

# Life before LUCA\*

Athel Cornish-Bowden and María Luz Cárdenas

Aix Marseille Univ, CNRS, BIP, IMM, Marseille, France

## AUTHORS' CONTACT INFORMATION

*email* acornish@imm.cnrs.fr

cardenas@imm.cnrs.fr

*telephone* + 33 491 16 41 38

## ARTICLE INFO

*Keywords:* Lynn Sagan, Lynn Margulis, LUCA, cenancestor, last universal common ancestor, origin of life, definition of life

**NOTE.** This file is printed from the final accepted version of the paper. The PDF file typeset by the Journal will be available later.

---

## ABSTRACT

We see the last universal common ancestor of all living organisms, or LUCA, at the evolutionary separation of the Archaea from the Eubacteria, and before the symbiotic event believed to have led to the Eukarya. LUCA is often implicitly taken to be close to the origin of life, and sometimes this is even stated explicitly. However, LUCA already had the capacity to code for many proteins, and had some of the same bioenergetic capacities as modern

---

\*This paper is dedicated to the memory of Lynn Sagan (Margulis), and especially of her paper "On the origin of mitosing cells".

organisms. An organism at the origin of life must have been vastly simpler, and this invites the question of how to define a living organism. Even if acceptance of the giant viruses as living organisms forces the definition of LUCA to be revised, it will not alter the essential point that LUCA should be regarded as a recent player in the evolution of life.

---

## 1. Introduction

In general I avoid the last 3 million years of evolution and any other studies that require detailed knowledge of mammalian, including human, biology. Why? Because political bias, hearsay and gossip are inevitable whereas in the first part of the evolution story (from 3800 until 3 million years ago) politics intervenes far less obtrusively. In pursuit of the story of life and its effects on planet Earth one can be more honest if the earliest stages of evolution are the objects of study (Margulis, 2007).

Lynn Margulis, pictured in Fig. 1 as she was in 2009, introduced a hypothesis about the origin of the Eukarya (Sagan, 1967) that brought about a major change in biological thought. Although little of what she wrote was completely new, as Lane (2017) discusses elsewhere in this special issue, it was the paper that forced biologists to address her ideas, and it continues to be influential after half a century. Its publication in this Journal stimulated considerable discussion, as did many of her other papers (Banerjee and Margulis, 1973; Margulis and Lovelock, 1974; Chapman and Margulis, 1998; Margulis et al., 2000; Chapman et al., 2000; Margulis, 2001; Margulis et al., 2006, 2009). Even if some details remain to be resolved, as discussed elsewhere in this special issue, her basic concept is now generally recognized to be correct. She continued to argue until the end of her life for the crucial role of symbiosis in evolution, and the title of an interview in the Spanish magazine SEBBM (Margulis, 2009) made this clear: “Symbiosis is the source of innovation in evolution”.

Margulis (2007) preferred to study what she regarded as the earliest events in evolution, most particularly the divergence from the Archaea that led to the Eukarya. We will argue in this paper, however, that this divergence, and the earlier divergence of the Eubacteria and the Archaea from one another, were late events in the development



Figure 1: Lynn Margulis in 2009. Reproduced from Margulis (2009) with permission of the Sociedad Española de Bioquímica y Biología Molecular.

of life after its original appearance, and that self-sustaining systems must have existed much earlier.

The last universal common ancestor of known life is often called LUCA or the *cen-ancestor*. The L in the name is important: it was not the *first* organism, but the *last* before the bifurcation that led to all modern organisms.

We do not doubt that living organisms existed long before LUCA. However, it is sometimes suggested to be an organism close to the origin of life, as in the titles of some recent papers: “How did LUCA make a living? Chemiosmosis in the origin of life” (Lane et al., 2010); “Origin of life: LUCA and extracellular membrane vesicles (EMVs)” (Gill and Forterre, 2015); “The routes of emergence of life from LUCA during the RNA and viral world: a conspectus” (Jheeta, 2015). An explicit equating of LUCA with the first organism is in the following account of the “RNA World” hypothesis (Wieczorek, 2012):

At some point, the RNA protocell “invented” DNA for storing genetic information and proteins to perform catalysis with greater efficiency. At this point we arrive at a form of life that is ancestral to all modern life on Earth — LUCA.

Likewise, in a more recent paper Weiss et al. (2016) describe LUCA as the link between abiotic systems and the first traces of life:

The last universal common ancestor (LUCA) is an inferred evolutionary intermediate that links the abiotic phase of Earth’s history with the first

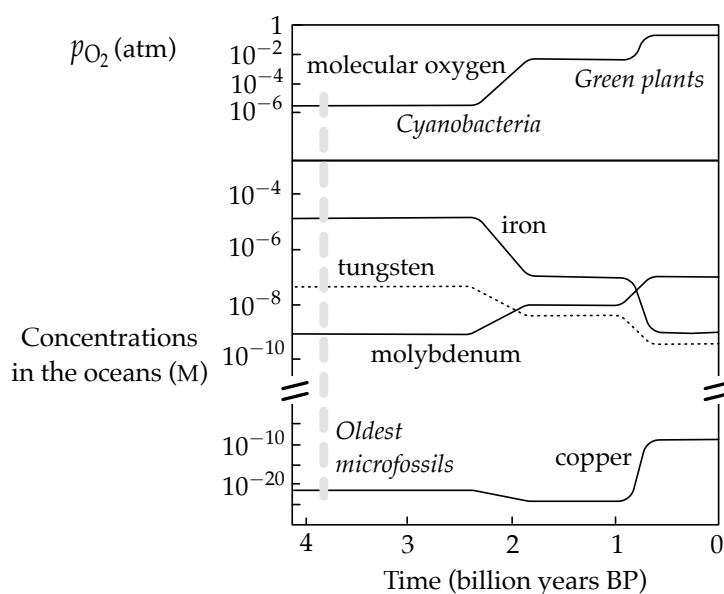


Figure 2: Changed availability of dissolved metals in the oceans with the increase in the partial pressure of molecular oxygen in the atmosphere. The curves are based on data of Anbar (2008), apart from the one for tungsten, which is drawn as an arbitrary reflection of that for molybdenum. Despite the great chemical similarity between these two elements, they differ in their solubility, and tungsten became less soluble in the oceans over the same period that molybdenum became more soluble (Pushie et al., 2014).

traces of microbial life in rocks that are 3.8–1.5 billion years of age.

## 2. Difficulties for defining LUCA

As Mariscal and Doolittle (2015) discuss, various definitions of LUCA have been proposed, including ones that postulate a community of different organisms (Acevedo-Rocha et al., 2013) rather than just one. Glansdorff et al. (2008) discussed many different views of LUCA, leading them to describe it as a “complex community of protoeukaryotes”.

LUCA is often considered to be at the point of bifurcation of the Eubacteria and the Archaea, as shown in Fig. 3. However, some authors, such as Martin et al. (2016) placed LUCA much earlier, following the suggestion (Martin and Russell, 2003) that the original separation between Archaea and Eubacteria was a divergence between different communities of cells in the prebiotic period, with “cells” bounded by minerals

rather than lipid membranes. Di Giulio (2011) likewise argued that LUCA is so ancient that it existed before life began, i.e. that it must be a *progenote*, but he also listed several authors, including Becerra et al. (2007) and Tuller et al. (2010), who took a variety of different views. Our own view is close to that of Tuller et al. (2010) that “LUCA appears to have been bacterial-like and had a genome size similar to the genome sizes of many extant organisms.” This means, of course, that we see LUCA as a living organism.

A major problem for unravelling the relationships between the three kingdoms of life has been to explain why the lipid membranes of Archaea, formed from *sn*-glycerol 1-phosphate, are so different from those of Eubacteria and Eukaryotes, which use its stereoisomer *sn*-glycerol 3-phosphate. Moreover, the enzymes that produce the two isomers are not homologous. Peretó et al. (2004) have examined the question in detail, and suggest that the original lipid membranes, and the pathways that produced them, were not stereospecific, but became so as more efficient enzymes evolved, and different isomers were subsequently lost in different lineages. This would imply that when LUCA existed natural selection had not had sufficient time to select highly stereospecific glycerol phosphate dehydrogenases, so it would not be surprising if there were mixed products. It was long thought that lipids based on *sn*-glycerol 1-phosphate were confined to Archaea, but glycerol 1-phosphate dehydrogenase exists in *Bacillus subtilis* (Guldan et al., 2008). Furthermore, *Bacillus subtilis* contains an enzyme, heptaprenylglyceryl phosphate synthase, that is clearly homologous to geranylgeranyl glyceryl phosphate synthase, which is involved in the production of Archaeal lipids (Peterhoff et al., 2012).

A case that can shed light on the problems raised by the differences of stereospecificity is that of the two types of specificity of lactate dehydrogenase in different organisms (Cristescu et al., 2008). Although both types of enzyme have achiral pyruvate as product, some are specific for L-lactate as substrate, others for its enantiomer D-lactate; for example, plants, insects and mammals have L-specific enzymes, whereas some other arthropods, such as spiders, have D-specific enzymes. The enzymes themselves belong to different families and show no detectable homology. It would be absurd to suggest on this basis that insects are more closely related to plants than they are to spiders, and, in fact, no such exotic hypothesis turns out to be necessary, as some organisms, including *Homo sapiens*, have both types of enzyme.

Suppose that at some future date all the Eubacteria become extinct. That is (fortunately) highly unlikely, but it is not impossible, and if it happened it would almost

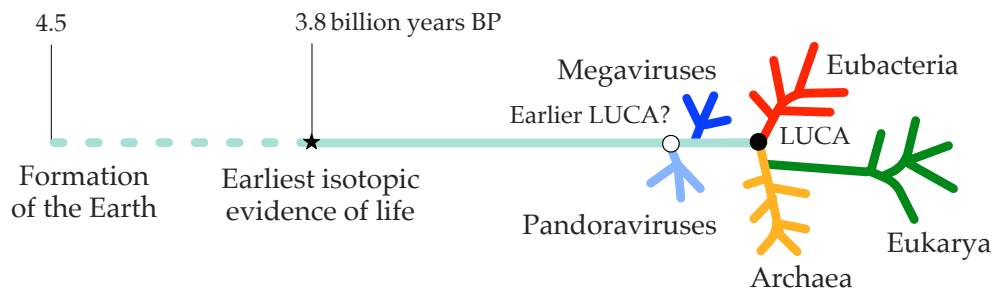


Figure 3: LUCA and the origin of life, which can be estimated from isotopic and fossil data (Joyce, 1991; Arndt and Nisbet, 2012) to date from 3.8 billion years ago. We regard LUCA as much more recent, situated at the bifurcation of the Eubacteria and the Archaea, as suggested by Tuller et al. (2010). This bifurcation is marked with a filled circle. However, if the giant viruses are accepted as living organisms or as descendants of living organisms, or if free-living relatives of these viruses are discovered, LUCA will need to be redefined as an earlier entity (open circle). Regardless of any such redefinition, any LUCA, with hundreds of genes coding for hundreds of proteins must be vastly more complicated than any self-sustaining system that existed at the origin of life. According to Glansdorff et al. (2008), “the first diagnostically identifiable Cyanobacteria are approximately 2.1 billion years old,” and we suggest that LUCA existed earlier, but in the absence of definite information we do not show a specific date in the figure.

certainly eliminate all the Eukarya at the same time, leaving only the Archaea as the masters of the world. LUCA would then be at a more recent point in evolution, corresponding to the time when the Archaea diverged from one another. Perhaps less unlikely, suppose that the search for a “shadow biosphere” (Cleland and Copley, 2005) reveals a few surviving members of a domain that diverged from the known microorganisms before the divergence of the Archaea and Eubacteria: LUCA will then need to refer to this earlier time. In some respects the giant viruses or their free-living ancestors already fulfil this role.

Consideration of giant viruses is important, and deserves attention, because viruses in general have not in the past usually been considered to be alive. The discovery of *Acanthamoeba polyphaga mimivirus*, or “Mimivirus” (Raoult et al., 2004) has reopened the argument, with, for example, Moreira and López-García (2009) saying that viruses are not alive, and Forterre (2010) saying that they are alive. The review of Moreira and López-García (2009) attracted seven comments in the same journal from authors

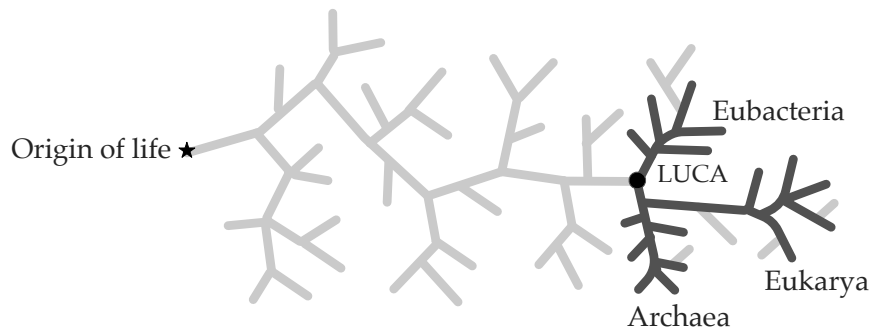


Figure 4: Life before LUCA. We make no suggestion that there were no living organisms before LUCA, the *last* common ancestor. On the contrary, there were probably many, as indicated in the greyed out part of the figure, but all of these are now extinct, as well as others that have become extinct more recently.

who disagreed with them, followed by a further response from López-García and Moreira (2009). Discovery of the other giant viruses (Arslan et al., 2011; Philippe et al., 2013; Legendre et al., 2014) has caused this argument to continue, but we agree with López-García and Moreira (2009) that viruses are not alive. However, the giant viruses have genomes that appear to predate LUCA: for example, two-thirds of the genome of *Pithovirus sibericum*, one of the giant viruses, appear to be completely different from those of the pandoraviruses or other known genomes (Legendre et al., 2014). If this is accepted as a living organism, therefore, LUCA will have to be defined as the point of bifurcation of the giant viruses from the line to the Eubacteria and Archaea. Even if it is not accepted there is still a problem, because *Pithovirus sibericum* has a predicted capacity to code for about 500 proteins, and the most reasonable explanation of why it needs so many is that it is descended from free-living organisms that lost their independence when they became parasites. Some of their free-living relatives may still be found as the “fourth domain of life” (Raoult et al., 2004), and, if they are, then LUCA will need to be moved back in time, regardless of whether viruses are accepted as organisms. However, recent work (Schulz et al., 2017) has cast doubt on the existence of a fourth domain of life, suggesting that the giant viruses originate from much smaller viruses by “piecemeal capture of eukaryotic translation machinery components”.

Figure 4 illustrates our view of the place of LUCA in the kingdom of life. Many living organisms, some of them ancestors of LUCA, existed before LUCA, but have become extinct, together with many that have become extinct more recently.

### 3. The origin of life

#### 3.1. What is life?

Discussion of the origin of life needs a definition of what life is, and we start by offering an answer to the question asked by Erwin Schrödinger (1944). We regard a living system as a network of processes that can maintain itself, with, therefore, a capacity to stay alive in spite of changing conditions. Some modern theories of how that is possible are briefly described in Section 4.

Properties such as reproduction and natural selection are important in biology as we know it today, but neither of these could occur with a system that was incapable of remaining alive for a significant period of time. The organisms that exist today contain numerous proteins that are clearly homologous, most notably ATP synthase, mentioned in the next section. The organisms themselves must therefore be homologous, and descend from common ancestors, the most recent of which is what we understand as LUCA.

The capacity to stay alive could in principle be satisfied without proteins, nucleic acids or other complicated molecules that appear essential to life today. Arriving at these molecules from much simpler ones at the origin of life surely required a very long period of natural selection.

#### 3.2. Catalytic capacity of LUCA

It hardly matters whether the giant viruses are regarded as alive or not, because it is impossible to believe that life started with a self-organizing system with many proteins.<sup>1</sup> Ouzounis et al. (2006) estimated that about 1000 protein families existed in LUCA, and in this special issue Harish and Kurland (2017) suggest that there were about 1300. We have no reason to doubt the validity of these numbers, but we note that they are far too large for an organism close to the origin of life. Even a system with just one gene and one encoded protein is far too complicated to have existed at the

---

<sup>1</sup>Giovannoni et al. (2005) show a range from 500 (*Mycoplasma genitalium*) to about 8000 (*Streptomyces coelicolor*) for the number of predicted proteins in 244 bacterial and archaeal species. However, the smaller genomes refer to parasitic organisms, and the smallest value for a free-living organism is about 1500 for *Thermoplasma acidophilum*. More recently, some insect symbionts have been found to have much smaller genomes than *Mycoplasma genitalium* (Moya et al., 2008), but these are not free-living organisms, and could not be.



origin of life, and in any case such a system could not be an organism, as it would have had no capacity of self-organization. As an example, all modern organisms depend on ATP synthase to use ion gradients across membranes to synthesize ATP, and all known ATP synthases are clearly homologous (Sousa et al., 2016). We can suppose, therefore, that LUCA used an ATP synthase for energy management. What we cannot suppose, however, is that the first organisms did the same, because ATP synthase is a large complicated enzyme that follows an elaborate mechanism (Doering et al., 1995). As it is a membrane-bound enzyme there is a suggestion that LUCA had a biological membrane, and not simply a mineral partition.

Regardless of how LUCA is defined, any self-organizing system that relied on coding of catalysts would fall foul of the paradox that there can be no error-correcting machinery without proteins, and no proteins without error correction (Eigen, 1971), a problem that Maynard Smith and Szathmary (1995) called *Eigen's paradox*.

The earliest self-sustaining organisms, which existed before LUCA, must have been vastly simpler than anything we can see today, and the first catalysts cannot have been as specific as protein enzymes. Instead, metals such as iron, zinc and molybdenum, which still act today as cofactors of protein enzymes, must have been the principal catalysts (Williams and Frausto da Silva, 2006). Even simpler catalysts, such as  $H^+$  and  $OH^-$  ions, which still play important catalytic roles in enzymes today (see Jencks, 1987), have always been available. The early atmosphere is believed to have lacked molecular oxygen,  $O_2$ , and in these conditions the abundant iron present on earth existed as  $Fe^{2+}$  ions, which were readily soluble in water. However, the increase in the  $O_2$  concentration in the atmosphere made the iron less and less available, as  $Fe^{3+}$  ions are much less soluble. The decrease in availability of iron was, however, accompanied by an increase in the availability of copper (Anbar, 2008), as seen in Figure 2, the oxidized  $Cu^{2+}$  ion being much more soluble than the reduced  $Cu^+$  ion. Tungsten can substitute for molybdenum in present-day enzymes without loss of activity (Schoepp-Cothenet et al., 2012), in keeping with the great chemical similarity between the two elements. However, they differ in solubility, and molybdenum has become more available as tungsten has become less. Although modern organisms use molybdenum more than they use tungsten, it is possible that the opposite was true at the beginning.

In general, we can suppose that the metals found as enzyme cofactors today also acted at the origin of life, but as far less specific and less active catalysts than they are when they are bound to proteins as cofactors.

### 3.3. Metabolism first, or replication first?

There has been considerable argument over whether metabolism arose first, or (RNA) replication (see, for example, Gabora, 2006). In general we find the metabolism-first scenario more plausible, but this cannot mean metabolism as we see it today, with reactions catalysed by protein enzymes. Metabolism must have started without protein enzymes, with chemical reactions either catalysed by metal ions or other simple catalysts, or not catalysed at all.<sup>2</sup> More specific catalysts, whether involving polynucleotides or polypeptides, could then favour the systems that contained them, both because they would allow reactions to occur faster, and, probably more important, because the greater specificity would eliminate some of the parasitic side reactions. We need to look, therefore, for chemical reactions that could plausibly occur in early life, such as the formose reaction (Boutlerow, 1861), in which formaldehyde reacts spontaneously to a mixture of sugars, most notably glucose. This approach has been applied in particular by Meléndez-Hevia et al. (1996, 2008), who explain, for example, how the tri-carboxylate cycle could have arisen without regarding it as “irreducibly complex” in the creationist sense. We suggest that evolution of metabolism and evolution of RNA replication must have occurred in parallel, because for RNA to exist there must be reactions that produce its components, and at the same time there could have been reactions that produce amino acids, which could polymerize to form polypeptides, and both polynucleotides and polypeptides are known to be capable of catalytic activity. We thus agree with the view indicated in Figure 2 of Glansdorff et al. (2008) that networks with *catalytic closure* existed before LUCA. This means that they can only be self-sustaining if they themselves produce the catalysts (other than simple ions) that they need, and agrees well with the theories of life that we discuss in the next section.

## 4. Theories of life

The principal modern theories of life —  $(M, R)$  systems (Rosen, 1991), autopoiesis (Maturana and Varela, 1980) and the chemoton (Gánti, 2003)<sup>3</sup> — are theories of *life*

---

<sup>2</sup>In the highly competitive world that exists today an organism without catalysts for essential processes would be outgrown by competitors that had them, but that would have been a far less important consideration in a world in which competitors did not exist.

<sup>3</sup>Gánti’s study of life started long before 2003, but his early publications, from Gánti (1971) onwards, are in Hungarian, and thus not easily accessible to most readers. Gánti (2003) is the first full account of his

rather than theories for the *origin of life*, but they are nonetheless relevant to the origin of life as they offer explanations of how simple systems can be self-sustaining. Two other theories, autocatalytic sets (Kauffman, 1986) and the hypercycle (Eigen and Schuster, 1977) are more specifically concerned with the origin of life. We do not share the opinion of Szostak (2012) that “attempts to define life do not help to understand the origin of life”, because we do not agree that one can study the origin of an entity without any definition of what the entity is.<sup>4</sup> All of these theories, which all imply catalytic closure (or *metabolic closure*), as defined earlier, are illustrated in outline in Figure 5.

We have discussed the differences and similarities between them elsewhere (Letelier et al., 2011; Cornish-Bowden, 2015). A recent newcomer to the field is Friston (2013), who used a highly mathematical argument to conclude that an ergodic system with a Markov blanket will inevitably result in life.<sup>5</sup> Although his arguments are different, his conclusion is similar to that of Kauffman (1986).

#### 4.1. Rosen’s (*M, R*) systems

In Rosen’s view a living system is a catalytically closed network of processes (Rosen, 1991). In his description, there is no particular implication of how large a system needs to be to be self-sustaining, but in our efforts to give concrete expression to his rather abstract presentation we have suggested that as few as two catalysts and three catalytic cycles can provide for the properties of *metabolism*, *replacement* and *organizational invariance* (Piedrafita et al., 2010) needed for metabolic closure, to persist in time and maintain identity

#### 4.2. Maturana and Varela’s autopoiesis and Gánti’s chemoton

Models of autopoiesis also include a rather small number of processes. Chemoton models tend to be larger, as Gánti included a specific (but rather rudimentary) *informa-*

---

ideas in English.

<sup>4</sup>A decade earlier Szostak seemed to be less hostile to the usefulness of a definition: “How simple can a cell be and still be considered as living? The answer depends on what we consider to be the essential properties of life. Defining life is notoriously difficult; its very diversity resists the confines of any compact definition” (Szostak et al., 2001).

<sup>5</sup>An *ergodic system* is a dynamic system in which the proportion of time that it spends in a particular state is the same as the probability that it will be found in that state at a random moment. A *Markov blanket* is the condition that all information about a random variable in a Bayesian network is contained within the set of nodes composed of its parents, children, and other parents of its children.

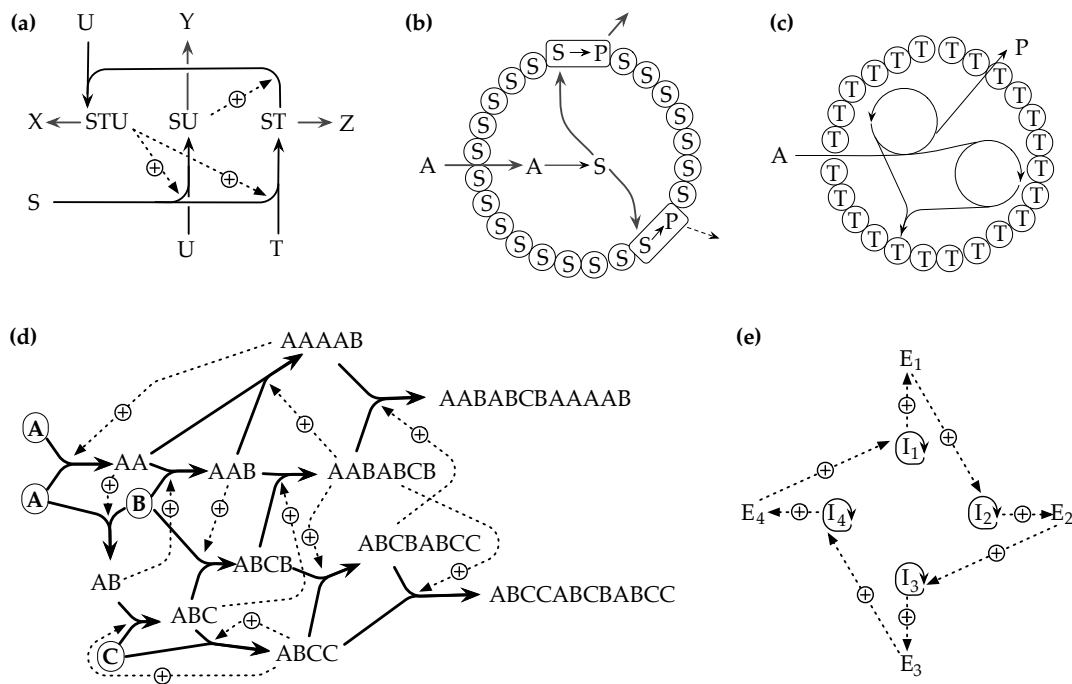


Figure 5: Modern theories of the essence of life. (a) An  $(M, R)$  system (Rosen, 1991), as interpreted by us (Piedrafita et al., 2010); (b) autopoiesis (Maturana and Varela, 1980); (c) a chemoton (Gánti, 2003); (d) an autocatalytic set (Kauffman, 1986); (e) a hypercycle (Eigen and Schuster, 1977). These are discussed elsewhere (Letelier et al., 2011), and a full discussion would be inappropriate here. The essential point is that they are all simple compared with a modern organism that has the capacity to code for hundreds of proteins. The autocatalytic set is conceptually very simple, despite its more complicated appearance.

tion cycle, but they are still very small compared with even the simplest known living organisms. Both of these, especially autopoiesis, emphasize the formation of a membrane, a feature missing from Rosen’s theory.

### 4.3. Kauffman’s autocatalytic sets

Kauffman’s original description of his autocatalytic sets implied that they are very much larger than the others we have mentioned, in terms of the numbers of distinct kinds of molecules that they contain (typically of the order of  $10^9$ ),<sup>6</sup> but they are con-

<sup>6</sup>The value of  $10^9$  is based on arguable assumptions, but it is interesting to compare it with experimental estimates (Ellington and Szostak, 1990; Sassanfar and Szostak, 1993) that roughly one in  $10^{10}$ – $10^{11}$  random-

ceptually much simpler, as he considered how a self-sustaining system with catalytic closure could arise from pure chance properties of its component molecules. However, more recent work (Hordijk and Steel, 2004) has shown that autocatalytic sets do not need to be as large as Kauffman's original analysis suggested.

## 5. Interlude: Lynn Margulis and controversy

Throughout her career, Lynn Margulis never tried to avoid controversy, and one cannot give a true picture of her character by censoring her most unpopular opinions: as an early example, her classic paper on the origin of the eukaryotes (Sagan, 1967) was against the conventional wisdom of the time, and generated much argument. Elsewhere in this issue, Doolittle (2017) discusses her collaboration with James Lovelock on the Gaia hypothesis (Margulis and Lovelock, 1974): this continues to be dismissed out of hand as nonsense by many scientists, though the fact that Margulis took it seriously is surely a reason not to do that.

Towards the end of her life Margulis defended Peter Duesberg and others who oppose the view that AIDS is a disease caused by HIV (Margulis, 2007; Margulis et al., 2009). In the first sentences of an unpublished document (Margulis, 2007) she made her position very clear, and not long before she died she reiterated her view in a Spanish magazine (Margulis, 2009):

Since Robert Gallo made his results public I have not been able to cite a single publication that satisfactorily proves to us microbiologists that there exists a complete correlation showing that HIV is responsible for the disease.<sup>7</sup>

Many scientists consider the opinions she expresses here to be at best mistaken and at worst positively dangerous, but the way to answer them is with detailed analysis in the scientific literature, not by suggestions that she was becoming senile in the last years of her life (Prothero, 2011).<sup>8</sup>

---

sequence RNA molecules folds in such a way as to create a specific binding site for small ligands such as organic dyes and ATP.

<sup>7</sup>Desde que Robert Gallo hizo públicos sus resultados, no he sido capaz de dar con una sola publicación que nos pruebe de una manera satisfactoria a nosotros, microbiólogos, que exista una correlación completa de que HIV sea el responsable de la enfermedad.

<sup>8</sup>"A highly respected and honored senior scientist, largely out of the mainstream and not up to date

## 6. Discussion

Evolving from a simple self-sustaining system to an organism with a coding capacity of hundreds of proteins is a huge step: even *Thermoplasma acidophilum*, with the smallest genome for a free-living organism, has about 1500 protein-coding genes (Giovannoni et al., 2005). We cannot at present know how much this number could be decreased while still having a viable free-living organism. Understanding how the transition to an organism with a large coding capacity can have happened is a more challenging problem than understanding how LUCA could have evolved to *Homo sapiens*. For the post-LUCA evolution we have at least a rough idea of how it happened and what mechanisms were involved, but pre-LUCA evolution is a black box, probably more difficult to understand than to understand how self-sustaining systems came to exist in the first place. If the theoretical ideas of Kauffman (1986) or Friston (2013) are valid then self-sustaining systems may be inevitable, but getting from these to LUCA still presents enormous difficulties, and some crucial points, such as the origin of the genetic code, remain speculative. Our view is that studying the pre-LUCA period and finding a plausible series of steps to LUCA is an urgent task, but it will not be solved as long as LUCA continues to be discussed as if it were an entity close to the origin of life.

### Acknowledgements

This work was supported by the CNRS. We thank Dr Wolfgang Nitschke of this Institute for helpful discussions.

### References

- Acevedo-Rocha, C. G., Fang, G., Schmidt, M., Ussery, D. W., Danchin, A., 2013. From essential to persistent genes: a functional approach to constructing synthetic life. *Trends Genet.*, 29, 273–279. doi:10.1016/j.tig.2012.11.001.
- Anbar, A. D., 2008. Oceans, elements and evolution. *Science*, 322, 1481–1483. doi:10.1126/science.1163100.
- Arndt, N. T., Nisbet, E. G., 2012. Processes on the young earth and the habitats of early life. *Ann. Rev. Earth Planet. Sci.*, 40, 521–549. doi:10.1146/annurev-earth-042711-105316.

---

with the recent developments (and perhaps a bit senile), makes weird pronouncements about their pet ideas”.

- Arslan, D., Legendre, M., Seltzer, V., Abergel, C., Claverie, J.-M., 2011. Distant Mimivirus relative with a larger genome highlights the fundamental features of Megaviridae. *Proc. Natl. Acad. Sci. USA*, 108, 17486–17491. doi:10.1073/pnas.1110889108.
- Banerjee, S., Margulis, L., 1973. Mitotic arrest by melatonin. *Exp. Cell Res.*, 78, 314–318. doi:10.1016/0014-4827(73)90074-8.
- Becerra, A., Delaye, L., Islas, S., Lazcano, A., 2007. The very early stages of biological evolution and the nature of the last common ancestor of the three major cell domains. *Ann. Rev. Ecol. Evol. Sys.*, 38, 361–379. doi:10.1146/annurev.ecolsys.38.091206.095825.
- Boutlerow, A., 1861. Formation synthétique d’une substance sucrée. *C.R. Acad. Sci.*, 53, 145–147. Modern English transliteration: Aleksandr Mikhailovich Butlerov.
- Chapman, M. J., Dolan, M. F., Margulis, L., 2000. Centrioles and kinetosomes: form, function, and evolution. *Q. Rev. Biol.*, 75, 409–429. doi:10.1086/393621.
- Chapman, M. J., Margulis, L., 1998. Morphogenesis by symbiogenesis. *Int. Microbiol.*, 1, 319–326.
- Cleland, C. E., Copley, S. D., 2005. The possibility of alternative microbial life on Earth. *Int. J. Astrobiol.*, 4, 165–173. doi:10.1017/S147355040500279X.
- Cornish-Bowden, A., 2015. Tibor Gánti and Robert Rosen: contrasting approaches to the same problem. *J. Theor. Biol.*, 381, 6–10. doi:10.1016/j.jtbi.2015.05.015.
- Cristescu, M. E., Innes, D. J., Stillman, J. H., Crease, T. J., 2008. D- and L-lactate dehydrogenases during invertebrate evolution. *BMC Evol. Biol.*, 8, 268. doi:10.1186/1471-2148-8-268.
- Di Giulio, M., 2011. The last universal common ancestor (LUCA) and the ancestors of Archaea and Bacteria were progenotes. *J. Mol. Evol.*, 72, 119–126. doi:10.1007/s00239-010-9407-2.
- Doering, C., Ermentrout, B., Oster, G., 1995. Rotary DNA motors. *Biophys. J.*, 69, 2256–2267. doi:10.1016/S0006-3495(95)80096-2.
- Doolittle, W. F., 2017. Darwinizing Gaia. *J. Theor. Biol.*, 000, 000–000. doi:10.1016/j.jtbi.2017.02.015.
- Eigen, M., 1971. Selforganization of matter and evolution of biological macromolecules. *Naturwissenschaften*, 58, 465–523. doi:10.1007/BF00623322.
- Eigen, M., Schuster, P., 1977. The hypercycle: a principle of natural self-organization. Part A: emergence of the hypercycle. *Naturwissenschaften*, 64, 541–565. doi:10.1007/BF0450633.
- Ellington, A. D., Szostak, J. W., 1990. *In vitro* selection of RNA molecules that bind specific ligands. *Nature*, 346, 818–822. doi:10.1038/346818a0.
- Forterre, P., 2010. Defining life: the virus viewpoint. *Orig. Life Evol. Biosph.*, 40, 151–160. doi:10.1007/s11084-010-9194-1.

- Friston, K., 2013. Life as we know it. *J. Roy. Soc. Interface*, 10, 20130475. doi:10.1098/rsif.2013.0475.
- Gabora, L., 2006. Self-other organization: why early life did not evolve through natural selection. *J. Theor. Biol.*, 241, 443–450. doi:10.1016/j.jtbi.2005.12.007.
- Gánti, T., 1971. *Az élet principiuma*. Gondolat, Budapest.
- Gánti, T., 2003. *The principles of life*. Oxford University Press, Oxford.
- Gill, S., Forterre, P., 2015. Origin of life: LUCA and extracellular membrane vesicles (EMVs). *Int. J. Astrobiol.*, 15, 7–15. doi:10.1017/S1473550415000282.
- Giovannoni, S. J., Tripp, H. J., Givan, S., Podar, M., Vergin, K. L., Baptista, D., Bibbs, L., Eads, J., Richardson, T. H., Noordewier, M., Rappe, M. S., Short, J. M., Carrington, J. C., Mathur, E. J., 2005. Genome streamlining in a cosmopolitan oceanic bacterium. *Science*, 309, 1242–1245. doi:10.1126/science.1114057.
- Glansdorff, N., Xu, Y., Labedan, B., 2008. The Last Universal Common Ancestor: emergence, constitution and genetic legacy of an elusive forerunner. *Biol. Direct*, 3, 29. doi:10.1186/1745-6150-3-29.
- Guldan, H., Sterner, R., Babinger, P., 2008. Identification and characterization of a bacterial glycerol-1-phosphate dehydrogenase: Ni<sup>2+</sup>-dependent AraM from *Bacillus subtilis*. *Biochemistry*, 47, 7376–7384. doi:10.1021/bi8005779.
- Harish, A., Kurland, C. G., 2017. Mitochondria are not captive bacteria. *J. Theor. Biol.*, 000, 000–000.
- Hordijk, W., Steel, M., 2004. Detecting autocatalytic, self-sustaining sets in chemical reaction systems. *J. Theor. Biol.*, 227, 451–461. doi:10.1016/j.jtbi.2003.11.020.
- Jencks, W. P., 1987. *Catalysis in Chemistry and Enzymology*. Dover Publications, New York. Original edition (1969) published by McGraw-Hill, New York.
- Jheeta, S., 2015. The routes of emergence of life from LUCA during the RNA and viral world: a conspectus. *Life*, 5, 1445–1453. doi:10.3390/life5021445.
- Joyce, G. F., 1991. The rise and fall of the RNA world. *New Biologist*, 3, 399–407.
- Kauffman, S. A., 1986. Autocatalytic sets of proteins. *J. Theor. Biol.*, 119, 1–24. doi:10.1016/S0022-5193(86)80047-9.
- Lane, N., 2017. Serial symbiosis or singular event at the origin of eukaryotes. *J. Theor. Biol.*, in press. doi:10.1016/j.jtbi.2017.04.031.
- Lane, N., Allen, J. F., Martin, W., 2010. How did LUCA make a living? Chemiosmosis in the origin of life. *BioEssays*, 32, 271–280. doi:10.1002/bies.200900131.
- Legendre, M., Bartoli, J., Shmakova, L., Jeudy, S., Labadie, K., Adrait, A., Lescot, M., Poirot, O., Bertaux, L., Bruley, C., Couté, Y., Rivkina, E., Abergel, C., Claverie, J.-M., 2014. Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. *Proc. Natl. Acad. Sci. USA*, 111, 4274–4279. doi:10.1073/pnas.1320670111.



- Letelier, J.-C., Cárdenas, M. L., Cornish-Bowden, A., 2011. From *L'Homme Machine* to metabolic closure: steps towards understanding life. *J. Theor. Biol.*, 286, 100–113. doi:10.1016/j.jtbi.2011.06.033.
- López-García, P., Moreira, D., 2009. Yet viruses cannot be included in the tree of life. *Nat. Rev. Microbiol.*, 7, 615–617. doi:10.1038/nrmicro2108-c7.
- Margulis, L., 2001. The conscious cell. *Ann. New York Acad. Sci.*, 929, 55–70. doi:10.1111/j.1749-6632.tb05707.x.
- Margulis, L., 2007. What is an HIV/AIDS denier? Or HIV/AIDS denialist? URL [http://www.robertogiraldo.com/reference/lynn\\_margulis\\_march\\_2007.html](http://www.robertogiraldo.com/reference/lynn_margulis_march_2007.html). (Accessed 22 June 2016, not formally published).
- Margulis, L., 2009. La simbiogénesis es la fuente de innovación en la evolución. *SEBBM*, 160, 26–29. URL <http://www.sebbm.com/pdf/160/e160.pdf>.
- Margulis, L., Chapman, M., Guerrero, R., Hall, J., 2006. The last eukaryotic common ancestor (LECA): acquisition of cytoskeletal motility from aerotolerant spirochetes in the Proterozoic eon. *Proc. Natl. Acad. Sci. USA*, 103, 13080–13085. doi:10.1073/pnas.0604985103.
- Margulis, L., Dolan, M., Guerrero, R., 2000. The chimeric eukaryote: origin of the nucleus from the karyomastigont in amitochondriate protists. *Proc. Natl. Acad. Sci. USA*, 97, 6954–6959. doi:10.1073/pnas.97.13.6954.
- Margulis, L., Lovelock, J., 1974. Biological modulation of earth's atmosphere. *Icarus*, 21, 471–489. doi:10.1016/0019-1035(74)90150-X.
- Margulis, L., Maniatis, A., MacAllister, J., Scythes, J., Brorson, O., Hall, J., Krumbein, W. E., Chapman, M. J., 2009. Spirochete round bodies Syphilis, Lyme disease & AIDS: resurgence of “the great imitator”? *Symbiosis*, 47, 51–58. doi:10.1007/BF3179970.
- Mariscal, C., Doolittle, W. F., 2015. Eukaryotes first: how could that be? *Phil. Trans. R. Soc. B*, 370, 20140322. doi:10.1098/rstb.2014.0322.
- Martin, W., Russell, M. J., 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. London Ser. B*, 358, 59–83.
- Martin, W. F., Weiss, M. C., Neukirchen, S., Nelson-Sathi, S., Sousa, F. L., 2016. Physiology, phylogeny, and LUCA. *Microb. Cell*, 3, 582–587. doi:10.15698/mic2016.12.545.
- Maturana, H., Varela, F., 1980. *Autopoiesis and cognition: the realisation of the living*. D. Reidel Publishing Company, Dordrecht.
- Maynard Smith, J., Szathmáry, E., 1995. *The Major Transitions in Evolution*. Freeman, Oxford.
- Meléndez-Hevia, E., Montero-Gómez, N., Montero, F., 2008. From prebiotic chemistry to cellular metabolism — the chemical evolution of metabolism before Darwinian natural selection. *J. Theor. Biol.*, 252, 505–519. doi:10.1016/j.jtbi.2007.11.012.

- Meléndez-Hevia, E., Waddell, T., Cascante, M., 1996. The puzzle of the Krebs citric acid cycle: assembling the pieces of chemically feasible reactions, and opportunism in the design of metabolic pathways during evolution. *J. Mol. Evol.*, 43, 293–303. doi:10.1007/BF02338838.
- Moreira, D., López-García, P., 2009. Ten reasons to exclude viruses from the tree of life. *Nat. Rev. Microbiol.*, 7, 306–311. doi:10.1038/nrmicro2108.
- Moya, A., Peretó, J., Gil, R., Latorre, A., 2008. Learning how to live together: genomic insights into prokaryote-animal symbioses. *Nat Rev Genet.*, 9, 218–229. doi:10.1038/nrg2319.
- 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. *Res. Microb.*, 157, 57–68. doi:10.1016/j.resmic.2005.06.015.
- Peretó, J., López-García, P., Moreira, D., 2004. Ancestral lipid biosynthesis and early membrane evolution. *Trends Biochem. Sci.*, 29, 469–477. doi:10.1016/j.tibs.2004.07.002.
- Peterhoff, D., Zellner, H., Guldan, R., H Merkl, Sterner, R., Babinger, P., 2012. Dimerization determines substrate specificity of a bacterial prenyltransferase. *Chembiochem*, 13, 1297–303. doi:10.1002/cbic.201200127.
- Philippe, N., Legendre, M., Doutre, G., Couté, Y., Poirot, O., Lescot, M., Arslan, D., Seltzer, V., Bertaux, L., Bruley, C., Garin, J., Claverie, J.-M., Abergel, C., 2013. Pandoraviruses: amoeba virus with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. *Science*, 341, 281–286. doi:10.1126/science.1239181.
- Piedrafita, G., Montero, F., Morán, F., Cárdenas, M. L., Cornish-Bowden, A., 2010. A simple self-maintaining metabolic system: robustness, autocatalysis, bistability. *PLoS Comput. Biol.*, 6, e1000872. doi:10.1371/journal.pcbi.1000872.
- Prothero, D., 2011. What is an HIV/AIDS denier? Or HIV/AIDS denialist? URL <http://www.skepticblog.org/2011/04/13/the-linus-pauling-effect/>. (Last accessed 28 April 2017, not formally published).
- Pushie, M. J., Cotelesage, J. J., George, G. N., 2014. Molybdenum and tungsten oxygen transferases — structural and functional diversity within a common active site motif. *Metallomics*, 6, 15–24. doi:10.1039/c3mt00177f.
- Raoult, D., Audic, S., Robert, C., Abergel, C., Renesto, P., Ogata, H., La Scola, B., Suzan, M., Claverie, J. M., 2004. The 1.2-megabase genome sequence of mimivirus. *Science*, 306, 1344–1350. doi:10.1126/science.1101485.
- Rosen, R., 1991. *Life itself*. Columbia University Press, New York.
- Sagan, L., 1967. On the origin of mitosing cells. *J. Theor. Biol.*, 14, 255–274. doi:10.1016/0022-5193(67)90079-3.
- Sassanfar, M., Szostak, J. W., 1993. An RNA motif that binds ATP. *Nature*, 364, 550–553. doi:10.1038/364550a0.

- Schoepp-Cothenet, B., van Lis, R., Philippot, P., Magalon, A., Russell, M. J., Nitschke, W., 2012. The ineluctable requirement for the trans-iron elements molybdenum and/or tungsten in the origin of life. *Sci. Rep.*, 2. doi:10.1038/srep00263.
- Schrödinger, E., 1944. *What is Life?* Cambridge University Press, Cambridge.
- Schulz, F., Yutin, N., Ivanova, N. N., Ortega, D. R., Lee, T. K., Vierheilig, J., Daims, H., Horn, M., Wagner, M., Jensen, G. J., Kyrpides, N. C., Koonin, E. V., Woyke, T., 2017. Giant viruses with an expanded complement of translation system components. *Science*, 356, 82–85. doi:10.1126/science.aal4657.
- Sousa, F. L., Nelson-Sathi, S., Martin, W. F., 2016. One step beyond a ribosome: the ancient anaerobic core. *Biochim. Biophys. Acta*, 1857, 1027–1038. doi:10.1016/j.bbabi.2016.04.284.
- Szostak, J., 2012. Attempts to define life do not help to understand the origin of life. *J. Biomol. Struct. Dyn.*, 29, 599–600.
- Szostak, J. W., Bartel, D. P., Luisi, P. L., 2001. Synthesizing life. *Nature*, 409, 387–390. doi:10.1038/35053176.
- Tuller, T., Birin, H., Gophna, U., Kupiec, M., Ruppin, E., 2010. Reconstructing ancestral gene content by coevolution. *Genome Res.*, 20, 122–132. doi:10.1101/gr.096115.109.
- Weiss, M. C., Sousa, F. L., Mrnjavac, N., Roettger, M., Nelson-Sathi, S., Martin, W. F., 2016. The physiology and habitat of the last universal common ancestor. *Nature Microbiol.*, 1, 16116. doi:10.1038/nmicrobiol.2016.116.
- Wieczorek, R., 2012. On prebiotic ecology, supramolecular selection and autopoiesis. *Orig. Life Evol. Biosph.*, 42, 445–450. doi:10.1007/s11084-012-9306-1.
- Williams, R. J. P., Fraústo da Silva, J. R. R., 2006. *The Chemistry of Evolution: the Development of our Ecosystem.* Elsevier, Amsterdam.