Progress in Biophysics and Molecular Biology 180–181 (2023) 28–36

Contents lists available at ScienceDirect

Progress in Biophysics and Molecular Biology

journal homepage: www.elsevier.com/locate/pbiomolbio

Origin of life: Drawing the big picture

Francisco Prosdocimi [a,*], Sávio Torres de Farias [b,c]

[a] Laboratório de Biologia Teórica e de Sistemas, Instituto de Bioquímica Médica Leopoldo de Meis, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil
[b] Laboratório de Genética Evolutiva Paulo Leminski, Centro de Ciências Exatas e da Natureza, Universidade Federal da Paraíba, João Pessoa, Paraíba, Brazil
[c] Network of Researchers on the Chemical Evolution of Life (NoRCEL), Leeds, LS7 3RB, UK

A B S T R A C T

Trying to provide a broad overview about the origin of life in Earth, the most significant transitions of life before cells are listed and discussed. The current approach emphasizes the symbiotic relationships that emerged with life. We propose a rational, stepwise scenario for the origin of life that starts with the origin of the first biomolecules and steps forward until the origins of the first cells. Along this path, we aim to provide a brief, though comprehensive theoretical model that will consider the following steps: (i) how nucleotides and other biomolecules could be made prebiotically in specific prebiotic refuges; (ii) how the first molecules of RNAs were formed; (iii) how the proto-peptidyl transferase center was built by the concatenation of proto-tRNAs; (iv) how the ribosome and the genetic code could be structured; (v) how progenotes could live and reproduce as "naked" ribonucleoprotein molecules; (vi) how peptides started to bind molecules in the prebiotic soup allowing biochemical pathways to evolve from those bindings; (vii) how genomes got bigger by the symbiotic relationship of progenotes and lateral transference of genetic material; (viii) how the progenote LUCA has been formed by assembling most biochemical routes; (ix) how the first virion capsids probably emerged and evolved; (x) how phospholipid membranes emerged probably twice by the evolution of lipid-binding proteins; (xi) how DNA synthesis have been formed in parallel in Bacteria and Archaea; and, finally, (xii) how DNA-based cells of Bacteria and Archaea have been constituted. The picture provided is conjectural and present epistemological gaps. Future research will help to advance into the elucidation of gaps and confirmation/refutation of current statements.

1. Introduction

The origin of life is one of the major open problems of biology and one of the three great, classic scientific problems: the others being the origin of the universe, and the origin of mind. Being of such relevance, researchers from different intellectual and cultural backgrounds have tried to tackle the question, constituting a transdisciplinary, complex research and an open, exploratory field of study. Here, we aim to review the most recent advances in the origins of life research program and provide a still conjectural, though broad, rational, and gradualist scenario for the origin of life in Earth (Fig. 1). Based on previous works and decades of study, we propose a complete scenario divided in four ages: (i) the prebiotic age; (ii) the age of FUCA and the origin of life; (iii) the age of progenotes; and (iv) the age of organisms. The whole scenario will be explained along this manuscript. Although there will be plenty of specific questions and details that will need further investigation, we believe to have reached a general model that may explain how life emerged from chemistry in the Hadean or Eoarchean Earth, from 4.5 to 3.6 billion years ago, shedding light to the most important transitions in prebiotic and early biotic evolution.

2. Origin of life: the top-down approach

The most relevant theories and methods applied to understand the origins of life can be divided in two main epistemological approaches: (i) from the bottom to the top, on which researchers start to consider how the atmospheric and geological conditions of early Earth allowed the emergence of biomolecules; and (ii) from top to the bottom, on which researchers use data from the current living organisms as indicatives of their shared ancestrality (Prosdocimi and Farias, 2019; Prosdocimi et al., 2018).

We are currently living in the genomics and post-genomics age of biological sciences. The comparison of DNA content among organisms (as well as other omics approaches) demonstrated to be capable to elucidate indeed many interesting regularities and allowed a better understanding about a fundamental question in genetics: the relationship between genotypes and phenotypes. Therefore, theoretical models that take genomic information on account are usually considered better, more concise, more reproducible, and more trustworthy. That is why, year after year, researchers from all over the world working in the origin of life use genomic data as basis to provide *bona fide* models for the evolution of life in Earth (Kim et al., 2008; Seufferheld and Caetano-Anollés, 2013; Prosdocimi et al., 2020). In that sense, the most

* Corresponding author.
E-mail address: prosdocimi@bioqmed.ufrj.br (F. Prosdocimi).

https://doi.org/10.1016/j.pbiomolbio.2023.04.005
Received 4 January 2023; Received in revised form 15 March 2023; Accepted 17 April 2023
Available online 18 April 2023

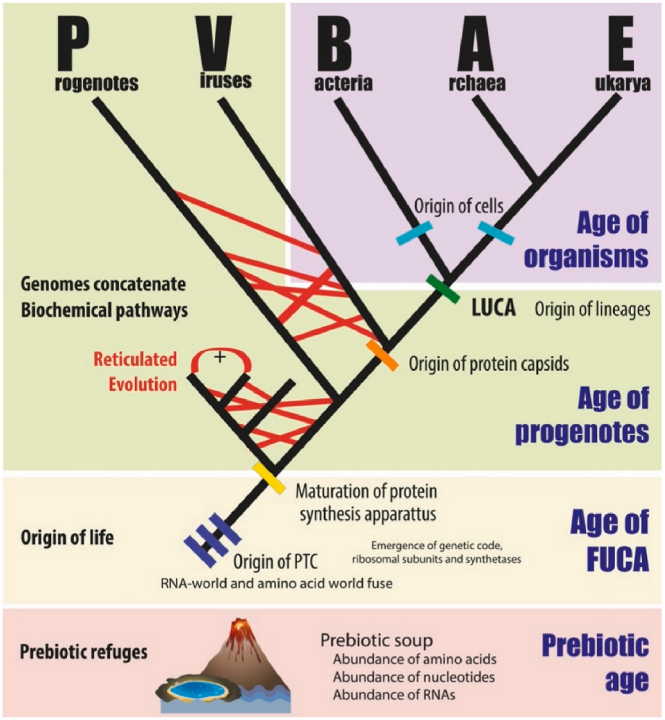

Fig. 1. The big picture: from prebiotic age to the age of organisms.

standard approach to appreciate the origin of life in the light of genomics deals with the study of all genomes available and checking *which genes and regulatory elements they all have in common*. Which genes every organism have and share with others? The rationale of this approach is Darwinian as it considers that all life in Earth is descendent of an early life that originated in the planet and give rise to complete set of biodiversity we see today. The genes in contemporary organisms are then seen as some sort of new and updated versions of ancient genes that lived a long time ago.

The whole set of biodiversity present in Earth basically consists of three great domains of organisms of life: the Bacteria, the Archaea, and the Eukarya. Although those domains have a cellular constitution, there are significant divergences that indicate their early differentiation in the root of the tree of life. Thus, instead of taking "each and every" organisms and checking which genomic features they all have in common, a task that would expand the computational power available, researchers often keep it simple and choose a manually curated set of organisms that represent the most significant clades inside those great domains. This method of working aims to reconstruct the genome of the first ancestor of those three cellular groups. Named LUCA by the renown American microbiologist Carl Woese as an acronym of "*Last Universal Common Ancestor*", this ancestor would evolve to give rise to the great domains of life (Woese et al., 1990). Thus, along the last decades, many of the most important works in the origins of life tried to discover the whole set of genes present in LUCA's genome (Lazcano and Forterre, 1999; Koonin, 2003; Delaye et al., 2005; Ouzounis et al., 2006; Mushegian, 2008; Mat et al., 2008; Kannan et al., 2013; Weiss et al., 2016).

A modern classic work on these matters was published in 2016 by a German group lead by the microbiologist William Martin (Weiss et al., 2016). Martin and his group downloaded the genomes of organisms in the three domains of life and published a masterpiece scientific work entitled "*The physiology and habitat of the last universal common ancestor*". The manuscript was published in July 2016, in Nature Microbiology, and presented a LUCA anaerobic, thermophilic, dependent of hydrogen, and capable to fix CO_2 and N_2. According to the authors, LUCA was an autotroph that inhabited a geochemically active environment (Weiss et al., 2016). For us, the most astonishing information provided by Martin and co-workers was the fact that LUCA presented 355 gene families, all of them precisely identified, provided as supplementary information and studied along this historical work. Previously, Arcady Mushegian predicted the number of LUCA's genes as 500–1000 (Mushegian, 2008) and Wai-Kin Mat and co-workers predicted 424 genes with 150–340 being essential (Mat et al., 2008).

3. Origin of life: the bottom-up approach

Once a LUCA-like entity presenting hundreds of genes could not appear from nothing, the question about the origin of life must go back in time. How has LUCA first appeared? How were these hundreds of genes structured and assembled in LUCA's genome? How did life emerge at first? Prebiotic chemists predicted the chemical constitution of abiotic Earth and tried to understand what lead to appearance of the first genes.

The bottom-up approaches to the origin of life aim to answer the questions, such like: what was the atmospheric composition of prebiotic Earth? And how the molecules existing prior to life came to bind together to produce the first monomers to the biological polymers? Two classical works are often invoked it the field, the first being proposed by the Russian biochemist Aleksandr Oparin, in 1924. Oparin created the concept of a "prebiotic soup" on which the basic elements and molecules for the origin of life were available to the initial assembly of peptides and nucleic acids (Oparin, 1924). Twenty-five years later, in 1953, the American chemist Stanley Miller made his experiment simulating the early Earth. Using an ingenious apparatus, Miller simulated the early Earth's atmosphere using water (H_2O), hydrogen (H_2), ammonia (NH_3), methane (CH_4), and hydrogen sulfide (H_2S) (Miller, 1953). Together with simulated heat and electric discharges, Miller observed that amino acids could be produced abiotically. Miller's work was a milestone for the field of prebiotic chemistry and several groups have been developing work inspired on his classical experiment. Due to these studies, it has been demonstrated that sugars and nitrogenous bases can be synthesized in environments that simulate the early conditions of the Earth (Ingar et al., 2003; Powner et al., 2007, 2009; Delidovich et al., 2014; Lamour et al., 2019; Kim and Kim, 2019; Yadav et al., 2020).

4. The RNA world

The work of Stanley Miller was published in the very same year on which James Watson and Francis Crick proposed the double-helix model for the structure of the DNA (Watson and Crick, 1953). Although initially DNA has been recognized as the most important nucleic acid in cells, responsible to control the metabolism and transmit genetic information, the relevance of RNA was raising year by year. Being simpler and presenting a more unstable structure, at some point researchers on molecular biology concluded that RNA was probably the first nucleic acid to originate (Woese, 1965; Crick, 1968; Orgel, 1968). DNA then came to be seen as a specialized chain of nucleic acids on which the information has been better protected, turning to be less reactivity to other molecules interactions (Forterre, 2005; Di Giulio, 2021).

Historically, amino acids were the first biomolecules shown to be produced abiotically (Miller, 1953). Therefore, most researchers started to theorize scenarios for the origin of life that were based only in peptides and proteins. However, with the advance of molecular biology, it became clear that nucleic acids also needed to be present in the context of early life. The Nobel laureated German biophysical chemist Manfred Eigen have proposed the origin of life based on protein hypercycles in 1971, but later he modified his theory and suggested that nucleic acids should have started life because they were better molecules to store information and replicate (Eigen, 1971).

Then, thirty-years after Miller's publication, it has been proved that RNAs were capable to perform catalysis (Kruger et al., 1982; Guerrier-Takada et al., 1983). And three years later, another Nobel laureated, the American biochemist Walter Gilbert published a single-page work in Nature describing the idea of an RNA-world (Gilbert, 1986). Gilbert

argumented that the mere existence of ribozymes was evidence that protein enzymes were not necessary to the origin of life.

The proposal of an RNA-world is often referred as one of the chicken-and-egg discussions in the origins of life. In that case, chicken geneticists argue that genetic encoding and RNA catalysis were necessary and sufficient to maintain the origin of life (Gilbert 1986; Schwartz 1995; Dworkin et al., 2003; Robertson and Joyce 2012). Contrarily, egg biochemists insisted that peptides and proteins were more important for the origin. They suggest that proteins were more important than RNA issues because (i) they present a higher level of chemical and 3D structures. This versability made peptides and proteins the main molecules that organized the protometabolism, being a hub of molecular contact. This is also verified by the low efficiency of RNA catalysis. Also, they argue that (ii) the absence of nucleotide molecules in prebiotic simulations and experiments (Di Giulio, 1997; Caetano-Anolles and Seufferheld, 2013) is evidence against the RNA-first hypothesis. Our proposal of a chemical symbiosis between nucleic acids and proteins as the most relevant event in the early origin of life (Prosdocimi et al., 2021b) acknowledge the relevance of both nucleic acids and proteins. This model agrees with both chicken geneticists and egg biochemists, suggesting that life could not emerge without any of these biomolecules. As RNAs were important to store information and evolve, proteins were key for the self-organization of molecular interactions that sparkled metabolism and allowed life to emerge.

The biochemists were correct about the difficult problem of producing nucleotides in prebiotic conditions. After decades of studies, scientific models still fail to provide a clear prebiotic scenario on which nucleotides could be produced in the amounts necessary to spark life (Le Vay and Mutschler, 2019; Prosdocimi et al., 2022). The mystery about the prebiotic synthesis of nucleotides still haunts the RNA-world advocates and provide a straightforward field of work in contemporary prebiotic chemistry.

In any case, even the biochemists now agree that it was important that nucleic acids could be produced in considerable amounts in prebiotic context. In a work entitled "Prebiotic chemical refugia" we review the most up to date information about nucleotide formation, citing dozens of works that try to produce riboses, purine and pyrimidine nucleobases, phosphates, and RNA polymers (Prosdocimi and Farias, 2019; Prosdocimi et al., 2022). Briefly, there is a possibility that ribose has come to Earth in comets or asteroids, nucleobases were produced in cycles of wet and dry in shallow ponds, phosphates could be formed in carbonate-rich lakes and RNA polymers could be formed in glaciers. These particular geochemical sites favored the production of nucleotides and were named prebiotic chemical refugia.

5. From the RNA world to the proto-PTC

Understanding the possible origins of the basic constituents found in biological systems is extremely important to glimpse the circumstances that led to the origin of biological systems. Currently, there are demonstrations for the abiotic synthesis of riboses (Meinert et al., 2016; Lazcano and Bada, 2003), nucleotidic bases (Burton et al., 2012; Menor-Salván and Marín-Yaseli, 2013), small peptides (Levy et al., 2000; Johnson et al., 2008), carbohydrates (Hollis et al., 2000), lipids (Lopez and Fiore, 2019; Damer and Deamer, 2015; Pasek and Lauretta, 2008), and other compounds that constitute the contemporary biological systems (Orgel, 2004; Becker et al., 2018). We conjecture that each of those prebiotic synthesis of compounds happened under microgeographic places (or refuges) that presented a specific set of favorable chemical conditions (Prosdocimi et al., 2022). Under a given refuge, biomolecules were assembled into nucleic acid polymers by self-assembly forces. Some authors suggest that transfer RNAs (tRNAs) were one of the first RNA 3D shapes capable to organize themselves (Gilbert 1986; Caetano-Anollés and Caetano-Anollés, 2016; Farias et al., 2016).

In a top-down approach to understand the sequence and topological features of early tRNAs, our group used phylogenetic ancestral reconstructions to produce putative forms of proto-tRNAs (Farias et al., 2014a,b). Those proto-tRNAs were aligned with the 23S ribosomal molecule in the Peptidyl Transferase Center (PTC) region, that corresponds to the catalytic site of the ribosome, a highly conserved ribozyme that catalyzes the binding of amino acids in every organism. The result showed that some proto-tRNAs had similarities with different parts of the PTC and these similarities were found to be in neighboring regions, suggesting that the 23S ribosomal molecule could have originated from the concatamerization of proto-tRNAs. (Farias et al., 2014, 2016). It has been also demonstrated that some of these proto-tRNA concatamers were 50% identical to the PTC.

If we recapitulate briefly some works dealing with the LUCA's genome, there has been always a consensus that one of the first biochemical routes that needed to emerge in early life were related to the informational machinery. Therefore, replication, transcription and translation routes needed to be present and functioning in the LUCA (Koonin, 2003). On the other hand, we will see that DNA probably did not exist in early Earth and the early nucleic acids were most likely composed of RNA (Forterre, 2005; Di Giulio, 2021). Therefore, transcription was not necessary as well. On the other hand, the protein synthesis apparatus also known as the translation machinery has been understood as having a key importance to the origin of life. This is because the genetic code (Prosdocimi and Farias, 2021; Farias et al., 2021) and a precise correlation between nucleic acids and peptides are at the center of this sort of chemical symbiosis that originated life (Prosdocimi et al., 2021a). A gene has no reason of existing if it cannot encode a protein or act as ribozyme. Even if the first RNA genes were possibly selected by being capable to be stable and replicate, the most consistent scenarios for the origin of life considers that translation emerged in the very beginning (Davidovich et al., 2009; Belousoff et al., 2010; Petrov et al., 2014, 2015; Farias et al., 2014a,b, 2016, 2017, 2020).

Although the translational machinery present in contemporary organisms is highly complex, with the presence of two ribosomal subunits formed by a complex set of dozens of proteins and a couple RNAs, models that describe its evolution from simple molecules have been proposed (Belousoff et al., 2010; Farias et al., 2014a,b). Many groups suggest a scenario on which the translation machinery evolved from the random formation of its catalytic center (Belousoff et al., 2010; Petrov et al., 2015). It is well known that the catalytic center of the ribosome is the Peptidyl Transferase Center (PTC) region, located in the main ribosomal RNA that forms the major ribosomal subunit. Petrov et al. (2015) comparatively analyzing the structure of several ribosomes, suggested that they were formed from PTC by accretion of structures that led to the complexification of their function. Xu and Wang (2021) analyzing the catalytic properties of rRNA without proteins, observed that portions of this molecule could bind two fragments of tRNAs located closely, and when analyzing the structure of these portions, they observed that they mimicked the structure of PTC. Bose et al. (2022) synthesized a proto-ribosome containing parts of the PTC and observed that some of these proto-ribosomes had catalytic activity, forming bonds between amino acids. Rivas and Fox (2023) analyzed in detail the results presented by Bose et al., (2022) and observed that the results suggest that the P site must have been the first to organize itself. They also indicate that the catalytic activity and the ability to form dimers is sequence dependent (Rivas and Fox, 2023). Together. these data reinforce the centrality of the ribosomal catalytic site in the initial organization of biological systems, as well as highlight the real possibility of the emergence of a catalytic RNA capable of binding two or more amino acids.

Some authors suggest that the PTC might be formed the concatenation of proto-tRNAs (Farias et al., 2014a,b; Root-Bernstein and Root-Bernstein 2015, Farias et al., 2017, Demongeot and Seligmann, 2022). Prosdocimi et al. (2020) found that the concatenation of proto-tRNAs encoding Proline, Tyrosine, Phenylalanine, Glutamine and

Glycine could form two types of secondary and tertiary structures: one that facilitated the catalysis of amino acids together, working as a primitive PTC and another that possibly facilitated its replication. In addition, these authors found out that the catalytic site formed by an Adenine is absolutely conserved in all domains of life and we provided another structural support to the idea (Prosdocimi et al., 2020).

6. From the proto-PTC to the ribosome

Although the complete evolution of the ribosome is still full of uncertainties, our model suggests that different population of RNA molecules would work as ribosomal or messenger RNAs, both being formed by proto-RNAs concatamers. Once a proto-PTC has been capable to bind and catalyze the fusion of amino acids, the amino acids in the pool started to be converted in small peptides. The initial formation of peptides by the proto-PTC would initially be quasi-random, considering the amounts and types of amino acids available in the prebiotic soup in a given moment (Davidovich et al., 2009). Some of those peptides would eventually bind back the proto-PTC and the ones that allowed the machinery to be more stable and efficient would be selected by Darwinian molecular forces to be maintained (Farias et al., 2016).

Initially, it seems that transfer RNAs would have their amino acid and anti-codons binding sites closely related in terms of structure. That is why a hydrophobicity relationship between amino acids and anti-codons has been observed (Guimarães et al., 2008). Also, it has been demonstrated that tRNA-aminoacyl synthetases recognize other sites besides the anticodon suggesting a structural co-evolution between tRNAs and these enzymes (Zamudio et al., 2019). The researcher Loren Williams and his group have divided the ribosome in eight different parts and provided a stepwise model showing how the ribosome has probably been formed under an accretion model, on which different layers have been structure one over the other along its evolution (Petrov et al., 2014, 2015; Petrov and Williams, 2015). Other relevant works also tried to understand the origin of ribosome and the translational machinery (Davidovich et al., 2009), some of them denying the origin of the ribosome from the PTC (Caetano-Anollés, 2015).

7. The chemical symbiosis' theory and the birth of FUCA

We conjecture that the origin of a proto-PTC from random concatamers of proto-tRNAs was probably the first step that linked the nucleic acids world to the amino acidic world (Farias et al., 2014a,b, 2016; Prosdocimi et al., 2020). Other authors also suggested that it was only with the establishment of the interaction between amino acids and nucleic acids that life became established (Lanier et al., 2017, Vitas and Dobovišek, 2018). The proto-PTC enhanced the relationship between two previously separated biomolecular worlds. The existence of PTC allowed that nucleic acids and peptides could start to interact for the benefit of both, allowing the emergence of order in nature (Prosdocimi et al., 2021). The production of small, quasi-random peptides allowed that the nucleic acids and some of these recently formed peptides become more stable. It also allowed those amino acids to form longer chains of proteins, capable to fold in complex structures and interact with other molecules in the prebiotic soup, increasing the number of bonds, linkages, and interactions.

We believe that this point on which the first PTC appeared was key in the history of life, being the time on which the First Universal Common Ancestor (FUCA) has been born (Prosdocimi et al., 2019). If LUCA was the last universal ancestor of cellular lineages and presented hundreds of genes, we would like to find the oldest ancestor from LUCA's lineage. And this would be the FUCA, the molecular structure on which the symbiosis of nucleic acids and proteins was originated and maturated. Even if the FUCA cannot be considered an organism, as it is formed simply by a naked form of RNA, it sparked life when connected the RNA-world to the peptidic-world. Those two worlds have been wandering in early Earth quite independently. First, a tRNA-like molecule bound an amino-acid and then another tRNAs with bounded amino acid concatenated to that first one. Through the concatamerization of proto-tRNAs, the proto-PTC was formed and capable to bind amino acids together, linking nucleic acids and proteins forever, and opening the way for the history of life in the planet Earth (Farias et al., 2021a,b).

Even if amino acids originated first and were more abundant, as most scenarios suggest, they would never become living if they did not bound nucleic acids chains. It is still under debate if the amino acids helped in the formation of the proto-tRNA concatamers of whether the tRNA structures allowed their binding to amino acids. Also, the RNA-world would never become living if the self-replicating RNA molecules and ribozymes did not bound amino acids to produce encoded peptides. In this scenario, the origin of life is a case of chemical symbiosis, and it happened when nucleic acids and peptides started to interact, crosstalk and stabilize each other's structures. When the proto-PTC emerged, a critical aspect of life started its way and FUCA was born (Prosdocimi et al., 2019).

8. Organic codes and the maturation of FUCA

This initial connection between nucleic acids, amino acids and peptides will need to evolve so that live could basis its nature under a process of encoding. Biosemiotics theories suggest that life can be defined by the interplay of codes (Emmeche, 1998). If this is the case, we must first understand what we mean by "codes". Thus, we need to clarify that one thing is the proposal that biology is defined by codes as organisms interact by coding/decoding relationships. A deeper issue is comprehending that living beings exist only because there is a coding processes that operates inside their bodies in multiple levels (Barbieri 2012, 2013, 2014, 2016; Hofmeyr, 2018). That is the difference between biosemiotics and code biology. Code biology was proposed by the Italian biologist Marcello Barbieri as a dissidence of biosemiotics (Barbieri, 2003). Code biology is interested in the interplay of codes, but it is even more interested in understanding the biological codes, studying them in detail and analyzing how they operate mechanistically inside the cells (Barbieri, 2014).

It is relevant at this point to understand clearly what we mean by an organic code. To define any organic code, we need that (i) the coding rules should not be dictated exclusively by the laws of physics and chemistry; (ii) the self-organization of "arbitrary rules" allow the correspondence between two previously nearly independent "worlds"; and, finally, (iii) the presence of a translator or adaptor (Barbieri, 2003). Our view of life suggest that it is based on multiple layers of processes of coding/decoding (Farias et al., 2020).

The first code to emerge, the one that characterized life as a process and allowed the maturation of FUCA was the genetic code. No matter how it has precisely evolved, the genetic code is an organic code and allowed the maturation of the connection between the world of nucleic acids and the world of amino acids and peptides.

8.1. From FUCA to the world of progenotes

Once the genetic code has been stablished, multiple populations of FUCA-like ribonucleoproteins started to be formed and evolve. Those proto-organisms were formed by "naked" ribonucleoprotein (RNP) molecules capable to translate their information into peptides. They were probably produced in some sort of liquid solution that formed pseudo-compartments based on the abundance of peptide-nucleotide droplets (Koga et al., 2011). After the maturation of the genetic code, those FUCA-like entities started to replicate and exchange genetic material. Time by time, considering RNA replication errors and the binding and recombination of FUCA entities to other RNA-world molecules, it emerged a new world of evolving RNP entities. The idea of progenotes described by Carl Woese is perfect to characterize those proto-organisms that lived in the origin of life context (Woese, 1998). Even if Woese did not consider them as RNP particles, we decided to inherit his concept as

we believe that this is similar to what he has imagined and to honor his works and theoretical constructions.

Under the current conjecture, the world of progenotes will endure along the entire transition from FUCA to LUCA. It was there that the metabolic routes have been created and evolved (Prosdocimi et al., 2019) based on peptides that bound other molecules in the prebiotic soup. The world of progenotes was an age on which free, naked RNP molecules created conditions to their living. Along this era, RNP molecules and their encoded peptides started to interact more directly with other metabolites present in the media. As life started as ribonucleoproteins, other molecules and metabolites in the pool came to be incorporated in life by interacting with peptides, RNAs and RNPs that constituted the progenotes.

All biochemical pathways started with the origin of binding properties between peptides, RNAs, RNPs and the metabolites available in the media. Probably, progenotes created randomly by the process of genetic drift and mutation would produce new peptide conformations and foldings. Those conformations would allow their binding to the molecules in the pool. In the case that these bindings were beneficial to the progenotes, helping them to endure or replicate, they would be incorporated in their building genomes by molecular natural selection. At that point, progenotes also interacted with one another and exchanged information by lateral transference. This way, the biochemical routes were also be accumulated by the symbiotic junction or collaboration between progenotes, in accordance to theories that suggested a polyphyletic origin of genes (Di Giulio, 2008).

As initial progenotes were mainly ribosomal in nature, it is possible to propose that further progenotes may not need to have the ribosomal molecules on them and might use the protein synthesis machinery of other fellow progenotes in the nearby. This relationship could be viewed as a symbiotic process, in the case that the progenote that use the ribosome of another provides some exchange of beneficial substances; or it may be viewed as a parasitic form of progenotes that merely use others to decode their genetic information.

Recent works in prebiotic chemistry suggest that many biochemical cycles and routes already existed without the presence of any biological catalyst (Keller et al., 2014; Ralser, 2018). For example, it has been observed that the glycolytic pathway, the citric acid cycle and other pathways such as the pentose pathway could operate by only physical and chemical influences (Keller et al., 2014; Muchowska et al., 2019a; 2019b). It is a field of exploration to understand whether progenotes could interact with those protobiotic cycles and use those organization to guide their own evolution (Prosdocimi and Farias, 2022).

Regarding the naked version of progenotes, we should note recent data have confirmed that groups of ribonucleoproteins can form some sort of colloidal liquids that create microenvironments inside cells (Koga et al., 2011; Gomes and Shorter, 2019; Zhao and Zhang, 2020). The presence of given sets of ribonucleoproteins allows the cell to have spatial organization and we propose that this phenomenon could also occur in the precellular environment. In the study of prions and prionic diseases, those soft compartments allow cells to perform some sort of compartmentalization that is often named as “liquid-liquid phase separation”. Thus, it may be the case that the own ribosomes could form some sort of these biomolecular condensates to allow metabolism to exist in a sort of liquid droplets. Also, some early progenotes may agglomerate and form those pseudo-compartments to shelter other “naked”, non-encapsulated, free progenotes.

9. Organic codes become encapsulated: from the era of progenotes to era of organisms

According to some calculi, the number of virion particles present in Earth is at the same magnitude to the number of stars in the entire Universe, estimated as 10^{23} (Suttle, 2013). If even a small fraction of those entities were actually parasites, life would not exist on Earth. Viruses are also agents of symbiosis, they allow that genomes stay dynamic and fluid, and sometimes their population becomes dysregulated, causing disruption in some system in the search to achieve a new level of homeostasis. If the system cannot find this further level of homeostasis, it may be destroyed by the over action of those viral populations. But this only happens as an exception to the relevant role these organisms have in the ecological regulation of biosphere dynamics (Prosdocimi et al., 2021b).

The question about the origin of viruses is often described by three scenarios: (i) viruses are seen as direct descendants of a molecular world before the emergence of cells; (ii) viruses are seen as products of a process of reduction of cellularity with loss of autonomy; and, (iii) they are seen as products of the escape of genetic material from cells along with the capture of the host proteins that form their capsids (Krupovic et al., 2019). Despite seeming to be mutually exclusive, these models for the origin of viruses must be seen as complementary, since currently the scenarios suggest a polyphyletic origin for these groups. Therefore, we propose that the viral strategy of life can be achieved by different paths. As viruses are essentially ribonucleoproteic and we propose that life started as this (Farias and Prosdocimi, 2022), Occam’s razor principle favors an early origin of virus over the proposal that cells originated first. Cells are much more complex entities and present an additional level of control to another biomolecules (*i. e.,* phospholipids). We will further consider the origin of cells from lipid-binding peptides. In the present work, we aim to focus our discussion on the virus first model.

Some authors proposed models for the primitive origin of the virus, Koonin and Martin (2005) suggest that viruses formed before cells from self-replicating molecular units. Nasir and Caetano-Anollés (2015) suggest based on structural analysis data that viruses emerged with modern cell ancestors, through the cell reduction process. Farias et al. (2019) suggest that viruses were descendants of an era where biological systems were not yet encapsulated. Here we suggest that virus could be originated under pre-cellular conditions by the self-assembly of capsid peptides that involved and protected small RNA genomes and could be eventually replicated by their connection to ribosomes.

According to the current conjecture, at the age of progenotes, proto-ribosomes were freely available in the proto-colloidal media and different progenote populations were capable to use those free ribosomes to decode their information. As we said before, some progenotes populations could have their own proto-ribosome as others might borrow proto-ribosomes from other fellow populations. In this model, if viruses nowadays “infect” cells is because they need to get in touch to the ribosome to be replicated (Farias et al., 2019; Prosdocimi et al., 2021b).

Another relevant concepts to have in mind when we understand that viruses most likely existed before cells are related to two important issues: (i) viruses are orders of magnitude simpler than cells (Occam’s razor); and (ii) the basic core of life and living organisms are related to processes involving ribonucleoproteins (Prosdocimi and Farias, 2022), not lipids. Although lipids will become important in the age of cells, at an early moment they were not relevant. Along the age of progenotes, as we saw above, populations of RNP would replicate, accumulate information by concatamerization, and interchange chemical information by lateral transference. Thus, the biological information were built step by step as soon as some populations would produce quasi-random peptides or ribozymes that bound available molecules in the media. This binding would higher the stability of both the molecule and the ligand, allowing natural selection to maintain the partnership that worked producing fitness (Prosdocimi and Farias, 2022).

At some point, it is possible to imagine that some of the quasi-random peptides produced would be able to bind to themselves or to some other simple peptide. Some of those peptidic molecules would then fit together according to their 3D folding and start to stack. This stacking would form piles or groups of molecules each time bigger, until they would become a sort of capsid capable to store and protect the nucleic acid that encoded the corresponding information (Prosdocimi et al., 2021b). Thus, the stacking proteins would probably bind themselves in

geometrical or fractal-like shapes and allow this sort of "genetic case" to develop. This hypothesis needs to be better confirmed, but it seems reasonable as many authors have now suggested that viruses probably originated before cells (Di Giulio, 2021).

With the origin of viral capsids, progenote populations started to become more encapsulated and the way on which molecular evolution operated was modified. Although lateral transference would still happen, the encapsulation allowed biological entities to become more separated, starting an "organismal age". This era of organisms followed the era of progenotes and natural selection started to work more in the way we see today, *i. e.*, operating in vertical lineages that would produce clades or groups of organismal inheritance (Prosdocimi and Farias, 2021).

If viruses inaugurated the organismal age by symbiotically linking together their biological information to produce large, self-replicative molecules encoding virion capsids, they have never had the capacity to link together and produce large entities. Until nowadays, viruses keep being tiny particles of nucleic acids encapsulated in protein coats without a cytoplasm. Even with the discovery of the giant viral genomes of Mimiviridae, they keep being small and incapable to grow symbiotically into large organisms (Claverie et al., 2006). At last, viruses should be understood as a polyphyletic strategy of life that has been achieved multiple times in the history of biological organisms (Farias et al., 2019; Prosdocimi et al., 2021b). Thus, even if the origin of some viral clades alive today could be probably traced back to the era of progenotes, it is clear that viral behavior can be achieved also by reduction of different sorts of cells (Forterre and Prangishvili, 2009; Harris and Hill, 2021). Those last ones should be considered as modern versions of viruses.

9.1. Castor and polux: the LUCA brothers

It is finally time to reconsider the idea of a LUCA. We saw that Eukaryotes have been recognized as a derived clade of Asgard Archaeas that made symbiotic relationships to ancestral alpha-proteobacteria (Spang et al., 2015; Zaremba-Niedzwiedzka et al., 2017). This means that cellular life was originated either by Bacterial or by Archaeal cell types. Both cell types evolved from progenotes populations.

When we analyze the biochemical routes in Bacteria and Archaea, it is interesting to note that both cell types are quite similar in most of their metabolic pathways as observed in the encoding of enzymes in their genomes. But there are critical differences. Two very specific genomic pathways cannot be said to be homologous between the two domains: (i) the DNA synthesis pathway (Di Giulio, 2021), and (ii) the lipid biosynthesis pathway (Sojo, 2019; Farias et al., 2021a,b). All the others being probably inherited from the same population of progenotes. Thus, we aim to suggest that these two pathways above were most likely the last ones to mature in the progenote populations that originated cells. More than that, we believe that LUCA should be understood as this population of progenotes that already presented most of biochemical pathways organized in a single genome, but neither DNA synthesis nor phospholipid synthesis were present in LUCA. Therefore, LUCA was a progenote, not a cellular entity (Farias et al., 2021a,b).

Following the assembly of most of biochemical pathways in the genome of the progenote LUCA, there should have had a separation between two populations of LUCA entities. From one side, evolved the archaeal cells. From the other, evolved the bacterial ones. In both clades, the origin of the membrane happened by the evolution of a very specific set of lipid-binding proteins (Sojo, 2019). Until nowadays, we know that biological membranes are made of lipids but also proteins. Proteins are broadly recognized as the second major component of plasma membranes, some of them being peripheral and acting just in one membrane side, while others being integral and crossing both sides of the phospholipids layer. Aiming to invert the logic of thought we follow the argument that membranes started with the evolution of lipid-binding proteins (Sojo, 2019). This becomes clear once again if we consider that biology is basically ribonucleoproteic (Farias and Prosdocimi, 2022). The components of lipids and carbohydrates, although they are nowadays of striking relevance to the perfect work of cells, they have been added later than the RNP core. Their addition created new layers of macrocode connections and allowed the origin of cells. The membranes of Bacteria are formed by standard phospholipids and bound by an ester chemical linkage to a molecule of glycerol that is phosphorylated at the C3 (glycerol-3-phosphate). On the other hand, the membranes of archaea are formed by branched, isoprenoid phospholipids bound by an ether linkage to a molecule of glycerol that is phosphorylated at C1 (glycerol-1-phosphate) (Koga, 2011). Besides, the biochemical routes to make those membranes are different. This fact suggests the parallel evolution of progenote LUCA populations that achieved cellularity by two alternative routes, both starting with lipid-binding proteins (Farias et al., 2021a,b). Lipid-binding proteins were therefore the doors through which phospholipid cellularity has emerged in biology.

Also, we have noticed that the enzymes responsible to perform DNA synthesis in Bacteria and Archaea are different. Although it is out of the scope here to tackle this question, it seems likely that DNA synthesis has been originated in viruses as proposed earlier (Forterre, 2005; Di Giulio, 2021). Possibly two different viral lineages of progenotes evolved these alternative routes to produced DNA as it is clear that DNA is a more stable media to store genomic information than RNA. This is mainly because its ribose is more stable and less reactive (deoxyribose) and it presents a double-helical structure, holding tight a backup of its own molecular information. At any time that the cellularized versions of LUCA meet the viral populations capable to produce DNA, the symbiosis would happen and benefit both Castor (the bacterial cell) and Polux (the archaeal cell). Thus, the acquisition of DNA synthesis was possibly the very last path that RNA genomes encapsulated in archaea and bacteria cells would acquire to produce cellular entities like the ones we know today. In any case, they could also have happened before cellularization, but later than the formation of the progenote LUCA on which the other pathways have already been assembled (Farias et al., 2021a,b).

Thus, there has never been a cellular LUCA, for which the denomination of last universal cellular ancestor or LUCellA is probably incorrect (Kim and Caetano-Anollés, 2012). The LUCA should be therefore understood as a derived progenote on which most biochemical routes have been accumulated, but neither DNA synthesis nor lipid synthesis.

10. Conclusion

In order to build a big picture for the origin of life in Earth, we need to focus on the most important molecular transitions happening from FUCA to LUCA. There will be details in the study of the origin of life that have little importance when one aims to build a broad scenario for the early days of living organisms, even if putative. More than anything, we aimed here to provide a biologically oriented overview for the origin of life. This seemed more important than discussing specific chemical characteristics of life in prebiotic Earth that are most likely contingent, such as the precise chemical scenarios that produced nucleobases, phosphoric acids, and riboses (Prosdocimi et al., 2022); the fact that this happened in either clay, water or somewhere else; or the fact that the amino acids in biology are levorotatory instead of dextrorotatory. Thus, instead of focusing on the chemical nature of molecules, we decided to focus on their biological nature.

Also, even if the scenario proposed here presents a significant number of gaps, those missing links can be further evaluated by new research that may discover: (i) how to make nucleotides prebiotically; (ii) how RNAs and tRNAs could be formed; (iii) how the proto-PTC has been built; (iv) how the genetic code has been structured; (v) how progenotes could live and reproduced as "naked" molecules of RNA; (vi) how peptides started to bind molecules in the prebiotic soup; (vii) how biochemical pathways evolved from those bindings; (viii) how genomes got bigger by the symbiotic relationship and concatenation of progenotes' genetic information; (ix) how the progenote version of LUCA has been formed; (x) how the first virion capsids have been formed; (xi) how

virion capsids evolved; (xii) how lipid-binding proteins produced phospholipid membranes; (xiii) how DNA synthesis have been invented; and, finally, (xiv) how DNA-based cells of bacteria and archaea have been constituted.

Each of these steps define complex research programs that should be seriously evaluated by the community interested in the origin of life. We look into the future to close those epistemological gaps and build a better scenario to understand this amazing topic that is the origin of life in Earth.

Declaration of competing interest

We declare that none people or organizations inappropriately influenced our work.

Acknowledgements

We are grateful to our friends, collaborators and researchers in the origin of life community for continuing discussions in meetings and seminars that allowed us to update and better ellaborate our model. In that sense, we would like to thank Marco V José, Myriam Palacios-Perez, Sohan Jheeta, Martin Dominik, Romeu Guimarães, Tuane Vieira, Juliana Cortines, Loren Williams, Ada Yonath, Antonio Lazcano, and Gustavo Caetano-Anollés. We also thank FAPERJ and CNPq for providing research productivity fellowships for STF and FP (CNE E-26/200.940/2022 and 306346/2022-2).

References

Barbieri, M., 2003. The Organic Codes: an Introduction to Semantic Biology. Cambridge University Press, Cambridge.

Barbieri, M., 2014. From biosemiotics to code biology. Biol Theory 9, 239–249. https://doi.org/10.1007/s13752-013-0155-6.

Barbieri, M., 2012. Code biology – a new science of life. Biosemiotics 5, 411–437.

Barbieri, M., 2013. The paradigms of biology. Biosemiotics 6, 33–59.

Barbieri, M., 2016. From the common ancestor to the first cells: the code theory. Biol. Theory. 11, 102–112.

Becker, S., Schneider, C., Okamura, H., Crisp, A., Amatov, T., Dejmek, M., Carell, T., 2018. Wet-dry cycles enable the parallel origin of canonical and non-canonical nucleosides by continuous synthesis. Nat. Commun. 11 (1), 163, 9.

Belousoff, M.J., Davidovich, C., Zimmerman, E., Caspi, Y., Wekselman, I., Rozenszajn, L., Shapira, T., Sade-Falk, O., Taha, L., Bashan, A., Weiss, M.S., Yonath, A., 2010. Ancient machinery embedded in the contemporary ribosome. Biochem. Soc. Trans. 38 (2), 422–427. https://doi.org/10.1042/BST0380422. Apr.

Bose, T., Fridkin, G., Davidovich, C., Krupkin, M., Dinger, N., Falkovich, A.H., Peleg, Y., Agmon, I., Bashan, A., Yonath, A., 2022. Origin of life: protoribosome forms peptide bonds and links RNA and protein dominated worlds. Nucleic Acids Res. 50 (4), 1815–1828. https://doi.org/10.1093/nar/gkac052.

Burton, A.S., Stern, J.C., Elsila, J.E., Glavin, D.P., Dworkin, J.P., 2012. Understanding prebiotic chemistry through the analysis of extraterrestrial amino acids and nucleobases in meteorites. Chem. Soc. Rev. 21 (16), 5459–5472, 41.

Caetano-Anollés, D., Caetano-Anollés, G., 2016. Piecemeal buildup of the genetic code, ribosomes, and genomes from primordial tRNA building blocks. Life 6 (4), 2. https://doi.org/10.3390/life6040043.

Caetano-Anollés, G., 2015. Ancestral insertions and expansions of rRNA do not support an origin of the ribosome in its peptidyl transferase center. Apr J. Mol. Evol. 80 (3–4), 162–165. https://doi.org/10.1007/s00239-015-9677-9. Epub 2015 Apr 12. PMID: 25864085; PMCID: PMC4555209.

Caetano-Anollés, G., Seufferheld, M.J., 2013. The coevolutionary roots of biochemistry and cellular organization challenge the RNA world paradigm. J. Mol. Microbiol. Biotechnol. 23 (1–2), 152–177. https://doi.org/10.1159/000346551.

Claverie, J.M., Ogata, H., Audic, S., Abergel, C., Suhre, K., Fournier, P.E., 2006. Mimivirus and the emerging concept of "giant" virus. Apr Virus Res. 117 (1), 133–144. https://doi.org/10.1016/j.virusres.2006.01.008. PMID: 16469402.

Crick, F.H.C., 1968. The origin of the genetic code. J. Mol. Biol. 38 (Issue 3), 367–379. https://doi.org/10.1016/0022-2836(68)90392-6. ISSN 0022-2836.

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 13 (1), 872–887, 5.

Davidovich, C., Belousoff, M., Bashan, A., Yonath, A., 2009. The evolving ribosome: from non-coded peptide bond formation to sophisticated translation machinery. Res. Microbiol. 160 (7), 487–492. https://doi.org/10.1016/j.resmic.2009.07.004. Sep.

Delaye, L., Becerra, A., Lazcano, A., 2005. The last common ancestor: what's in a name? Orig. Life Evol. Biosph. 35 (6), 537–554. https://doi.org/10.1007/s11084-005-5760-3.

Delidovich, I.V., Simonov, A.N., Taran, O.P., Parmon, V.N., 2014. Catalytic formation of monosaccharides: from the formose reaction towards selective synthesis. ChemSusChem 7 (7), 1833–1846. https://doi.org/10.1002/cssc.201400040.

Demongeot, J., Seligmann, H., 2022. Evolution of small and large ribosomal RNAs from accretion of tRNA subelements. Biosystems 222, 104796. https://doi.org/10.1016/j.biosystems.2022.104796.

Di Giulio, M., 1997. On the RNA world: evidence in favor of an early ribonucleopeptide world. J. Mol. Evol. 45 (6), 571–578. https://doi.org/10.1007/pl00006261.

Di Giulio, M., 2021. The late appearance of DNA, the nature of the LUCA and ancestors of the domains of life. Apr Biosystems 202, 104330. Epub 2020 Dec 19. PMID: 33352234.

Di Giulio, M., 2008. The origin of genes could be polyphyletic. Dec 15 Gene 426 (1–2), 39–46. https://doi.org/10.1016/j.gene.2008.07.024. Epub 2008 Jul 29. PMID: 18706983.

Dworkin, J.P., Lazcano, A., Miller, S.L., 2003. The roads to and from the RNA world. J. Theor. Biol. 7 (1), 127–134, 222.

Eigen, M., 1971. Self organization of matter and the evolution of biological macromolecules. Naturwissenschaften 58, 465–523. https://doi.org/10.1007/BF00623322.

Emmeche, C., 1998. Defining life as a semiotic phenomenon. Cybern. Hum. Knowing 5 (1), 3–17, 1998.

Farias, S.T., Prosdocimi, F., 2022. RNP-world: The ultimate essence of life is a ribonucleoprotein process. Genet. Mol. Biol. 45 (3 Suppl 1), e20220127 https://doi.org/10.1590/1678-4685-GMB-2022-0127.

Farias, S.T., Prosdocimi, F., Caponi, G., 2021a. Organic codes: a unifying concept for life. Acta Biotheor. 69 (4), 769–782. https://doi.org/10.1007/s10441-021-09422-2.

Farias, S.T., Gaudêncio Rêgo, T., José, M.V., 2017. Peptidyl transferase center and the emergence of the translation system. Life 7 (2), 21. https://doi.org/10.3390/life7020021. Apr 25.

Farias, S.T., Jheeta, S., Prosdocimi, F., 2019. Viruses as a survival strategy in the armory of life. Oct 14 Hist. Philos. Life Sci. 41 (4), 45. https://doi.org/10.1007/s40656-019-0287-5. PMID: 31612293.

Farias, S.T., José, M.V., 2020. Transfer RNA: the molecular demiurge in the origin of biological systems. Prog. Biophys. Mol. Biol. 153, 28–34. https://doi.org/10.1016/j.pbiomolbio.2020.02.006. Jul.

Farias, S.T., Rêgo, T.G., José, M.V., 2014a. Origin and evolution of the peptidyl transferase center from proto-tRNAs. FEBS Open Bio 4, 175–178. https://doi.org/10.1016/j.fob.2014.01.010. Published 2014 Feb 8.

Farias, S.T., Rêgo, T.G., José, M.V., 2014b. Origin and evolution of the peptidyl transferase center from proto-tRNAs. FEBS Open Bio 4, 175–178. https://doi.org/10.1016/j.fob.2014.01.010. Feb 8.

Farias, S.T., Rêgo, T.G., José, M.V., 2016. tRNA core hypothesis for the transition from the RNA world to the ribonucleoprotein world. Mar 23 Life 6 (2), 15. https://doi.org/10.3390/life6020015. PMID: 27023615; PMCID: PMC4931452.

Farias, S.T., Rêgo, T.G., José, M.V., 2021b. Origin of the 16S ribosomal molecule from ancestor tRNAs. J. Mol. Evol. 89 (4–5), 249–256. https://doi.org/10.1007/s00239-021-10002-8.

Forterre, P., 2005. The two ages of the RNA world, and the transition to the DNA world: a story of viruses and cells. Sep-Oct Biochimie 87 (9–10), 793–803. https://doi.org/10.1016/j.biochi.2005.03.015. Epub 2005 Apr 12. PMID: 16164990.

Forterre, P., Prangishvili, D., 2009. The origin of viruses. Sep Res. Microbiol. 160 (7), 466–472. https://doi.org/10.1016/j.resmic.2009.07.008. Epub 2009 Jul 30. PMID: 19647075.

Gilbert, W., 1986. Origin of life: the RNA world. Nature 319, 618. https://doi.org/10.1038/319618a0.

Gomes, E., Shorter, J., 2019. The molecular language of membraneless organelles. May 3 J. Biol. Chem. 294 (18), 7115–7127. https://doi.org/10.1074/jbc.TM118.001192. Epub 2018 Jul 25. PMID: 30045872; PMCID: PMC6509512.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., Altman, S., 1983. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. Cell 35 (3 Pt 2), 849–857.

Guimarães, R.C., Moreira, C.H., de Farias, S.T., 2008. A self-referential model for the formation of the genetic code. Theor. Biosci. 127 (3), 249–270. https://doi.org/10.1007/s12064-008-0043-y. Aug.

Harris, H.M.B., Hill, C., 2021. A place for viruses on the tree of life. Jan 14 Front. Microbiol. 11, 604048. https://doi.org/10.3389/fmicb.2020.604048. PMID: 33519747; PMCID: PMC7840587.

Hofmeyr, J.S., 2018. The first Special Issue on code biology - a bird's-eye view. Biosystems 164, 11–15. https://doi.org/10.1016/j.biosystems.2017.12.007.

Hollis, J.M., Lovas, F.J., Jewell, P.R., 2000. Interstellar glycolaldehyde: the first sugar. ApJL 540, L107–L110.

Ingar, A.A., Luke, R.W., Hayter, B.R., Sutherland, J.D., 2003. Synthesis of cytidine ribonucleotides by stepwise assembly of the heterocycle on a sugar phosphate. Chembiochem 4 (6), 504–507. https://doi.org/10.1002/cbic.200300554.

Johnson, A.P., Cleaves, H.J., Dworkin, J.P., Glavin, D.P., Lazcano, A., Bada, J.L., 2008. The Miller volcanic spark discharge experiment. Science 17 (5900), 322, 404.

Kannan, L., Li, H., Rubinstein, B., Mushegian, A., 2013. Models of gene gain and gene loss for probabilistic reconstruction of gene content in the last universal common ancestor of life. Dec 19 Biol. Direct 8, 32. https://doi.org/10.1186/1745-6150-8-32. PMID: 24354654; PMCID: PMC3892064.

Keller, M.A., Turchyn, A.V., Ralser, M., 2014. Non-enzymatic glycolysis and pentose phosphate pathway-like reactions in a plausible Archean ocean. Mol. Syst. Biol. 10 (4), 725. https://doi.org/10.1002/msb.20145228. Published 2014 Apr 25.

Kim, H.J., Kim, J., 2019. A prebiotic synthesis of canonical pyrimidine and purine ribonucleotides. Astrobiology 19 (5), 669–674. https://doi.org/10.1089/ast.2018.1935.

Kim, K.M., Caetano-Anollés, G., 2012. The evolutionary history of protein fold families and proteomes confirms that the archaeal ancestor is more ancient than the ancestors of other superkingdoms. Jan 27 BMC Evol. Biol. 12, 13. https://doi.org/10.1186/1471-2148-12-13. PMID: 22284070; PMCID: PMC3306197.

Kim, K.M., Sung, S., Caetano-Anollés, G., Han, J.Y., Kim, H., 2008. An approach of orthology detection from homologous sequences under minimum evolution. Oct Nucleic Acids Res. 36 (17), e110. https://doi.org/10.1093/nar/gkn485. Epub 2008 Aug 1. PMID: 18676448; PMCID: PMC2553584.

Koga, S., Williams, D.S., Perriman, A.W., Mann, S., 2011. Peptide-nucleotide microdroplets as a step towards a membrane-free protocell model. Nat. Chem. 3 (9), 720–724. https://doi.org/10.1038/nchem.1110.PMID:21860462. Aug 7.

Koga, Y., 2011. Early evolution of membrane lipids: how did the lipid divide occur? Mar J. Mol. Evol. 72 (3), 274–282. https://doi.org/10.1007/s00239-011-9428-5. Epub 2011 Jan 23. PMID: 21259003.

Koonin, E.V., 2003. Comparative genomics, minimal gene-sets and the last universal common ancestor. Nat. Rev. Microbiol. 1 (2), 127–136. https://doi.org/10.1038/nrmicro751.PMID:15035042. Nov.

Koonin, E.V., Martin, W., 2005. On the origin of genomes and cells within inorganic compartments. Trends Genet. 21 (12), 647–654. https://doi.org/10.1016/j.tig.2005.09.006.

Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E., Cech, T.R., 1982. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. Cell 31 (1), 147–157. https://doi.org/10.1016/0092-8674(82)90414-7.

Krupovic, M., Dolja, V.V., Koonin, E.V., 2019. Origin of viruses: primordial replicators recruiting capsids from hosts. Nat. Rev. Microbiol. 17 (7), 449–458. https://doi.org/10.1038/s41579-019-0205-6.

Lamour, S., Pallmann, S., Haas, M., Trapp, O., 2019. Prebiotic sugar formation under nonaqueous conditions and mechanochemical acceleration. Life 9 (2), 52. https://doi.org/10.3390/life9020052. Published 2019 Jun 20.

Lanier, K.A., Petrov, A.S., Williams, L.D., 2017. The central symbiosis of molecular biology: molecules in mutualism. J. Mol. Evol. 85 (1–2), 8–13. https://doi.org/10.1007/s00239-017-9804-x.

Lazcano, A., Bada, J.L., 2003. The 1953 Stanley L. Miller experiment: fifty years of prebiotic organic chemistry. Orig. Life Evol. Biosph. 33 (3), 235–242.

Lazcano, A., Forterre, P., 1999. The molecular search for the last common ancestor. Oct J. Mol. Evol. 49 (4), 411–412. PMID: 10485998.

Le Vay, K., Mutschler, H., 2019. The difficult case of an RNA-only origin of life. Emerg Top Life Sci 3 (5), 469–475. https://doi.org/10.1042/ETLS20190024.

Levy, M., Miller, S.L., Brinton, K., Bada, J.L., 2000. Prebiotic synthesis of adenine and amino acids under Europa-like conditions. Icarus 145 (2), 609–613.

Lopez, A., Fiore, M., 2019. Investigating prebiotic protocells for A comprehensive understanding of the origins of life: a prebiotic systems chemistry perspective. Life 7 (2), 9. Pii: E49.

Mat, W.K., Xue, H., Wong, J.T., 2008. The genomics of LUCA. Front. Biosci. 13, 5605–5613. https://doi.org/10.2741/3103.PMID:18508609. May 1.

Meinert, C., Myrgorodska, I., de Marcellus, P., Buhse, T., Nahon, L., Hoffmann, S.V., d'Hendecourt Lle, S., Meierhenrich, U.J., 2016. Ribose and related sugars from ultraviolet irradiation of interstellar ice analogs. Science 8 (6282), 208–212, 352.

Menor-Salván, C., Marín-Yaseli, M.R., 2013. A new route for the prebiotic synthesis of nucleobases and hydantoins in water/ice solutions involving the photochemistry of acetylene. Chemistry 10 (20), 6488–6497, 19.

Miller, S.L., 1953. A production of amino acids under possible primitive earth conditions. Science 117 (3046), 528–529. https://doi.org/10.1126/science.117.3046.528.PMID:13056598. May 15.

Muchowska, K.B., Chevallot-Beroux, E., Moran, J., 2019. Recreating ancient metabolic pathways before enzymes. Bioorg. Med. Chem. 27 (12), 2292–2297. https://doi.org/10.1016/j.bmc.2019.03.012.

Muchowska, K.B., Varma, S.J., Moran, J., 2019b. Synthesis and breakdown of universal metabolic precursors promoted by iron [published correction appears in Nature. 2021 May;593(7857):E3]. Nature 569 (7754), 104–107. https://doi.org/10.1038/s41586-019-1151-1.

Mushegian, A., 2008. Gene content of LUCA, the last universal common ancestor. Front. Biosci. 13, 4657–4666. https://doi.org/10.2741/3031.PMID:18508537. May 1.

Nasir, A., Caetano-Anollés, G., 2015. A phylogenomic data-driven exploration of viral origins and evolution. Sci. Adv. 1 (8), e1500527 https://doi.org/10.1126/sciadv.1500527.

Orgel, L.E., 1968. Evolution of the genetic apparatus. J. Mol. Biol. 38 (3), 381–393. https://doi.org/10.1016/0022-2836(68)90393-8.

Orgel, L.E., 2004. Prebiotic chemistry and the origin of the RNA world. Crit. Rev. Biochem. Mol. Biol. 39 (2), 99–123.

Oparin, A.I., 1924. Proiskhozhdenie Zhizni. Izd. Moskovskii Rabochii, Moscow (in Russian). Translated as "*The Origin and Development of Life* (NASA TTF-488). Washington: D.C.L GPO, 1968".

Ouzounis, C.A., Kunin, V., Darzentas, N., Goldovsky, L., 2006. A minimal estimate for the gene content of the last universal common ancestor–exobiology from a terrestrial perspective. Jan-Feb Res. Microbiol. 157 (1), 57–68. https://doi.org/10.1016/j.resmic.2005.06.015. Epub 2005 Dec 19. PMID: 16431085.

Pasek, M., Lauretta, D., 2008. Extraterrestrial flux of potentially prebiotic C, N, and P to the early Earth. Orig. Life Evol. Biosph. 38 (1), 5–21.

Petrov, A.S., Bernier, C.R., Hsiao, C., Norris, A.M., Kovacs, N.A., Waterbury, C.C., Stepanov, V.G., Harvey, S.C., Fox, G.E., Wartell, R.M., Hud, N.V., Williams, L.D., 2014. Evolution of the ribosome at atomic resolution. Jul 15 Proc. Natl. Acad. Sci. U. S. A. 111 (28), 10251–10256. https://doi.org/10.1073/pnas.1407205111. Epub 2014 Jun 30. PMID: 24982194; PMCID: PMC4104869.

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. U. S. A. 112 (50), 15396–15401. https://doi.org/10.1073/pnas.1509761112. Dec 15.

Petrov, A.S., Williams, L.D., 2015. The ancient heart of the ribosomal large subunit: a response to Caetano-Anolles. Apr J. Mol. Evol. 80 (3–4), 166–170. https://doi.org/10.1007/s00239-015-9678-8. Epub 2015 Apr 16. PMID: 25877522.

Powner, M.W., Anastasi, C., Crowe, M.A., Parkes, A.L., Raftery, J., Sutherland, J.D., 2007. On the prebiotic synthesis of ribonucleotides: photoanomerisation of cytosine nucleosides and nucleotides revisited. Chembiochem 8 (10), 1170–1179. https://doi.org/10.1002/cbic.200700098.

Powner, M.W., Gerland, B., Sutherland, J.D., 2009. Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. Nature 459 (7244), 239–242. https://doi.org/10.1038/nature08013.

Prosdocimi, F., Farias, S.T., 2019. A emergência dos sistemas biológicos: uma visão molecular sobre a origem da vida. ArteComCiencia, 1 ed. Rio de Janeiro. https://www.amazon.com.br/dp/B08WTHJPM4.

Prosdocimi, F., Farias, S.T., 2020. Vírus: reinterpretando a história natural e sua importância ecológica (in Portuguese). Helius 3 (2), 3 1791–1811.

Prosdocimi, F., Farias, S.T., 2021. Life and living beings under the perspective of organic macrocodes. Biosystems 206, 104445. https://doi.org/10.1016/j.biosystems.2021.104445.

Prosdocimi, F., Farias, S.T., 2022. Entering the labyrinth: a hypothesis about the emergence of metabolism from protobiotic routes. Oct Biosystems 220, 104751. https://doi.org/10.1016/j.biosystems.2022.104751. Epub 2022 Aug 5. PMID: 35940497.

Prosdocimi, F., de Farias, S.T., José, M.V., 2022. Prebiotic chemical refugia: multifaceted scenario for the formation of biomolecules in primitive Earth. Nov Theor. Biosci. 141 (4), 339–347. https://doi.org/10.1007/s12064-022-00377-7. Epub 2022 Aug 30. PMID: 36042123.

Prosdocimi, F., Farias, S.T., Cortines, J., José, M.V., 2021b. The misleading history of VIRUSes. Preprints, 2021040743. https://doi.org/10.20944/preprints202104.0743.v1.

Prosdocimi, F., Jheeta, S., Torres de Farias, S., 2018. Conceptual challenges for the emergence of the biological system: cell theory and self-replication. Oct Med. Hypotheses 119, 79–83. https://doi.org/10.1016/j.mehy.2018.07.029. PMID: 30122496.

Prosdocimi, F., José, M.V., de Farias, S.T., 2021a. The theory of chemical symbiosis: a margulian view for the emergence of biological systems (origin of life). Acta Biotheor. 69, 67–78. https://doi.org/10.1007/s10441-020-09388-7.

Prosdocimi, F., José, M., Farias, S., 2019. The first universal common ancestor (FUCA) as the earliest ancestor of LUCA's (last UCA) lineage. Chapter 3. In: Book: Evolution, Origin of Life, Concepts and Methods. Pierre Pontarotti. Springer. https://doi.org/10.1007/978-3-030-30363-1_3. October.

Prosdocimi, F., Zamudio, G.S., Palacios-Pérez, M., Torres de Farias, S., V José, M., 2020. The ancient history of peptidyl transferase center formation as told by conservation and information analyses. Aug 5 Life 10 (8), 134. https://doi.org/10.3390/life10080134. PMID: 32764248; PMCID: PMC7459865.

Ralser, M., 2018. An appeal to magic? The discovery of a non-enzymatic metabolism and its role in the origins of life. Biochem. J. 475 (16), 2577–2592. https://doi.org/10.1042/BCJ20160866. Published 2018 Aug 30.

Robertson, M.P., Joyce, G.F., 2012. The origins of the RNA world. Cold Spring Harbor Perspect. Biol. 1 (5), a003608 https://doi.org/10.1101/cshperspect.a003608, 4.

Root-Bernstein, M., Root-Bernstein, R., 2015. The ribosome as a missing link in the evolution of life. J. Theor. Biol. 367, 130–158. https://doi.org/10.1016/j.jtbi.2014.11.025.

Rivas, M., Fox, G.E., 2023. How to build a protoribosome: structural insights from the first protoribosome constructs that have proven to be catalytically active. RNA (N. Y.) 29 (3), 263–272. https://doi.org/10.1261/rna.079417.122.

Seufferheld, M.J., Caetano-Anollés, G., 2013. Phylogenomics supports a cellularly structured urancestor. J. Mol. Microbiol. Biotechnol. 23 (1–2), 178–191. https://doi.org/10.1159/000346552. Epub 2013 Apr 18. PMID: 23615204.

Sojo, V., 2019. Why the lipid divide? membrane proteins as drivers of the split between the lipids of the three domains of life. Bioessays 41 (5), e1800251. https://doi.org/10.1002/bies.201800251.

Spang, A., Saw, J.H., Jørgensen, S.L., Zaremba-Niedzwiedzka, K., Martijn, J., Lind, A.E., van Eijk, R., Schleper, C., Guy, L., Ettema, T.J.G., 2015. Complex archaea that bridge the gap between prokaryotes and eukaryotes. May 14 Nature 521 (7551), 173–179. https://doi.org/10.1038/nature14447.Epub2015May6. PMID: 25945739; PMCID: PMC4444528.

Schwartz, A.W., 1995. The RNA World and its origins. Planet. Space Sci. 43 (1–2), 161–165.

Suttle, C.A., 2013. Viruses: unlocking the greatest biodiversity on Earth. Genome 56 (10), 542–544. https://doi.org/10.1139/gen-2013-0152.

Vitas, M., Dobovišek, A., 2018. In the beginning was a mutualism - on the origin of translation. Orig. Life Evol. Biosph. 48, 223–243.

Watson, J.D., Crick, F.H., 1953. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Apr 25 Nature 171 (4356), 737–738. https://doi.org/10.1038/171737a0. PMID: 13054692.

Weiss, MC, Sousa, FL, Mrnjavac, N, Neukirchen, S, Roettger, M, Nelson-Sathi, S, Martin, WF, 2016 Jul 25. The physiology and habitat of the last universal common ancestor. Nat Microbiol 1 (9), 16116. https://doi.org/10.1038/nmicrobiol.2016.116. PMID: 27562259.

Woese, C.R., 1965. On the evolution of the genetic code. Proc. Natl. Acad. Sci. U. S. A. 54 (6), 1546–1552. https://doi.org/10.1073/pnas.54.6.1546.

Woese, C., 1998. The universal ancestor. Proc. Natl. Acad. Sci. U. S. A. 95 (12), 6854–6859. https://doi.org/10.1073/pnas.95.12.6854.

Woese, C.R., Kandler, O., Wheelis, M.L., 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc. Natl. Acad. Sci. U. S. A. 87 (12), 4576–4579. https://doi.org/10.1073/pnas.87.12.4576.

Xu, D., Wang, Y., 2021. Protein-free ribosomal RNA scaffolds can assemble poly-lysine oligos from charged tRNA fragments. Biochem. Biophys. Res. Commun. 544, 81–85. https://doi.org/10.1016/j.bbrc.2021.01.036.

Yadav, M., Kumar, R., Krishnamurthy, R., 2020. Chemistry of abiotic nucleotide synthesis. Chem. Rev. 120 (11), 4766–4805. https://doi.org/10.1021/acs.chemrev.9b00546.

Zamudio, G.S., Prosdocimi, F., de Farias, S.T., José, M.V., 2019. A neutral evolution test derived from a theoretical amino acid substitution model. J. Theor. Biol. 467, 31–38. https://doi.org/10.1016/j.jtbi.2019.01.027.

Zaremba-Niedzwiedzka, K., Caceres, E.F., Saw, J.H., Bäckström, D., Juzokaite, L., Vancaester, E., Seitz, K.W., Anantharaman, K., Starnawski, P., Kjeldsen, K.U., Stott, M.B., Nunoura, T., Banfield, J.F., Schramm, A., Baker, B.J., Spang, A., Ettema, T.J., 2017. Asgard archaea illuminate the origin of eukaryotic cellular complexity. Jan 19 Nature 541 (7637), 353–358. https://doi.org/10.1038/nature21031. Epub 2017 Jan 11. PMID: 28077874.

Zhao, Y.G., Zhang, H., 2020. Phase separation in membrane biology: the interplay between membrane-bound organelles and membraneless condensates. Oct 12 Dev. Cell 55 (1), 30–44. https://doi.org/10.1016/j.devcel.2020.06.033. Epub 2020 Jul 28. PMID: 32726575.