were also most likely a pool of organisms. Notably, “universal ancestors” in Woese (1998) were what Woese and Fox (1977) called “progenotes” earlier or FUCA here. For similar interpretations, see Koonin 2014a; Gogarten and Deamer 2016; Egel 2017.

³ In fact, despite the legendary differences between Carl Woese and Lynn Margulis, they agreed that “speciation itself is a product of evolution” (Margulis and Sagan 2002, 143), or that only after the crossing of the “Darwinian Threshold” did speciation become possible (Woese 2002, 8744).

there must be survivors, to paraphrase Szathmary and Maynard Smith (1997).

Second, natural selection can operate without replication or even metabolism, as long as different molecules, complexes, and vesicles have differential survival rate within a system. Therefore, natural selection, a non-Darwinian kind admittedly, had operated during the prebiotic phase and long before the crossing of the “Darwinian Threshold”. In fact, Darwinian selection had been a product of the non-Darwinian phase (Woese 2002; Damer and Deamer 2015; cf. Tessera 2018).

Third, merger and acquisition, via breaking-and-re-encapsulation, endocytosis, endosymbiosis, and processes similar to them, had been a far more powerful force of variation and selection in the pre-Darwinian period of evolution that led to LUCA and long before eukaryogenesis (cf. de Duve 2005b). Endosymbiosis therefore had been a far more foundational force in evolution than even Lynn Margulis and many of her supporters had appreciated (e.g., Sagan 1967; Margulis 1981; 1991; Margulis and Sagan 2002; Norris and Raine 1998; O’Malley 2014). Our thesis thus goes beyond Woese’s (1998; 2002) emphasis of horizontal gene transfer (HGT) and actually subsumes HGT with endosymbiosis (see also Woese and Fox 1977; Vetsigian et al 2006).⁴

As I shall attempt to show below, combing these three new perspectives with other perspectives and evidence sheds important new light upon the origin of FUCA, the singular water-shedding moment in the evolution of life.

Of course, the origin(s) of FUCA may well be one of those puzzles that we simply cannot stop thinking about but can never come up with a definitive answer (Luisi 2016; Sutherland 2017). This article does not pretend to provide a definitive answer either. Rather, the purpose is to show that taking a different starting point will allow us to arrive at a more valid interpretation of existing data and evidence. More importantly, this new perspective may allow us to resolve some apparent controversies and point to more fruitful inquiries.

Three criteria should be applied to any new perspective regarding the origin of cell. First, it should provide a more coherent and consistent organization of existing data with fewer ad hoc hypotheses than existing ones, because it integrates new or previously underappreciated

⁴ Interestingly, starting with a quite different perspective, Woese and Fox (1977, 5) 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”.

data (Lanier and Williams 2017). Second, it should resolve more controversies with fewer ad hoc hypotheses than existing ones (Pereto 2005; Fry 2011). As becomes clear below, the new perspective resolves almost all major contradictions and inconsistencies within Carl Woese's theory regarding the same puzzle, as advanced in his millennium series (esp. Woese 1998; 2002). Finally and specifically for the origin(s) of FUCA, because we cannot establish the exact timing of the major events leading to FUCA, we should focus more on the sequence of events. Establishing a more valid sequence of events within this extraordinary process while maintaining consistency with thermochemical, geological, and biological principles provides us with a crucial baseline for evaluating competing theories and designing future research.

Seven Premises

I first outline seven simple premises. Except the last two, the other five premises contain key innovative statements. As becomes clear in the next section, these premises lead us to some new hypotheses for understanding the origin of the first cell.

1. Survival comes before replication, and certainly before division or reproduction (with or without genetic replication): before the coming of reproducer and replicator, there must be survivors.⁵ This law that survival or persistence comes before replication and reproduction holds most forcefully during the phase of prebiotic evolution. In short, an entity, be it a compound, a complex, or a vesicle, has to exist and then stay (or "survive") within the (pre-)biotic system before it can become part of life, even if it cannot metabolize or replicate (Pascal and Pross 2016; Toman and Flger 2017). In other words, survival, metabolism, and replication were not coupled for much of the early pre-biotic evolution and their coupling came as a product of prebiotic evolution. Indeed, even after the coming of the first protocell, an organism is a survival machine first and foremost and a replication and reproduction machine second.⁶
2. Before the first cell (and perhaps well after), variation had been the primary means toward survival (and replication). In other words, from the very beginning of bioorganic evolution

⁵ Clearly, a cell can survive for a while without replication.

⁶ As I shall show elsewhere, this principle holds critical implications for understanding other key puzzles, from an organism's lifespan, to the lifespan of cell lineages, cancer, apoptosis, and reproduction.

that eventually led to FUCA and LUCA, the evolutionary process was mostly about gaining more molecular and hence functional diversities so that FUCA or LUCA could survive in more diverse environment with a more potent arsenal. Moreover, variation back then was not generate by genetic mutation (which did not exist for a long time), but by two rather different processes: (1) prebiotic chemical synthesis, polymerization, and stereochemical mutualism; (2) absorption, via breaking-and-re-encapsulation, primitive endocytosis (e.g., simple engulfing), primitive endosymbiosis, and processes similar to them. Notably, endosymbiosis (as fusion or merger) and endocytosis (as acquisition) entail extensive “horizontal biomolecule transfer” (HBMT) than merely HGT does: HBMT subsumes HGT in the strict sense. In fact, Woese’s (1998; 2002) emphasis of HGT in the origin of FUCAs is valid only if he meant HBMT with HGT. HBMT was therefore the more pivotal and pervasive process than HGT.

3. A period of non-Darwinian but natural selection operated before the coming of Darwinian selection: the former produced the latter (cf. Tessera 2018). There were four major non-Darwinian selection mechanisms, which most likely had appeared in chronic order.
 - a) The first non-Darwinian selection mechanism, which is purely chemical, operates upon biomolecules themselves and their capacities for forming polymers and complexes. Here, the yardsticks of “fitness” include steady supply from abiotic synthesis (i.e., availability), kinetic and thermochemical stability or persistence (Pascal and Pross 2016; Toman and Flegr 2017), (water) solubility, polymerization, and stereochemical “mutualism” for forming complexes (Lanier, Petrov, and Williams 2017; Vitas and Dobovišek 2018).
 - b) The second non-Darwinian selection mechanism, which is both chemical and physical, selects biomolecules that can not only stabilize vesicles but also make vesicles somewhat more permeable (e.g., Wei and Pohrille 2011; Black et al 2013).
 - c) The third non-Darwinian selection mechanism operates upon the different capacities of different bioorganic complexes to interact with each other and the different capacities of different vesicles to merge and fuse with or engulf other complexes and vesicles (Damer and Deamer 2015).
 - d) The fourth non-Darwinian selection mechanism operates upon vesicles that now

approach protocells. Already stabilized (or intact) vesicles that can absorb useful ingredients or components from the environment via endosymbiosis or endocytosis and then invent primitive metabolism and replication will hold critical selection advantage. Here, the key yardstick of “fitness” was survival, absorption, growth, and division, first without and then with primitive metabolism and genetic replication (Norris and Raine 1998; Szostak et al 2001; Zhu et al 2012).

4. FUCAs came to exist via breaking-up and then re-encapsulating of vesicles and via endosymbiosis and endocytosis by vesicles, drawing useful ingredients or components from “global invention” (Woese 2002). *Put it differently, FUCAs did not come to exist via de novo evolution within individual protocells: 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 this underpinned by merger and fusion, rather than HGT alone, that FUCAs came to possess a proto-machinery of survival and a proto-machinery of replication within the same protocell.
5. Once FUCAs came to possess both a proto-machinery of survival and a proto-machinery of replication (both machineries of course require some kind of metabolism), survival and replication began to co-evolve with each other. Along the way, FUCAs continued to absorb useful ingredients and synthesize them into more complex, versatile, and effective macromolecules, including more complex proteins and RNAs. For this phase, however, a tight coupling of survival and replication might not hold any selective advantage. Indeed, the opposite might have been true: being more promiscuous and hence more flexible provides a protocell with significant advantage for survival. It is mostly because of this key dynamics that FUCAs did not have a genealogical history, but only a physical-chemical one (Woese 1998; 2002).
6. FUCAs competed against each other. After a period of time during which survival and replication co-evolved with each other, some of the FUCAs eventually became protocells in which survival and replication are more tightly coupled and smoothly regulated. Proto-cellular division (of vesicle) with tightly coupled genetic replication therefore came rather late in the evolutionary process that led to the first cells (Rasmussen et al, 2009,

xv-xvi). Once evolved, however, these protocells with division and genetic replication already being tightly coupled would come to enjoy an enormous advantage, and it was these protocells that eventually became the LUCA that came to move into diverse environment and eventually colonize the whole biosphere.

7. The evolution of the first proto-cell may well require a sustained increase in its capacity of harnessing and generating energy (Lane et al, 2015).

Core Hypotheses

I now derive the key hypotheses from the seven premises. I divide the hypotheses into two parts: before and after FUCAs. Obviously, hypotheses in the second part draw more from existing discussions, but even here, our new starting premises offer new perspectives. Figure 1 schematically summarizes the whole process from the origin(s) of FUCA to LUCA.

Evolution before FUCA

1. 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 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 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 availability partly depend on the relative easiness of synthesizing them from simple precursors and protection from UV light.
2. In order for simple bioorganic molecules to be assembled into more complex hetero-biomolecules, they also 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 that only some

amino acids or peptides can bind with simple RNAs or that only some peptides can form α -helices and then insert themselves into lipid membranes to make lipid membranes more permeable. Moreover, once some mutualism or stereochemical compatibility is fixed, it becomes extremely difficult to change or unravel, because such a change may be lethal. A key implication of stereochemical compatibility or 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 have been selected.

3. Self-organizing amphiphiles tend to form vesicles (Pohorille and Deamer 2009). Here, as long as a vesicle can retain its basic structure, float within a solution, absorb ingredients from its environment (e.g., via proto-endocytosis), and merge with other vesicles, it succeeds in surviving in the system. Most likely, such vesicles also had the capacity of “dividing” without either reproduction or genetic replication. Rather, they divided via pinching or budding due to enlargement of size by 1) absorbing more lipids, peptides, and other bioorganic molecules, 2) merging with and engulfing other vesicles, and 3) synthesizing new polymers within (Zhu and Szostak 2009; Budin and Szostak 2010; Budin et al 2012; Kurihara et al 2015; Armstrong et al 2018).⁷ Of course, such vesicles were also the substrate for endosymbiosis by other vesicles. *During this stage, survival and division was not coupled with active replication or even sophisticated metabolism.*
4. Although 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 very early on. The primitive translation apparatus and genetic code originated from this co-evolution of amino acid/peptide with RNA (Wong 1975; Wolf and Koonin 2007; Francis 2011; 2013; 2015; Sengupta and Higgs 2015; Koonin and Novozhilov 2017). During the co-evolution of amino acid/peptide with RNA, precision in RNA replication is not necessarily an advantage. Rather, during this stage of evolution, the key is to make more RNAs without precision because this increases the

⁷ Here, it is critical to differentiate “division” of liposomes due to surface chemical forces from division that depends on replication of genetic material, either RNA or DNA (e.g., Zhu and Szostak 2009).

structural diversity and hence the functional diversity of RNAs (Horning and Joyce 2017). With more diverse structures and functions, RNAs will be able to support the production of more diverse peptides with different properties. This mutually reinforcing increase in structure and function leads to peptides and proteins with more diverse structures and functions and paves the way for the more complex ribonucleoprotein (RNP) world, the standard genetic code (SGC), and eventually a more versatile metabolism system.

5. For a period of time in the prebiotic evolution, the evolution of peptide-lipid bi-layered membrane and the evolution of RNA-peptide (as the proto-translation machinery) might have proceeded independently from each other. Indeed, the two processes might have operated in different locations (e.g., different terrestrial hydrothermal ponds/fields) or even one came earlier than the other (Mulkiđjanian et al 2012; Damer and Deamer 2015; Koonin 2014b, 35-36). Eventually, however, these two processes had to come together to make the first proto-cell, and the moment in which these two processes merged with each other was the watershed event: it laid the foundation of the first proto-cell or FUCA (Schrum et al 2010; Pressman et al 2015; Cantine and Fournier 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 surface of the peptide-lipid vesicle.
6. FUCAs continued to absorb other fellow FUCAs, vesicles, RNAs, peptides, and other biomolecules from its external environment. Indeed, different FUCAs must have competed against each other for these ingredients in order to divide and populate within the system, as a form of non-Darwinian selection process (Chen et al 2004; Wei and Porhille 2011; Zhu et al 2012; Adamala and Szostak 2013; Black et al 2013; Kurihara et al 2015; Armstrong et al 2018). As a result, co-evolution of these molecules within FUCAs accelerated (Norris and Raine 1998; Wilson et al 2014; Saha et al 2018). Most critically, within FUCAs, metabolism, survival, division, and replication of genetic material came to be coupled with each other. The coupling was loose, however: FUCAs had yet to cross the Darwinian Threshold (Koonin 2014a; Egel 2017; cf. Woese 1998; 2002).
7. Within a stable microenvironment provided by FUCA's protocell 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 FUCAs, therefore, metabolism came to support the synthesis of more diverse and elaborate proteins and RNAs that could reinforce metabolism within a better regulated membrane system. Together, they came to form more sophisticated machineries of replication and survival and allow FUCAs to evolve toward LUCAs. Replication, or more precisely, fairly accurate replication of genetic materials (either RNA or DNA) supported by a replication apparatus could have only evolved within a protocell such as FUCA. Most likely, FUCA initially contained only short RNAs rather than a single long RNA molecule.

From FUCA to LUCA

Once the survival machinery (i.e., metabolism supported by proteins within a membrane) and the replication machinery (now supported by both proteins and RNAs) were coupled with each other within a protocell, they began to co-evolve with each other. This co-evolutionary process within FUCA laid the foundation for all subsequent evolutionary processes.

1. FUCAs were a pool of “progenotes”, and they struggled for existence. Eventually, one or few of the FUCAs came to possess the right and tight coupling of metabolism, translation machinery, division with genetic replication, and energy efficiency. These luck few became the LUCAs: their progenies came to dominate the whole system. Because LUCAs possessed a relatively tight coupling of cell division with genetic replication, they had crossed the Darwinian Threshold (Woese 1998; 2002). By LUCAs, a translation system with the full SGC had also “crystallized” (Woese 1998; 2002; Koonin 2014a). More likely than not, this whole process became only possible within protocells with a regulated membrane rather than directly from the “naked” RNA world (Pohorille and Deamer 2009; Lombard et al 2012; see also Koonin 2014a).
2. Due to some kind of geological ruptures or accidents (e.g., overflowing of a hydrothermal field or pond), some LUCAs were transferred to a new environment. This event started the making of bacteria and Archaea as the two foundational domains, including their metabolism, lipid membrane (i.e., from heterochiral to homochiral, see below), DNA to

RNA transcription, and DNA replication etc. (Koonin 2014b; Lombard et al 2012).

3. During the FUCA to LUCA period, a transition from RNA to DNA via RNA-DNA hybrid had begun. LUCA, however, began to diffuse into two different environments before the transition from a RNA-DNA hybrid system to a full DNA genetic system was complete. As a result, DNA replication evolved twice independently in the two primary domains after their divergence (Leipe et al 1999). The whole transition from RNA to DNA as genetic material therefore had two phases: a phase of RNA-DNA hybrid and a phase of full transition to a DNA-only system. The former lasted from FUCA to LUCA and beyond, even if some components of a primitive DNA-replication machinery were already in place within FUCA (i.e., before LUCA). The latter started after the divergence of LUCA into bacteria and Archaea and continued well after.
4. Only in LUCA 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 (Margulis 1981). Moreover, only in LUCA did HGT gradually become more harmful (Jain et al 1999), and a defense system against HGT came to exist (Koonin 2017; see also Koonin and Dajila 2013). In other words, the arms race between HGT and defense against genetic invasions by genetic parasites (or replicons) via HGT most likely began only after FUCA. Before LUCA, every replicon had their chance of becoming part of LUCA's genetic core. Some inevitably got lucky, and some inevitably did not (Koonin et al 2017). Once genetic parasites came to exist, however, they began to play a critical role in driving the evolution of their hosts (Koonin and Dajila 2013; Koonin 2016)

Evidences: Chemical and Biological-Functional

Because FUCAs were a community of proto-cells and the long evolutionary process after FUCAs with rampant HGT had erased or at least obscured most of the early genetic footprints (Fournier et al 2015), empirical evidence for any theory regarding the origin of the cell must be mostly chemical and biological-functional rather than genetic.⁸ This section presents the

⁸ 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 (Ranea et al 2006).

chemical and biological-functional evidence that supports our core hypotheses, in addition to some genetic evidence. The next section goes on to show that our new theory resolves some key controversies regarding the origin of life, thus providing another source of support.

1. The universality of the standard genetic code (SGC) is indisputable. As part of the core translation apparatus, SGC's indisputable universality in only one out of 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) can only be explained by the coming of amino acid/peptide-RNA interaction very early on, long before the coming of DNA replication and DNA to RNA transcription. The possibility that the proto-translation machinery came to operate via initial chemical mutualism between amino acid/peptide with RNA and then peptide/protein and RNA co-evolved thereafter is now almost universally accepted (e.g., Wong 1975; 1981; Petrov et al 2015; Kovacs et al 2017; Yarus 2017). Moreover, almost all the one hundred or so universally conserved genes were within the translation apparatus (Koonin 2003; Puigbo et al 2009; Koonin 2014a). This fact suggests that a RNA-amino acid/peptide world came to exist very early on, even though RNAs can self-replicate and possess other catalytic activities (Kovacs et al 2017; Saad 2018).
 - a) The fact that many ancient proteins are connected to nucleotides and RNAs also suggests an early rather than a late RNA-peptide world (Ma et al 2008).
 - b) This possibility is further strengthened by the fact that small and simple peptides derived from the core part of ribosome proteins can enhance RNA polymerase ribozyme (RPR) catalytic activities, with the smallest size being merely ten amino acids (Tagami et al 2017). As such, a primitive ribosome-like apparatus with both RNAs and peptides must have come to exist quite early on (Wolf and Koonin 2007).
 - c) 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. Indeed, viroids so far have only been found in higher

plants whereas viruses have been found in all three domains.⁹ This fact is also more consistent with the notion that nuclear acids (mostly likely, RNA) came to interact with peptides or proteins quite early on.

2. The fact that a primitive translation system and eventually the whole SGC had evolved quite early on strongly suggests that vesicles with permeable membrane came to exist quite early on (Pohorille & Deamer 2009; Mulkidjanian et al 2009; Lombard et al 2012; Koonin 2014a). It also strongly suggests the possibility that these vesicles were able to merge with each other via proto-endosymbiosis. Simply put, it would have been extremely difficult for a single proto-cell or vesicle to evolve the whole SGC: the evolution of SGC most likely requires input from “global invention” (Woese 1998; 2002). In contrast, with HBMT, which is much more than acquiring a genome, vesicles within a pool of vesicles or protocells could have easily drawn from and brought together “global inventions”, not only in genetic materials but also in metabolism and other ingredients, within the community of progenotes. Obviously, HBMT assumes vesicles with proto-membrane that can retain molecules within whereas Woese’s HGT does not.
3. The possibility that HBMT via proto-endosymbiosis and proto-endocytosis has been a key mechanism in evolution of FUCA and LUCA, in addition to (and well before) eukaryogenesis, is strongly supported by several sets of evidence.
 - a) Many artificial vesicles indeed can grow and divide without replication, by simply absorbing either ingredients or other (mimi-)vesicles (Kurihara et al 2015; Saha and Chen 2014; see also Pohorille and Deamer 2009; and other contributions in Rasmussen et al 2009). They can also undergo structural changes under different conditions (e.g., different PHs, hydration-and-dehydration as the wet-and-dry cycle, freeze-and-thaw cycle), thus facilitating merger and acquisition (e.g., Deamer and Barchfield 1982; Qiao et al 2016).
 - b) The notion that mitochondria and chloroplast and other plastid-like structures (e.g., liposomes, hydrogenosomes, and mitosomes) had evolved as relics of endosymbiosis (i.e., its substrates and leftovers) has now been firmly established (Sagan 1967; Margulis 1981; Zimorski et al 2014; Martin et al 2014; *J. Theo. Biol.*, special issue

⁹ The fact that viroids have only been found in higher plants suggests that they have a more recent origin.

2017). This fact suggests that endosymbiosis and endocytosis had a very ancient origin and was more widespread in pre-cellular evolution (Woese and Fox 1977). If endosymbiosis and endocytosis could still operate after the origin of cell, shouldn't they be more prevalent before cell? The fact that bacterial and Archaea cells today can still conjugate and then exchange plasmids or other replicons indicates that these cells retain some of the key components for extensive endosymbiosis and endocytosis in ancient time.¹⁰

- c) LUCA might have been an already quite complex organism with several hundred genes (Doolittle 2000; Koonin 2003; Ranea et al 2006). This fact is more consistent with the possibility that LUCA came from several FUCA lineages via several steps of merger and acquisition than with the possibility that LUCA came to evolve from a single FUCA lineage all by itself.
 - d) RNA and lipid can interact with each other, thus making the fusion of RNA-peptide and the lipid-peptide vesicle quite likely. In fact, positive-strand RNA viruses can readily work with membranes during replication and infection by “hijacking” or recruiting membrane lipids (Miller and Krijinse-Locker 2008; van der Schaar et al 2016; Altan-Bonnet 2017).
4. In the evolutionary process leading to LUCA and beyond, HBMT thus has two distinct phases. The first phase, before the divergence of bacteria and Archaea from LUCA, was mostly, if not exclusively, via fusion and absorption based on endosymbiosis and endocytosis. The second phase, mostly in the form of replicons jumping from one organism to another, came after the “crystallization” of the two principal domains. HGT only dominated this second phase. Even in the second phase, however, endosymbiosis played a key role, especially in eukaryogenesis (Sagan 1967; Margulis and Sagan 2002).
- a) Woese's (1998; 2002; also Vetsigian et al 2006) vision that HGT was a critical force in the evolution of SGC does not have a viable mechanism other than HBMT as merger and acquisition (Koonin 2014a, 202). Put it simply, how can a RNA-peptide molecule without some sort of vesicle import and retain other RNA molecules? Once

¹⁰ In this sense, eukaryogenesis was a unique event only because it produced a eukaryotic cell but not because endosymbiosis only operated this once (Booth & Doolittle 2015; cf. Lane and Martin 2015).

we replace HGT with HBMT, which subsumes the former, we can resolve all the contradictions and inconsistencies in Woese's (1998; 2002) treatises on the origin of the cell (for an insightful review, see Koonin 2014a). Indeed, Woese (2002, 8744) came close to admit this possibility: "*Were that organization (of a progenote as a proto-cell) simple and modular enough, all of the componentry of a cell could potentially be horizontally displaced over all time.*" Quite evidently, such a possibility is not so compatible with HGT but entirely compatible with HBMT as the central mechanism leading to the origin of the cell.

- b) The massive amount of shared genetic material between Archaea and bacteria is more consistent with the possibility that in early evolution, HBMT via fusion and absorption had been a more powerful force than HGT via replicons jumping from one cell or organism to another.
 - c) HBMT subsumes Lynn Margulis and Dorion Sagan's (2002) notion of "acquiring a genome": HBMT is much more than "acquiring a genome".¹¹ There was a "genetic takeover" in the origin of cell (Cairns-Smith 1987), but it came via HBMT based on endosymbiosis and endocytosis rather than genetic invasion or HGT.
 - d) With the exception of megaviruses, most viruses and other replicons (e.g., plasmids) have a small genome. Although one can argue that all the replicons came to their small genome size via loss of genes, it is equally, if not more, plausible that replicons really were just small genetic fragments. Notably, as double-stranded DNA viruses, megaviruses must have originated after the origin of DNA replication. Hence, megaviruses do not prove that early replicons or protocells possessed big genomes. In short, HGT via replicons could not have brought together a large genome that could have sustained a complex life as LUCA or even FUCA.
5. Our thesis suggests that very early on, only some early replicons were selected to occupy the central genetic apparatus of the first cells whereas many were left out. Some of these

¹¹ Here, ironically, while Woese was mostly genetic-centric and hence vehemently against the mechanism of endosymbiosis regarding the origin of the cell, Lynn Margulis was also genetic-centric here despite her insistence of endosymbiosis for eukaryogenesis and the evolution after that great event. Woese was against endosymbiosis because he wanted to preserve his original tree of life based on 16s ribosome RNA, which stakes a three primary domains of life, whereas Margulis' theory suggests only two primary domains of life. Recent data strongly suggests only two primary domains of life, with eukaryotes arose from the merger of bacteria and Archaea (Lombard et al 2012; Williams et al 2015; Koonin 2014a; Eme et al 2017).

left-out replicons then came to exist as viruses. In other words, genetic parasites had been the inevitable result of a selection process (Koonin et al 2017). A possible explanation behind this early divergence of fate among replicons might have been that only some replicons were helpful for the protocells whereas many others were not or even harmful. In other words, viruses might have been selected out of the first protocells precisely because they do not carry fitness-enhancing functions for the first cells. Only this logic can in turn explain the presence of viral defenses across all three domains quite early on: one only needs to defend against potential harm (Koonin 2017).

- a) Such defense system must come after the crystallization of a proto-genome, and most likely, after the coming of FUCA, if not LUCA. Moreover, such a defense system implies the presence of cell membrane, again supporting the hypothesis that the proto-cell membrane was a critical dimension to the evolution of the first cell.
 - b) Some, if not most, positive single-stranded RNA viruses [(+)ssRNA viruses] might have originated from those RNA-peptide complexes or RNA molecules that were not integrated into FUCA's proto-genome due to their disruptive properties (Koonin and Dalja 2013).¹² This possibility best explains the possible "Big Bang" origins of (+)ssRNA viruses (Koonin et al 2008), the great diversity of this group of virus, and why (+)ssRNA viruses remain the largest group of viruses.
6. If survival comes before replication, then molecules that can support the survival of an organism are equally critical, if not more so, when comparing to components in the replication (and transcription) apparatuses. Besides the bilayered lipid membrane itself as a protective apparatus, some other protective mechanisms should have been selected for very early on. Put it differently, for any organism or proto-cell, some kind of stress-response machinery is necessary. This is most likely in the form a stress response mechanism and apparatus, or "heat shock response (HSR)" as known today. Very likely, HSR was an early invention as requisites for life on the edge, for the sake of survival.
- a) Major components within HSR, such as Hsp100, Hsp90, Hsp70, Hsp60, and other small heat-shock proteins (HSPs), are highly conserved across the three domains

¹² Unfortunately, because viruses and other replicons tend to evolve rather rapidly, their evolution informs us little about the evolution of FUCAs and LUCAs.

(Richter et al 2010). For instance, DnaK, the prokaryotic version of Hsp70, shares about 60% sequence identity with its eukaryotic counterpart. Likewise, Hsp60 is present in all three domains. Most likely, HSPs were perhaps some of the earliest proteins to be firmly integrated into FUCAs. Only this logic can explain why so many HSPs are highly conserved across all three domains with such high fidelity. The fact that mitochondria possess HSPs also suggests that stress response was critical for protecting this vital organelle. In short, survival demands FUCAs to possess a stress response system quite early on

- 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. Indeed, lipid membranes are stabilized and regulated by peptides and this mutualism between membrane lipids and membrane proteins is a key part of the foundation of survival. Several observations suggest that Hsp12 may be one of the molecular fossils from this period. In its soluble form, yeast Hsp12 is unfolded or unstructured (Richter et al 2010, 254). Yet, when it interacts with lipids, it becomes α -helical. More critically, Hsp12, when binding membranes, stabilizes membranes by decreasing membrane fluidity. In terms of sequence, 50% of Hsp12's sequence is made of five amino acids: Ala, Asp, Glu, Gly, and Lys. Remarkably, among the five amino acids, four of them (i.e., Ala, Asp, Glu, and Gly) belong to the "first amino acids". More importantly, both Ala and Gly tend to form transmembrane (TM) alpha-helices via Ala/Gly-X-X-X-Ala/Gly whereas Asp can stabilize α -helices by capping α -helix TM domain (Francis 2013). Ala/Gly-X-X-X-Ala/Gly sequences also facilitate dimerization. Together, these facts suggest that machineries that regulate cell's interaction with and response toward the environment came very early on, at least at FUCA. After all, a key aspect of the co-evolution of membrane and membrane proteins was about protecting the membrane (Mulkidjanian et al 2009). So far, no Hsp12 homolog has been discovered beyond yeast and fungus, but this may be due to the fact that membrane proteins are less conserved than water-solvable proteins (Sojo et al 2016).

- c) The fact that molecules that ensure survival came rather early is also consistent with the possibility that life most likely began in a stressful environment (Damer and Deamer 2015; Lane et al 2010; Mulkidjanian et al 2012), and life forms that can survive in other less-demanding environment (i.e., cooler) came later.
7. Ubiquitin (UB) is universally conserved in eukaryotes while ubiquitin-like proteins (UBLs) (including small ubiquitin-related modifiers SUMOs) 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 “small but versatile” beta-grasp fold that is common to UB, UBLs, and URM (Iyer et al 2006; Burroughs et al 2006). These proteins may be another set of proteins that were integrated into FUCAs (Staley and Fuerst 2017).
- a) UB, UBLs, URM, and SUMOs 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. The fact that deletion of Nedd-8, a UBL with high UB homologue (58% amino acid identity), is lethal in a variety of organisms from yeast to mice strongly suggests a critical value for survival (van der Veen and Ploegh 2012, 329-332).
- b) UB and UBL (in the Urm1 - Uba4 system, which is much closer to THiS and THiF) might have been a ‘molecular fossil’ from the more ancient sulfur-transfer pathway (Hochstrasser 2009, 425; van der Veen and Ploegh 2012, 342-3). Indeed, metal 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. The fact that UB and UBLs are small proteins also suggests an ancient origin.

Controversies Resolved, Partly and Possibly

It is highly unlikely that we can establish the exact timing of each major evolutionary event leading toward FUCAs. But we may be able to establish the order for some of the events. Once the order of these events is established, we can resolve some key controversies

and identify more fruitful directions for future research (for earlier reviews, see Pereto 2005; Glansdorff et al 2008; Fry 2011; Koonin 2014a; 2014b; Spitzer 2018; Cantine and Fouriner 2018). This section details the key controversies that our theory can help resolve. In contrast, some debates may never be confidently resolved and hence less useful.

1. The question which comes first, the RNA world (and then a RNA-peptide world) or the peptide-lipid membrane, is one of those unfruitful controversies because it will never be firmly resolved. Our thesis suggests that vesicles came to exist very early on and that vesicles or protocells had made subsequent evolutionary processes possible (Lombard et al 2012; Koonin 2014a; Deamer 2016; Cantine and Fournier 2018). Most critically, our thesis implies that the most decisive step was the merging of these two pathways that paved the way for the eventual coupling of survival, metabolism, and replication.
2. During the process leading to FUCAs and LUCAs, the process was mostly about *gain* in structure and function so that FUCAs and LUCAs could survive in more diverse environment with a more potent weaponry system. Our thesis also suggests that before FUCAs/LUCA, metabolic innovation was equally critical, if not more critical than genetic ones (e.g., Margulis 1981; O'Malley 2014). More likely than not, streamlining or simplification via loss of functions and genes came after the divergence of bacteria and Archaea from LUCA, when “evolutionary temperature” had cooled down considerably (Woese 1998). As de Duve (2005b, 163, fn. 3) put it pithily, “There can be no reduction without prior ‘complexification.’” Two additional facts suggest that streamlining or simplification via loss of functions and genes came after the divergence of bacteria and Archaea, including 1) SGC had evolved in two phases, an early phase with early amino acids and a later phase with late amino acids (Koonin 2010; Francis 2011; 2013; 2015), and 2) many proteins in LUCA performed multiple functions (Ranea et al 2006).
3. The possibility that streamlining or simplification via loss of functions and genes came after the divergence of bacteria and Archaea from LUCA goes against the thesis that the FUCA or LUCA was a proto-eukaryotic cell, and bacteria, Archaea, and Eukaryotes came to exist via genetic reduction or loss of function in proteins and genes (e.g., Glansdorff et al 2008; see also Forterre and Philippe 1999; Philippe and Forterre 1999). Recent

advances also strongly suggest a Tree of Life (ToL) with bacteria and Archaea as the two primary domains (Williams et al 2003; López-García & Moreira 2015; Spang et al 2017; Dacks et al 2016; Eme et al 2017). By all likelihood, the hypothesis that FUCA or LUCA was a proto-eukaryotic cell can now be confidently ruled out.

4. Intuitively, when survival comes first and merger and acquisition, vesicles as proto-cells could not afford to be too picky. Indeed, some degree of heterogeneity was perhaps crucial for the evolutionary process (Szostak 2011). When this is the case, FUCAs and LUCAs were most likely to have peptide-lipid proto-membranes that were heterochiral with both isoprenoid-based and fatty-acid-based phospholipids (Wächtershäuser 2003; Pereto et al 2005; Lombard et al 2012; Shimada and Yamagishi 2011). Homochiral membranes came only after the divergence of bacteria and Archaea from LUCA.
5. The notion that FUCAs came together through HBT based on endosymbiosis and endocytosis that brings together different vesicles containing different components suggests a heterotrophic origin of FUCA and LUCA. Autotrophic life was only achieved after a long period of heterotrophic evolution (Oparin 1953; Fry 2011; Damer and Deamer 2015). The leading autotrophic hypothesis that LUCA emerged from “hydrothermal alkaline vents (HTAV)” as an autotrophic protocell within the chambers of HTAV and then escaped from the vent has numerous difficulties (Mulkiđjanian et al 2012; Gogarten and Deamer 2016; Jackson 2016; 2017; cf. Martin et al 2008; Lane et al 2010; Lane 2017; Weiss et al 2016). Indeed, by insisting that all the good things for LUCA must be present in one place (i.e., HTAV), the HTAV hypothesis approaches a theory banking on miracles.
6. 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 a marine one (Mulkiđjanian et al 2012; Damer and Deamer 2015; cf. Weiss et al 2016). Most critically, TAGTFs allow the “wet-and-dry” cycle, which in turns drive vesicles through the cycle of breaking up (old vesicles) and then re-forming (new vesicles) by changing the concentration of ions and other ingredients or the overall physical micro-environment within the “warm little pond” (Deamer and Barchfeld 1982; Budin et al 2012; Damer and Deamer 2015; Zhu and Szostak 2009). Along the way, new and old biomolecules will be repacked and recombined within new vesicles, essentially working as a physical and

chemical process of recombination. This process of breaking-up and [re]encapsulation thus overcomes part of the physical barrier against HBMT imposed by vesicles (Damer and Deamer 2015). The fact that quite a few pathways can propel this environmentally induced vesicle recombination (e.g., concentration, wet-and-dry, freeze-and-thaw, or even redox) and that different vesicles containing different biomolecules have different capacities of growth (via absorption and in-taking) and division strongly suggests that such pathways might have been powerful forces of variation and then selection in the evolution of FUCAs (Chen and Szostak 2004; Damer and Deamer 2015; Zhu and Szostak 2009; Budin and Szostak 2011; Budin et al 2012; Qiao et al 2016). Again, this possibility supports the “TAGTF” thesis while going against the “HTAV” thesis (see also Da Silva et al 2015; Pearce et al 2017; Milshteyn et al 2018 for additional evidence).

7. If life did originate in TAGTFs, then a geological rupture or an inundation induced by a high tide could have washed FUCAs from a “warm little pond” (or a geothermal field) to the ocean and ignited the process of life eventually conquering the whole biosphere (Mulkidjanian et al 2012; Damer and Deamer 2015).
8. Our interpretation questions the possibility that DNA replication came from DNA virus rather than FUCA/LUCA with some kind of rudimentary DNA replication machinery (Forterre 2006). 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. Hence, DNA replication evolved twice, once in bacteria and once in Archaea (Leipe et al 1999).
9. Without a (primitive) defense system (e.g., the CRISP defense system in bacteria), primitive cells are extremely vulnerable to genetic invasion. In contrast, acquiring an even primitive defense system against genetic invasion inevitably reduces the rate of HGT. The acquiring of a defense system against genetic invasion, most likely after FUCAs, therefore marks the coming of HGT as a reduced form of HBMT.
10. Because cell division via pinching or budding in the beginning could not have been very precise, asymmetrical cell division (ACD) was perhaps an inevitable outcome. Moreover, ACD might have been selected very early on because it provides one of the daughter cells with a higher probability of survival. Such an arrangement holds important advantage

over the scenario of having two daughter cells with the same probability of survival.

Concluding Remarks: Unsolved Mysteries and Ways Forward

This article advances some radical new hypotheses regarding the origin of the first cells that seems to be consistent with much of the evidence and help resolve some key controversies. My emphasis of merger and acquisition via endosymbiosis and endocytosis does not challenge the central mechanism of VSI but rather constituting an extension of VSI. I argue that endosymbiosis and endocytosis represent a powerful force for generating variations that were not limited to genetic, but also metabolic and perhaps structural. Moreover, before LUCAs, metabolic and structural innovations were more critical than genetic ones.

Our thesis thus calls for a more foundational extension of the Neo-Darwinian Modern Synthesis. For the Modern Synthesis, variation can only be genetic and mostly within cells. Yet, before LUCAs, variation can be both non-genetic (e.g., metabolic and structural) and imported. Darwinian evolution itself had been a product of pre-Darwinian/cellular evolution. Moreover, endosymbiosis and endocytosis can be understood as a force of selection. The profound implication for evolution of endosymbiosis and endocytosis therefore must be more appreciated (O'Malley 2014; cf. Margulis 1981; 1991).

Of course, we still leave many questions, from the divergence of the two primary domains, to the origin of the eukaryotes, and the evolution of the individual phylum unaddressed. Nonetheless, we can now decipher several exciting directions forward.

First, some interesting, if not entirely decisive, experiments can be conceived. Among others, three will be of particular interest: 1) Artificial vesicles can perform endosymbiosis; 2) Artificial vesicles can indeed perform endocytosis by absorbing not only small molecules (e.g., amino acids) but also engulfing bioorganic complexes (e.g., an RNA-peptide complex); and 3) Artificial vesicles can not only absorb but also exclude RNAs.

Second, within a population of artificial vesicles aided by a steady supply of amino acids and nucleotides and environmental changes triggered by a wet-and-dry cycle, a freeze-and-thaw cycle, or some other physical and chemical changes, a rudimentary genetic code can indeed evolve. Such a rudimentary genetic code may be some

probabilistic association between some amino acids and some short (oligo-)nucleotides. Experiments along this direction may constitute powerful evidence that the model outlined here (Fig. 1) might have indeed operated in the origin(s) of LUCA with the standard genetic code.

Third, structural phylogenomic analysis with HSPs that are responsible for protecting the membrane (e.g., Hsp12) being of special interest. Other useful molecular markers may include UBLs and UBMLs. The possibility that beta-grasp fold was a RNA-binding domain with connections to RNA metabolism also suggests that its origin has been ancient. Of course, molecular phylogeny analysis based on sequence will be of increasingly limited value as we reach further and further back into the origin of life, even if we get the techniques of such analyses right (Shen et al 2017). A better way forward will be searching for structural homologies. But even this technique has its limits.

Declaration

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The author declares that he has no competing interests.

Funding

The search has not received any funding.

Availability of data and materials

Not applicable.

Authors' contributions

ST developed the theory and wrote the paper.

Acknowledgements

For insightful comments and suggestions, I thank Bruce Damer, David Deamer, and Eugene Koonin. Jin He and Chang-an Liu provided outstanding research assistance.

Author details

School of International Relations and Public Affairs (SIRPA), Fudan University, Shanghai, China. twukong@fudan.edu.cn

Abbreviations

HBMT: horizontal biomolecules transfer; HGT: horizontal gene transfer; HTAV: hydrothermal alkaline vent; TAGTF: terrestrial anoxic geothermal fields

References

- Adamala, K., Szostak, J.W., 2013. Competition between model protocells driven by an encapsulated catalyst. *Nat. Chem.* 5, 495-501, <http://dx.doi.org/10.1038/nchem.1650>.
- Altan-Bonnet, N., 2017. Lipid tales of viral replication and transmission. *Trends in Cell Biol.* 27 (3), 201-213, <http://dx.doi.org/10.1016/j.tcb.2016.09.011>.
- Armstrong, D. L., 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. *Astrobiol.* 18 (4), 419-430, <http://dx.doi.org/10.1089/ast.2016.1615>.
- Black, R. A., Blosser, M. C., Stottrup, B. L., Tavakley, R., Deamer, D. W., Keller, S. L., 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 (33) 13272-13276, DOI: [10.1073/pnas.1300963110](https://doi.org/10.1073/pnas.1300963110)
- Booth, A., Doolittle, W.F., 2015. Eukaryogenesis, how special really? *Proc. Natl. Acad. Sci. USA.* 112 (33), 10278–10285.
- Bouchard, F., 2014. Ecosystem evolution is about variation and persistence, not populations and reproduction. *Bio. Theo.* 9, 382–91.
- Budin, I., Debnath, A., Szostak, J.W., 2012. Concentration-driven growth of model protocell membranes. *J. Am. Chem. Soc.* 134 (51): 20812–20819
- Budin, I., Szostak, J.W., 2010. Physical effects underlying the transition from primitive to modern cell membranes. *Proc. Natl. Acad. Sci. USA.* 108, 5249–5254.
- Burroughs, M.A. et al. Small but versatile: the extraordinary functional and structural diversity of the β -grasp fold, *Bio. Direct* 2, 18, <http://dx.doi.org/10.1186/1745-6150-2-18>.
- Cairns-Smith, A.G., 1987. *Genetic Takeover*. Cambridge: Cambridge University Press.
- Cantine, M.D., Fournier, G P., 2018. Environmental adaptation from the origin of life to the last universal common ancestor. *Orig. Life. Evol. Biosph.*, 48, 35–54, <http://dx.doi.org/10.1007/s11084-017-9542-5>.
- Chen, I.A., Roberts, R.W., Szostak, J.W., 2004. The emergence of competition between model protocells. *Science*, 305, 1474–1476.
- Da Silva, L., Maurel, M, C., Deamer, D., 2015. Salt-promoted synthesis of RNA-like

molecules in simulated hydrothermal conditions. *J Mol Evol* 80: 86–97. doi: 10.1007/s00239-014-9661-9.

Dacks, J.B. et al., 2016. The changing view of eukaryogenesis — fossils, cells, lineages and how they all come together. *Cell Sci.* 129, 3695–3703.

Damer, B., 2016. A Field Trip to the Archaean in Search of Darwin’s Warm Little Pond. *Life* 2016, 6, 21; doi:10.3390/life6020021

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.

de Duve, Christian. 2005a. The onset of natural selection. *Nature* 433: 581-582.

de Duve, Christian. 2005b. Singularities: landmarks on the pathways of life. Cambridge: Cambridge University Press.

Deamer, D. W., Barchfeld G. L. 1982. Encapsulation of macromolecules by lipid vesicles under simulated prebiotic conditions. *J Mol Evol* 18 (3): 203-6.

Doolittle, W.F., 2000. The nature of the universal ancestor and the evolution of the proteome. *Curr Opin Struct Biol* 10, 355–358

Doolittle, W.F., 2014. Natural selection through survival alone, and the possibility of Gaia. *Biol Phil*, 29 (3): 415-23.

Doolittle, W.F., Inkpen S.A., 2018. Processes and patterns of interaction as units of selection: an introduction to ITSNTS thinking, *Proc Natl Acad Sci USA*. 115 (16), 4006-4014. DOI: 10.1073/pnas.1722232115

Ely, R., 2017. ‘Parabiogenic Evolution’: from stochasticity in geochemical and subsequent processes to genes, genomes and modular cells. Preprints (www.preprints.org), <http://dx.doi.org/10.20944/preprints201710.0153.v1>.

Eme, L., Spang, A., Lombard, J., Stairs, C. W., Ettema, T J. G. 2017. Archaea and the origin of eukaryotes. *Nat. Rev. Microbio.*, 15, 711-723.

Forterre P, Philippe H, 1999. Where is the root of the universal tree of life? *Bioessays*, 21, 871-879.

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

(10), 3669-2674. DOI: www.pnas.org/cgi/doi/10.1073/pnas.0510333103

Fournier, G. P., Andam, C. P., Gogarten, J. P., 2015. Ancient horizontal gene transfer and the last common ancestors. *BMC Evo. Bio*, 15, 70.

Francis, B. R. 2011. An alternative to the RNA world hypothesis. *Trends in Evol. Biol.*, doi:10.4081/eb.2011.e2

Francis, B. R. 2013. Evolution of the Genetic Code by Incorporation of Amino Acids that Improved or Changed Protein Function. *J Mol Evol* (2013) 77:134–158, DOI 10.1007/s00239-013-9567-y

Francis, B. R. 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, doi:10.3390/life5010467

Fry, I., 2011. The Role of Natural Selection in the Origin of Life. *Orig Life Evol Biosph* 41, 3–16, <http://dx.doi.org/10.1007/s11084-010-9214-1>.

Garson, J., 2017. A Generalized Selected Effects Theory of Function, *Phil. Sci.*84 (2), 523–543.

Glansdorff, N., Xu, Y., Labedan, B. 2008. The Last Universal Common Ancestor: emergence, constitution and genetic legacy of an elusive forerunner. *Biol. Direct*, 2008, 3, 29 <http://dx.doi.org/10.1186/1745-6150-3-29>.

Hochstrasser, M. 2009. Origin and function of ubiquitin-like proteins. *Nature* 458, 422-29.

Horning, D. P., Joyce, G. F., 2017, Amplification of RNA by an RNA polymerase ribozyme. *Proc. Natl. Acad. Sci. USA* 113 (35), 9786–9791.

Iyer, L.M., Burroughs, A.M., Aravind L., 2006. The prokaryotic antecedents of the ubiquitin-signaling system and the early evolution of ubiquitin-like β -grasp domains. *Gen. Bio.* 2006, 7, R60, <http://dx.doi.org/10.1186/gb-2006-7-7-r60>.

Jackson, J.B., 2017. Ancient living organisms escaping from, or imprisoned in, the vents? *Life*, 7, 36, <http://dx.doi.org/10.3390/life7030036>.

Jain, R., Rivera, M.C., Lake, J.A. 1999. Horizontal gene transfer among genomes: the complexity hypothesis. *Proc. Natl. Acad. Sci. USA* 96: 3801-6; <http://dx.doi.org/10.1073/pnas.96.7.3801>.

Johann Peter Gogarten, David Deamer, 2016, Is LUCA a thermophilic progenote. *Nat. Micro*

1, 16229

Joyce, G. F., 2002. The antiquity of RNA-based evolution. *Nature* 418, 214–221.

Koonin, E. V., Wolf, Y. I., Nagasaki, K., Dolja, V. V. 2008. The Big Bang of picorna-like virus evolution antedates the radiation of eukaryotic supergroups. *Nat. Rev. Microbiol.* 6, 925-939. doi:10.1038/nrmicro2030

Koonin, E.V., 2003. Comparative genomics, minimal gene-sets and the last universal common ancestor. *Nat Rev Microbiol* 1, 127–136.

Koonin, E.V., 2014a. Carl Woese's vision of cellular evolution. *RNA Biology* 11 (3), 197-204.

Koonin, E.V., 2014b. The origin of cellular life. *Antonie van Leeuwenhoek* 106, 27-41.

Koonin, E.V., 2016. Viruses and mobile elements as drivers of evolutionary transitions. *Phil. Trans. R. Soc. B* 371: 20150442. <http://dx.doi.org/10.1098/rstb.2015.0442>

Koonin, E.V., 2017. Evolution of RNA- and DNA-guided antiviral defense systems in prokaryotes and eukaryotes: common ancestry vs convergence. *Bio. Direct.* 12, 5, <http://dx.doi.org/10.1186/s13062-017-0177-2>.

Koonin, E.V., Dolja, V.V., 2013. A virocentric perspective on the evolution of life. *Curr. Opin. Virol.* 3 (5), 546–557.

Koonin, E.V., Novozhilov, A.S., 2017. Origin and evolution of the universal genetic code. *Annu. Rev. Genet.* 27 (51), 45-62. <http://dx.doi.org/10.1146/annurev-genet-120116-024713>.

Koonin, E.V., Wolf, Y.I., 2012. Evolution of microbes and viruses: a paradigm shift in evolutionary biology? *Front Cell Infect Microbiol.* 2, 119, <http://dx.doi.org/10.3389/fcimb.2012.00119>.

Koonin, E.V., Wolf, Y.I., Katsnelson, M. I. 2017. Inevitability of the emergence and persistence of genetic parasites caused by the evolutionary instability of parasite-free states. *Biol Direct* 12 (31), DOI 10.1186/s13062-017-0202-5.

Kovacs, N.A., Petrov, A.S., Lanier, K.A., Williams, L.D., 2017. Frozen in time: the history of proteins. *Mol. Biol. Evol.* 34, 1252–1260.

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. Comm.* 6, <http://dx.doi.org/10.1038/ncomms9352>.

Lane, N., Allen, J. F., Martin W., 2010. How did LUCA make a living? Chemiosmosis in the

- origin of life. *Bio Essays* 32, 271–280.
- Lane, N., Martin, W.F., 2015. Eukaryotes really are special, and mitochondria are why. *Proc. Natl. Acad. Sci. USA* 112, E4823.
- Lanier, K. A., Williams, L. D. 2017. The origin of life: models and data. *J Mol Evol.* 84, 85–92.
- Laura Eme, Anja Spang, Jonathan Lombard, Courtney W. Stairs, Thijs. G. J. Ettema, 2017. Archaea and the origin of eukaryotes. *Nat. Rev. Micro* 15, 711-723
- Leipe, D.D., Aravind, L., Koonin, E.V., 1999. Did DNA replication evolve twice independently? *Nucleic Acids Res.* 27 (17), 3389-401.
- Lombard, J, López-García, P., Moreira, D., 2012. The early evolution of lipid membranes and the three domains of life. *Nat. Rev. Microbio.* 10, 507-515.
- López-García, P., Moreira, D., 2015. Open questions on the origin of eukaryotes. *Trends Ecol. Evol.*30, 697–708.
- Luisi, P. L., 2016. *The Emergence of Life*. Oxford: Oxford University Press.
- Ma, B-G. et al. 2008. Characters of very ancient proteins. *Biochem. & Biophys. Res. Comm.* 366: 607-611.
- Margulis, L., 1981. *Symbiosis in Cell Evolution*. New York: W. H. Freeman.
- Margulis, L., 1991. Symbiogenesis and Symbiogenesis. Symbiosis as a source of evolutionary innovation: speciation and morphogenesis, eds Margulis L, Fester R (MIT Press, Cambridge, MA), 1–13.
- Margulis, L., Sagan, D., 2002. *Acquiring a genome: a theory of the origins of Species*. New York: Basic Books.
- Martin, W., Russell, M., 2003. On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells, *Phil. Trans. R. Soc. Lond. B*, 358, 59–85, <http://dx.doi.org/10.1098/rstb.2002.1183>.
- Miller, S., Krijnse-Locker, J., 2008. Modification of intracellular membrane structures for virus replication. *Nat. Rev. Microbiol.*, 6, 363-274.
- Milshcheyn, D., Damer, B., Havig, J., Deamer, D. 2018. Amphiphilic Compounds Assemble into Membranous Vesicles in Hydrothermal Hot Spring Water but Not in Seawater. *Life* 8 (2),

<https://www.mdpi.com/2075-1729/8/2/11>

Moerira, D., Lopes-Garcia, 2009. Ten reasons to exclude viruses from the tree of life, *Nature Reviews Microbiology* 7, 306-311.

Mulkidjanian, A.Y., Bychkov, A.Y., Dibrova, D.V., Galperin, M.Y., Koonin, E.V., Origin of first cells at terrestrial, anoxic geothermal fields. *Proc. Natl. Acad. Sci. USA* 2012; 109, E821-30, <http://dx.doi.org/10.1073/pnas.1117774109>

Mulkidjanian, A.Y., Galperin, M.Y., Koonin, E.V., 2009. Co-evolution of primordial membranes and membrane proteins. *Trends. Biochem. Sci.* 34, 206.

<http://dx.doi.org/10.1016/j.tibs.2009.01.005>

Norris, V., Raine, D. J., 1998. A fission-fusion origin for life. *Orig Life Evol Biosph.* 28, 523–537.

O'Malley, M. A., 2014. Endosymbiosis and its implications for evolutionary theory. *Proc. Natl. Acad. Sci. USA*, www.pnas.org/cgi/doi/10.1073/pnas.1421389112.

Oparin, A.I., 1953. *The Origin of Life*. 2nd. New York: Dover.

Orgel, L.E., 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, <http://dx.doi.org/10.1007/s11084-016-9494-1>.

Pearce, B. K. D., Pudritz, R. E., Semenov, D. A., Henning, T. K., 2017. Origin of the RNA world: The fate of nucleobases in warm little ponds. *Proc. Natl. Acad. Sci. USA.* 114 (43), 11327-11332.

Pereto, J., 2005. Controversies on the origin of life. *Inter. Microbiol.* 8.23-31.

Petrov, A.S., Gulen, B., Norris, A.M., Kovacs, N.A., Bernier, C.R., Lanier, K.A., Fox, G.E., Harvey, S.C., Wartell, R.M., Hud, N.V., Williams, L.D., 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 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 (7), 449-456, <https://doi.org/10.1016/j.resmic.2009.06.004>.

Pressman, A., Celia B., Irelen A. C., 2015. The RNA world as a model system to study the

- origin of life. *Curr. Biol. Rev.* 25, <http://dx.doi.org/10.1016/j.cub.2015.06.016>
- Puigbò, P., Wolf, Y.I., Koonin, E.V., 2003. The tree and net components of prokaryote evolution.
- Puigbò, P., Wolf, Y.I., Koonin, E.V., 2009. Search for a ‘Tree of Life’ in the thicket of the phylogenetic forest. *J Biol*, 8, 59, <http://dx.doi.org/10.1186/jbiol159>
- Qiao, H., et al. 2017. Encapsulation of Nucleic Acids into Giant Unilamellar Vesicles by Freeze-Thaw: a Way Protocells May Form. *Orig Life Evol Biosph* 47 (4): 499–510.
- Ranea, J. A., Sillero, A., Thornton, J. M., Orengo, C. A. 2006. Protein superfamily evolution and the last universal common ancestor. *J Mol Evol.* 63, 513–525, <http://dx.doi.org/10.1007/s00239-005-0289-7>.
- Rasmussen, S., et al, eds., 2009. Protocells: bridging nonliving and living matter. Cambridge, M. A., MIT Press.
- Richter, R., Haslbeck, M., Buchner, J. 2010. The heat shock response: life on the verge of death. *Mol Cell* 40, 253-266.
- Robertson, M. P., Joyce, G. F. 2012. The origins of the RNA world. *Cold Spring Harb Perspect Biol.* 4, a003608, [http:// dx.doi.org/10.1101/cshperspect.a003608](http://dx.doi.org/10.1101/cshperspect.a003608).
- Saad, N. Y., 2018. A ribonucleopeptide world at the origin of life. *J. Syst. Evol.* 56 (1), 1-13. <https://doi.org/10.1111/jse.12287>
- Sagan, L., 1967. On the origin of mitosing cells. *J. Theoret. Biol.* 14, 225-274.
- Saha, R., Pohorille, A., Chen, I. A. 2014. Molecular crowding and early evolution. *Orig. Life Evol. Biosph.* 44, 319–324.
- Saha, R., Verbanic, S., Chen, I. A. 2018. Lipid vesicles chaperone an encapsulated RNA Aptamer. *Nat. Comm.* 9, 2313, [http:// dx.doi.org/10.1038/s41467-018-04783-8](http://dx.doi.org/10.1038/s41467-018-04783-8).
- Schrum, J. P., Zhu, T. F., Szostak, J. W. 2010. The origins of cellular life. *Cold Spring Harb Perspect Biol*, 2, a002212.
- Sengupta, S., and Higgs, P.G., 2015. Pathways of genetic code evolution in ancient and modern organisms. *J. Mol. Evol.* 80 (5-6), 229-43. <http://dx.doi.org/10.1007/s00239-015-9686-8>.
- Shimada, H., Yamagishi, A., 2011. Stability of heterochiral hybrid membrane made of bacterial sn-G3P lipids and archaeal sn-G1P lipids. *Biochem.* 50 (19), 4114–4120.

- 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(11), 2874–2884, [http:// dx.doi.org/10.1093/molbev/msw164](http://dx.doi.org/10.1093/molbev/msw164).
- Spitzer, J. 2018. Emergence of life on Earth: a physicochemical jigsaw puzzle. *J Mol Evol*, 84, 1–7, <http://dx.doi.org/10.1007/s00239-016-9775-3>.
- Staley, J. T., and Fuerst, J. A. 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.
- Sutherland, J.D., 2017. Studies on the origin of life—the end of the beginning. *Nature Rev. Chem.* 1 (1)
- Szathmáry, E., Marynard S.J., 1997. From replicators to reproducers: the first major transitions leading to life. *J. Theor. Biol.* 187 (4), 555-71.
- Szostak, J.W., 2011. An optimal degree of physical and chemical heterogeneity for the origin of life? *Phil. Trans. R. Soc. B*, 366, 2894–2901, [http:// dx.doi.org/10.1098/rstb.2011.0140](http://dx.doi.org/10.1098/rstb.2011.0140).
- Szostak, J.W., Bartel, D. P., Luisi, P.L., 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. [http:// dx.doi.org/10.1038/nchem.2739](http://dx.doi.org/10.1038/nchem.2739).
- Tessera, M. 2018. Is pre-Darwinian evolution possible? *Biol. Direct* (2018) 13:18 <https://doi.org/10.1186/s13062-018-0216-7>
- Tocheva, E. I., Ortega, D.R., Jensen, G.J., 2016. Sporulation, bacterial cell envelopes and the origin of life. *Nat. Rev. Microbiol.* 14, 535-542.
- Toman, J., Flger, J., 2017. Stability-based sorting: The forgotten process behind (not only) biological evolution. *J. Theor. Biol.* 435, 29–41.
- Van der Schaar, H. M. et al. 2016. Fat(al) attraction: picornaviruses usurp lipid transfer at membrane contact sites to create replication organelles. *Trends Microbiol.* 24 (7): 535-546, <http://dx.doi.org/10.1016/j.tim.2016.02.017>.
- Van der Veen, A.G., Ploeh. H.L., 2012. Ubiquitin-like proteins. *Annu. Rev. Biochem.* 81. 323-57, [http:// dx.doi.org/10.1146/annurev-biochem-093010-153308](http://dx.doi.org/10.1146/annurev-biochem-093010-153308).
- Vetsigian, K., Woese, C., Goldenfeld, N., 2006. Collective evolution and the genetic code.

- Proc. Natl. Acad. Sci. USA; 103:10696-701, <http://dx.doi.org/10.1073/pnas.0603780103>.
- Vitas, M., Dobovisek, A., 2018. In the beginning was a mutualism - on the origin of translation. *Orig. Life. Evol. Biosph.* <https://doi.org/10.1007/s11084-018-9557-6>.
- Wächtershäuser, G., 2003. From pre-cells to Eukarya—a tale of two lipids. *Mol. Microbiol.* 47, 13–22.
- Wei, C., Pohorille, A. 2011. Permeation of Nucleosides through Lipid Bilayers. *J. Phys. Chem. B.* 115, 3681-3688. dx.doi.org/10.1021/jp112104r
- Weiss, M., et al. 2016 The physiology and habitat of the last universal common ancestor. *Nat. Microbiol.* DOI: 10.1038/NMICROBIOL.2016.116
- Williams, R. J., Fraústo, Da Silva, J. .J., 2003. Evolution was chemically constrained, *J. Theor. Biol.* 220, 323–343.
- Wilson, M. A., Wei, C., Pohorille, A. 2014. Towards Co-Evolution of Membrane Proteins and Metabolism. *Orig Life Evol Biosph*, 44: 357–361, DOI 10.1007/s11084-014-9393-2
- Woese, C.R., 1998. The universal ancestor. *Proc. Natl. Acad. Sci. USA*, 95, 6854-9, <http://dx.doi.org/10.1073/pnas.95.12.6854>.
- Woese, C.R., 2002. On the evolution of cells. *Proc. Natl. Acad. Sci. USA*, 99, 8742-7, <http://dx.doi.org/10.1073/pnas.132266999>.
- Woese, C.R., Fox, G.E., 1977. The concept of cellular evolution. *J Mol. Evol.*, 10, 1-6, <http://dx.doi.org/10.1007/BF01796132>.
- Woese, C.R., Goldenfeld, N., 2009. How the microbial world saved evolution from the scylla of molecular biology and the charybdis of the modern synthesis. *Microbiol. Mol. Biol. Rev.*, 73, 14-21, <http://dx.doi.org/10.1128/MMBR.00002-09>.
- Wolf, Y.I., Koonin, E.V., 2007. On the origin of the translation system and the genetic code in the RNA world by means of natural selection, exaptation, and subfunctionalization. *Bio. Direct.* 2, 14, <https://doi.org/10.1186/1745-6150-2-14>.
- Wong, J.T.F. 1975. A Co-Evolution Theory of the Genetic Code. *Proc. Natl. Acad. Sci. USA*, 72, 1909–1912.
- Wong, J.T.F. 1981. Coevolution of genetic code and amino acid biosynthesis. *Trends Biochem. Sci.* 6, 33–36.
- Xing, X. S., Chris, T.H., Antonis, R., 2017. Contentious relationships in phylogenomic studies

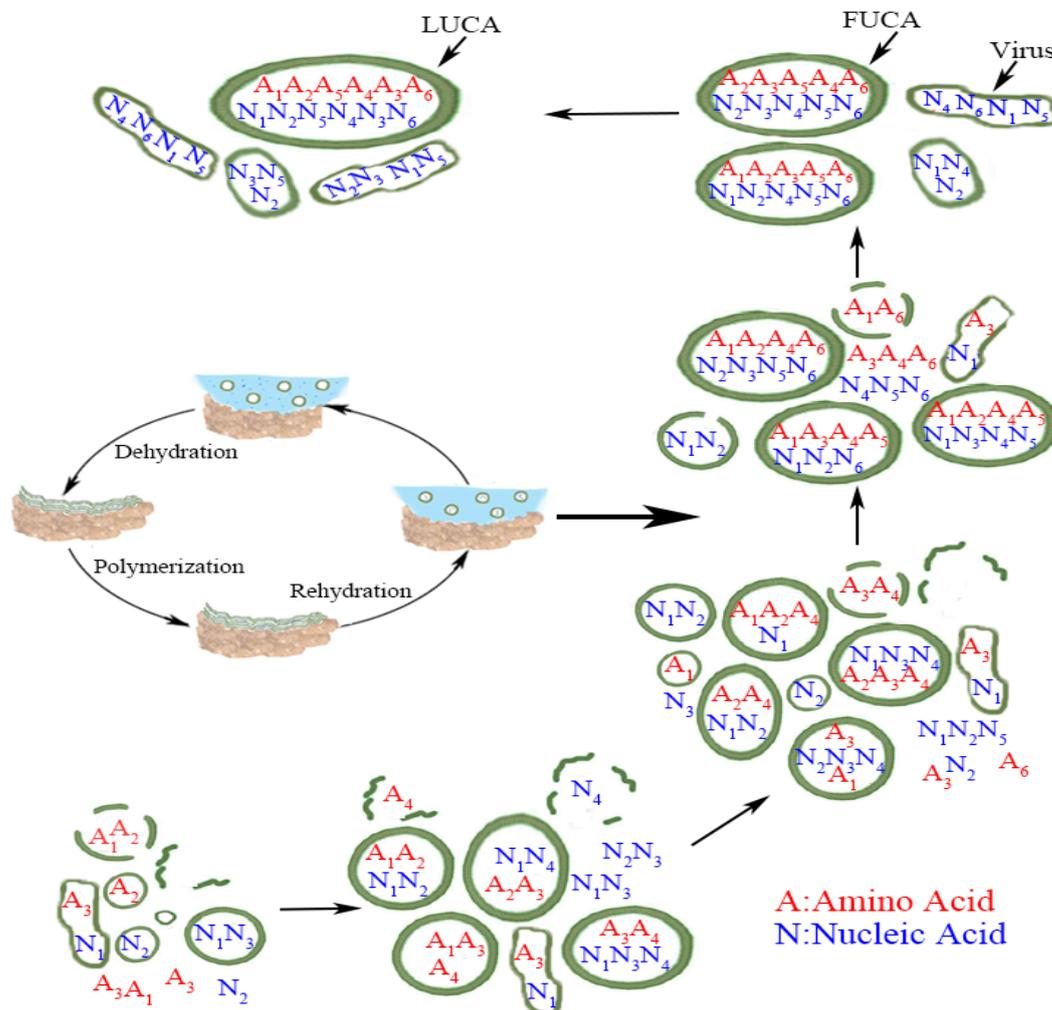
can be driven by a handful of genes. *Nat. Eco. Evol.* <http://dx.doi.org/10.1038/s41559-017-0126>.

Yarus, M., 2017. The genetic code and RNA-Amino acid affinities. *Life*, 7, 13, <http://dx.doi.org/10.3390/life7020013>.

Zhu, T.F., Adamala, K., Zhang, N., Szostak, J.W., 2012. Photochemically driven redox chemistry induces protocell membrane pearling and division. *Proc. Natl. Acad. Sci. USA* 25, 9828–9832.

Zhu, T.F., Szostak, J.W., 2009. Coupled growth and division of model protocell membranes. *J. AM. Chem. Soc.* 131, 5705–5713, <http://dx.doi.org/10.1021/ja900919c>.

Zimorski, V., Ku, C., Martin, W. F., Gould, S. B. 2014. Endosymbiotic theory for organelle origins. *Curr. Opin. Microbiol.*, 22, 38–48.

Figure 1. From FUCA 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 (SGC) had evolved most completely by the time of LUCA. The less than exact matching amino acid and nucleic acid within FUCA and vesicles before FUCA denotes the evolutionary path of SGC from a rudimentary form to a mature form in LCUA.

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. The wet-and-dry cycle (on the left of the diagram) might have played a key role in driving the process of breaking-and-re-encapsulation cycle that facilitates the merger and acquisition by vesicles. The wet-and-dry cycle part within the figure is adapted from Bruce Damer and David Deamer (2015) with permission.