Symbiosis in eukaryotic evolution

Purificación López-García¹, Laura Emé, David Moreira¹

¹ Ecologie Systématique Evolution, CNRS, Université Paris-Sud, Université Paris-Saclay, AgroParisTech, 91400 Orsay, France

² Centre for Comparative Genomics and Evolutionary Bioinformatics, Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, Canada NS B3H 4R2

ARTICLE INFO

Keywords:
- Eukaryogenesis
- Symbiosis
- Syntrophy
- Archaea
- Mitochondria
- Eukaryotic origins

ABSTRACT

Fifty years ago, Lynn Margulis, inspired by early twentieth-century ideas that postulated a symbiotic origin for some eukaryotic organelles, proposed a unified theory for the origin of the eukaryotic cell based on symbiosis as evolutionary mechanism. Margulis was profoundly aware of the importance of symbiosis in the natural microbial world and anticipated the evolutionary significance that integrated cooperative interactions might have as a mechanism to increase cellular complexity. Today, we have started fully appreciating the vast extent of microbial diversity and the importance of syntrophic metabolic cooperation in natural ecosystems, especially in sediments and microbial mats. Also, not only the symbiogenetic origin of mitochondria and chloroplasts has been clearly demonstrated, but improvement in phylogenomic methods combined with recent discoveries of archaebial lineages more closely related to eukaryotes further support the symbiogenetic origin of the eukaryotic cell. Margulis left us in legacy the idea of ‘eukaryogenesis by symbiogenesis’. Although this has been largely verified, when, where, and specifically how eukaryotic cells evolved are yet unclear. Here, we shortly review current knowledge about symbiotic interactions in the microbial world and their evolutionary impact, the status of eukaryogenetic models and the current challenges and perspectives ahead to reconstruct the evolutionary path to eukaryotes.

1. Introduction

In 1967, Lynn Margulis (Sagan) published her famous On the origin of mitosing cells (Sagan, 1967), where she set up the basis of her Serial Endosymbiotic Theory for the origin of the eukaryotic cell (Margulis, 1981, 1996). In her manuscript, Margulis revived previous ideas proposing the endosymbiotic origin of chloroplasts from ‘blue-green algae’ (cyanobacteria) (Mereschkowsky, 1905, 1910) and mitochondria from purple bacteria (alphaproteobacteria) (Wallin, 1927); and further put forward the idea that the eukaryotic flagellum derived from symbiotic bacteria (spirochete-like). Although Konstantin Mereschkowsky had also advocated for the endosymbiotic origin of the eukaryotic nucleus (Mereschkowsky, 1910), Margulis discarded this possibility without nonetheless proposing a clear mechanism for the origin of the nucleus in an original heterotrophic amoeboid host (Sagan, 1967). Later, after the recognition of archaea as a third phylogenetic domain of life (Woese and Fox, 1977; Woese et al., 1990) showing significant similarities in terms of informational processes (replication, transcription, translation) with eukaryotes (Rivera et al., 1998; Thiergart et al., 2012), she postulated that the host of the future mitochondrion derived from a thermoacidophilic wall-less archaeon similar to contemporary Thermoplasma spp. (Euryarchaeota) (Margulis, 1981, 1996), an idea that she adopted from Dennis Searcy (Searcy, 1992). The nucleus would have evolved autogenously (not by symbiosis) in a chimeric archaeal-bacterial cell (Margulis et al., 2000).

The strength and innovative character of Margulis’ ideas lay in the proposal of a unified theory based on symbiosis as evolutionary mechanism for the origin of the eukaryotic cell: eukaryogenesis by symbiogenesis. Strongly criticized at the time, at least part of these ideas became credible with the advent of molecular phylogeny that, among its earliest achievements, allowed demonstrating that conserved genes in chloroplast and mitochondrial genomes clustered with, respectively, cyanobacterial and alphaproteobacterial genes in phylogenetic trees (Schwartz and Dayhoff, 1978). The endosymbiotic origin of these membrane-bound organelles became mainstream science and it is now well established how their bacterial ancestors evolved, undergoing a process of genome reduction, impacting their host and leading to a variety of organelle derivatives (Iyall et al., 2004; Gray, 2012; Müller et al., 2012; Schwartz and Dayhoff, 1978). However, the origin of her eukaryotic ‘nucleocytoplasm’ was much controversial; neither the Thermoplasma-like nature of the host nor (far less) the
symbiotic origin of the flagellum were ever considered seriously. In her view, a metabolic symbiosis based on interspecific sulfur transfer was established between a *Thermoplasma*-like archaean, which generated hydrogen sulfide, and a spirochete, which oxidized sulfide to sulfur and, at the same time, provided motility to the symbiotic consortium (Margulis et al., 2000), being at the origin of the eukaryotic cytoskeleton and the mitotic apparatus (Margulis et al., 2006; Sagan, 1967). Margulis argued that, similarly to mitochondria and chloroplasts, which retained reduced genomes, a remnant genome should be found in connection with the (9+2) microtubular basal bodies associated to eukaryotic flagella (Sagan, 1967). But such remnant genomes were never found and phylogenomic analyses in the genomics era did not show any particular similarity between eukaryote and spirochete genomes. Thus, while the endosymbiotic origin of chloroplasts and mitochondria in an eukaryotic host became widely accepted, the symbiotic origin of the eukaryotic cell itself was not and Margulis' ideas stood alone for almost three decades before other symbiogenetic models made it through to the scientific arena.

Lynn Margulis had a classical, exceptional knowledge on the morphology, biology and ecology of microorganisms, which led her to actively participate in attempts to establish global classification systems for microorganisms. She sustained that life should be classified in two superkingdoms or domains (prokaryotes and eukaryotes) and five kingdoms (Margulis, 1992; Margulis and Schwartz, 1998; Whittaker and Margulis, 1978), and co-edited the famous Handbook of Protocista (Margulis et al., 1990) widely used by eukaryotic microbiologists. Paradoxically, although she recognized the importance of DNA comparisons in establishing a natural (evolutionary) classification system for all organisms (Sagan, 1967), she was reluctant to adopt (Margulis and Guerrero, 1991) the Carl R. Woese's three-domain classification system that resulted from molecular phylogenetic analyses of universally conserved genes (Woese and Fox, 1977; Woese et al., 1990). The molecular revolution initiated by Woese using 16S rRNA genes had three major outcomes that have subsequently impacted many areas in biology, from taxonomy and systematics to microbial ecology and comparative genomics. He showed that it was possible to build universal phylogenetic trees for all cellular life and hence establish a natural biological classification system. In doing so, he incidentally discovered that a group of prokaryotic cells, the archaea, defined a distinct phylogenetic group (a third domain of life) (Woese and Fox, 1977; Woese et al., 1990). And finally, as corollary of his work, it was in principle possible to explore microbial diversity in natural environments by amplifying and sequencing conserved marker genes, thus sidestepping the well-known problem of culture bias (only a tiny fraction of microorganisms are amenable to culture in the laboratory). The amplification, cloning and sequencing of 16S and 18S rRNA genes from many different environments led in the 1990s and early 2000s to the realization that the diversity of both prokaryotic and eukaryotic microorganisms was much more important than ever thought (Moreira and López-Garcia, 2002; Pace, 1997). Today, large-scale metabarcoding analyses based on amplicon high-throughput sequencing, together with single cell genomics and metagenome-based genome reconstruction not only reinforce the view of an extraordinarily important microbial diversity but at the same time validate the three domains of life (Archaea, Bacteria, Eucarya) that Woese identified forty years ago (Eme and Doolittle, 2015; Hug et al., 2016; Rinke et al., 2013; Yarza et al., 2014).

This possibility to study novel, uncultured organisms in the wild has revealed crucial to advance in the resolution of the long-term query of eukaryotic origins. Notwithstanding Margulis symbiogenetic ideas, the eukaryogenetic model that prevailed in the last part of the twentieth century and until recently was that of an autogenous origin of all typical eukaryotic features but mitochondria and chloroplasts (of endosymbiotic origin) in a proto-eukaryotic lineage sister to archaea (de Duve, 2007; Jekely, 2003; Poole and Penny, 2007). The discovery that mitochondrial protists lost (or modified) mitochon-
microbial ecology determinants (Moran and Sloan, 2015; Mushegian and Ebert, 2016).

But what is exactly symbiosis? 'Living together' by etymology, symbiosis has two major definitions in biology. One is looser and refers to the co-existence and dependency of at least one organism upon a partner. Symbiosis would then accommodate three types of interactions depending on the effect on the host’s fitness: mutualism (fitness increase), commensalism (no fitness change), and parasitism (fitness decrease) (Mushegian and Ebert, 2016). In reality, the line between mutualism and parasitism is thin and these kinds of interactions are often seen as a continuum, sometimes depending on the environmental conditions. Examples of parasite–mutualist transitions are well-documented in plant mycorrhiza (Paszkowski, 2006), Wolbachia-insect interactions (Hosokawa et al., 2010) and even in intrinsic genetic parasites such as plasmids or viruses (Bao and Roosinck, 2013). The other type of definition equals symbiosis and mutualism, so that symbiosis implies a mutually beneficial interaction. This is the type of definition that Margulis used when she referred to symbiosis. In her view, the holobiont properly meant an integrated symbiotic consortium (obligate mutualism) that behaves as a unit of evolution (Guerrero et al., 2013). Eukaryogenesis could only be understood in this context.

We are progressively discovering that mutualistic symbioses (hereafter, symbioses) are plentiful in nature and, in particular environments such as sediments or microbial mats with steep redox gradients, they may be the rule. Some examples follow.

2.2. Eukaryote – prokaryote symbioses

This type of symbioses is often easy to identify since, in many cases, it implies endosymbiosis within a larger host and, in the case of multicellular organisms, the frequent evolution of dedicated structures, such as bacteriocytes in aphids, trophosomes in deep-sea polychaetes or root nodules in plants. Many of these symbioses are the result of long co-evolutionary processes that have shaped the genotypes and phenotypes of the different partners involved.

Prokaryotic symbioses with animals have been widely documented (Moya et al., 2008). They have been particularly well studied in insects, where they can have astounding impacts not only in nutritional but also in reproductive aspects of their biology. Thus, vertically-transmitted alphaproteobacterial endosymbionts of the genus Wolbachia impact sex determination and can induce feminization, parthenogenesis, male killing and sperm-egg incompatibility (Cordaux et al., 2011; Werren et al., 2008). Comparative genomic analyses are starting to reveal the underlying basis for these effects, which is sometimes linked to horizontal gene transfer from the bacterium to the host genome. For instance, a 3-Mb insert of a feminizing Wolbachia genome was recently transferred into the pillbug nuclear genome and its occurrence correlates with the female sex (Leclercq et al., 2016). Also well-studied symbioses are those involving aphids and other phloem-sap feeding insects with different bacteria that complement their imbalanced, sugar-rich diet by supplying amino acids. A famous symbiotic couple is that of aphids with Buchnera aphidicola, an endosymbiotic gammaproteobacterium living in specialized aphid cells (bacteriocytes). These bacterial endosymbionts, being vertically inherited and having low population sizes, are prone to genetic drift. Their genes are fast-evolving and their genomes, AT-biased and progressively reduced (McCutcheon and Moran, 2012; Moran and Bennett, 2014). Genome reduction can lead to the loss of essential genes for amino-acid biosynthesis. When this happens, a second, less reduced endosymbiont that complements the lost biosynthetic pathways can be acquired and maintained in tripartite symbiosis (Perez-Brocal et al., 2006) or it can be simply lost and replaced by a new symbiont (Koga and Moran, 2014). The replacement of a reduced-genome symbiont is not frequent and can be achieved experimentally (Moran and Yun, 2015).

Less well-known, but equally important are symbioses of bacteria and marine animals. The microbial communities that inhabit sponges, making up to one third of their biomass, are specific to their hosts and many of their bacterial and archaeal members may be true symbionts (Hentschel et al., 2012). Members of the phylum Poribacteria are specifically associated with sponges and seem to be able to fix carbon and contain putative symbiotic factors including adhesins and other proteins potentially mediating host-microbe interactions (Siegl et al., 2011). Similarly, microbial communities in corals are highly diverse and, although several members may be true mutualists (Hernandez-Agreda et al., 2016), the clear established case is that of euphotic-zone corals and dinoflagellate algae of the genus Symbiodinium (Roth, 2014), which also occur in some bivalves and anemonies. Climate-change induced ocean acidification and warming appear to correlate with symbiont loss, leading to coral bleaching and death, and seriously affecting coral-reef ecosystems. The counterpart of these phototrophic symbioses, whereby fixed carbon is handed over to the host in exchange for nutrient collection and the availability of a stable environment, is that of chemotrophic symbioses in deep-sea or sediment fauna. Soon after the discovery of deep-sea vents and their cohort of exotic fauna in the late 1970s, symbiotic bacteria filling out completely the modified gut of the giant worm Riftia pachyptila led to propose a chemohautrophic symbiosis (Cavanaugh et al., 1981). Today the basis for this symbiosis is well known; the gammaproteobacterial symbiont, acquired anew at each generation, fixes carbon gaining energy by the oxidation of H₂S, while the polychaete worm transports O₂, NO₃⁻ and H₂S for the needs of its otherwise metabolically versatile symbiont (Röbøld et al., 2008). Similar chemotrophic symbioses also abound in the gills of deep-sea bivalves and other deep-sea metazoans and, more generally, ecto- and endosymbionts are frequently associated to animals living in oxygen-deprived marine settings, from the deep-sea to coastal sediments settings (Dubilier et al., 2008). These most often correspond to sulfide- or sulfur-oxidizing Gamma- or Epsilonproteobacteria, methanotrophic Gammnaproteobacteria or sulfate-reducing Deltaproteobacteria (Dubilier et al., 2008). In many cases, these symbioses are multiple, involving diverse endosymbionts with distinct metabolic capabilities. For instance, clones of the genus Bathymodiolus harbor dual symbioses with methanotrophic and thiotrophic bacteria (Duperron et al., 2007). The case of the oligochaete Olavius algarvensis is most remarkable; this worm, lacking mouth, gut and nephridia, hosts co-existing sulfide-oxidizing (Gammaproteobacteria) and sulfate-reducing (Deltaproteobacteria) bacteria and benefits from the versatile metabolism of its hosted, sulfur-cycling consortium (Woyke et al., 2006).

Microbial symbioses with plants are also widespread and possibly explain their evolutionary and ecological success. Over 80% plants establish symbioses with specific fungi that provide humidity, nitrogen and phosphorous to the roots in exchange for fixed carbon. The mycorrhizal symbiosis is thought to be at the origin of terrestrial colonization by plants (Field et al., 2015), the algal ancestor of land plants likely being pre-adapted for symbiosis (Delaux et al., 2015). Almost as important are nitrogen-fixing symbioses established with bacteria from various phyllogenetic groups, the most successful and best-known of which are the alphaproteobacterial Rhizobiales. Rhizobium-legume symbioses have co-evolved leading to specific signaling processes that involve, among others, immunity suppression during the establishment and maintenance of the symbiosis (Geurts et al., 2016; Gourion et al., 2015).

Prokaryotic symbioses established with protists are more difficult to identify because many protists being grazers, it is difficult to distinguish endosymbionts from prey. Also, although many classical works describe the recurrent presence of endobionts in protists, they often fail to inquire further about their role and many of these potential endosymbionts remain cryptic. Known endosymbionts in microbial eukaryotes often involve photosynthetic cyanobacteria, nitrogen-fixing/recycling bacteria, or methanogenic archaea (in some hydrogenosome-bearing anaerobic protists) (Nowack and Melkonian, 2010). Symbioses with cyanobacteria occur frequently, and one of them,
particularly successful, gave rise to the chloroplast more than 1 billion years ago (see below). An intriguing symbiosis occurs in oligotrophic oceans between marine haptophyte algae and the derived, nitrogen-fixing U-CYNA cyanobacteria, which have secondarily lost the ability to make photosynthesis (Zehr et al., 2016). Like in the case of animals, and although not well-known, protist-prokaryote symbioses are particularly abundant in suboxic environments (Bernhard et al., 2000). Hydrogen-transfer appears important, as in methanogenic symbioses of ciliates and other anaerobic protists or the ectosymbiosis established by the breviate *Lentisia limosa* with EPSilonproteobacteria of the genus *Arcobacter* (Hamann et al., 2016), also frequent epibionts in deep-sea vent fauna, including alvinellids and shrimps. Similarly, sulfide-oxidizing epilproteobacterial epibionts are associated to *Symbiontida* euglenozoans in anoxic sediments (Edgcomb et al., 2011b). Multiple symbioses, including sometimes ecto- and endosymbionts, such as in the case of oyxnomads with spirochetes and Bacteroidales (Noda et al., 2009), have also been described. Some multiple symbioses include three or more partners. For instance, a ciliate from deep, anoxic, sulfide-rich sediments contains at least three different endosymbionts in specialized membrane-bound sub-cellular regions, including two sulfate-reducing Deltaproteobacteria (related to the families Desulfulbaceae and Desulfbacteraceae) and a methanogenic archaeeon; it possibly also includes one endosymbiotic Bacteroidetes and a Type I methanotroph (Edgcomb et al., 2011a). This further supports the idea that multiple symbioses offer to the host the benefit of synergistic metabolisms in this type of anoxic environments.

2.3. Prokaryote-prokaryote syntrophy

Most symbioses imply metabolic interactions. This is particularly true in the case of prokaryote-prokaryote symbioses. Microbial syntrophy implies obligitory mutualistic metabolic cooperation. This type of symbiosis is especially important, sometimes mandatory, in low-energy environments lacking strong electron acceptors and where many endergonic reactions can become exergonic only when one partner acts as an electron sink for the other (Morris et al., 2013). Syntrophic interactions mediated by interspecies hydrogen or formate transfer are possibly the most conspicuous and profuse in the planet, being essential in the anaerobic conversion of organic matter down to methane in anoxic sediments worldwide (McInerney et al., 2009). In addition of playing key roles in the intermediate steps of the carbon cycle, syntrophic interactions can also involve the exchange of organic, sulfur- or nitrogen-containing molecules (Morris et al., 2013).

Symbiotic partners can be so well integrated that it is sometimes possible to isolate and study such consortia (though not the individual partners). Classical examples are *Methanococcus omelianski*, formed by an ethanol fermenter producing acetate and hydrogen and a methanogen using that hydrogen to reduce CO2 to CH4 (Bryant et al., 1967), or phototrophic consortia of the type *Chlorochromatium* (van Embree et al., 2015). Genomic analysis show that the non-motile chlorobi can sense and move the consortium towards light and provide sulfide that is used as electron donor for photosynthesis (Cerqueda-Garcia et al., 2014; Liu et al., 2013). Nonetheless, identifying syntrophic partners in natural ecosystems is difficult due to the high diversity and complexity of microbial communities in this kind of settings, and to the fact that a large part of this diversity corresponds to divergent bacteria and archaeal taxa (including the recently discovered Asgard clades) for which the metabolic potential is only beginning to be explored via metagenomics and single-cell genomics (Baker et al., 2016; Biddle et al., 2011, 2006; Rinke et al., 2013; Sousa et al., 2016). However, unveiling cryptic microbial symbioses in complex environments like sediments, soils or mats, will require ultimately the visualization of metabolic exchange, which can be achieved by a variety of sophisticated techniques including fluorescent in situ hybridization coupled to secondary ion mass spectrometry or bioorthogonal noncanonical amino acid tagging coupled to fluorescence-activated cell sorting (Hatzenpichler et al., 2016; Orphan, 2009).

Deltaproteobacteria and Gram positive bacteria together with methanogen-related archaea are the better-known and most prominent groups of prokaryotes involved in syntrophic interactions in sediments and similar environments worldwide (McInerney et al., 2008). Deltaproteobacteria are metabolically versatile; most are sulfate-reducing, but some have secondarily lost this ability in favor of anaerobic fermentation or aerobic heterotrophy, while some are predatory, like the bdellovibrios (Madigan et al., 2014). Deltaproteobacterial fermenters are usually engaged in syntrophic partnership with methanogenic archaea, some genera being obligatory syntrophic (*Syntrophus, Syntrophobacter*), as some Firmicutes are (*Syntrophomonas*). Interestingly, this type of mutualism can evolve extremely fast, as has been shown by experimental evolution of co-cultures of the sulfate-reducing *Desulfovibrio vulgaris* and the archaean *Methanococcus maripaludis*. When these organisms, with no known history of previous interaction, were grown in co-culture, the deltaproteobacterium stopped reducing sulfate (which consumes hydrogen) and started to ferment, liberating hydrogen that was used by the methanogen for its metabolism. This syntrophy was efficient (evolved co-cultures grew up to 80% faster and were up to 30% more productive in terms of biomass per mole of substrate) and evolved in less than 300 generations (Hillesland and Stahl, 2010).

In addition, sulfate-reducing deltaproteobacteria also establish syntrophic interactions with methanotrophic archaea (methanogen-related AMME lineages that most likely perform partly reversed methanogenesis). Anaerobic oxidation of methane (AOM) was thought to be impossible because energetically unfavorable, until it was discovered that this could be achieved by syntrophic consortia in marine sediments worldwide (Boetius et al., 2000; Orphan et al., 2001; Thauer and Sima, 2008). AOM is also crucial for the global carbon cycle, being responsible for the transformation of a large part of methane in sediments (Wegener et al., 2016). Remarkably, direct cell-cell electron transfer between by nanowire-like structures has been demonstrated in this type of consortia (Wegener et al., 2015). Direct interspecies electron transfer has raised a lot of attention, because some organisms (cable bacteria) can exchange electrons over long distances; it typically occurs in anaerobic methane-producing and methane-consuming communities (Lovley, 2016).

Metagenomic and metatranscriptomic analyses are starting to provide clues into syntrophy for anaerobic metabolic cooperation (Sieber et al., 2012). This kind of studies often reveals the involvement of more than two syntrophic partners in a consortium, as is the case in AOM (Pernthaler et al., 2008). Indeed, it appears that multi-member syntrophic communities are generally favored in nature even in environments not as challenging as anoxic ecosystems, since they reduce the individual metabolic burden in favor of cross-feeding and synergistic growth. This is suggested by experimental evolution of co-existing strains that tend to reduce the biosynthetic cost of amino acid synthesis promoting cooperative interactions, which are stronger for the more costly amino acids (Mee et al., 2014). Although not well understood, multidimensional interactions are also important in natural anaerobic methanogenic communities involved in carbon cycling (Embree et al., 2015).

Most known prokaryote-prokaryote symbioses involve cell-cell contact but not direct endosymbiosis. However, rare cases of endosymbiosis of prokaryotes within prokaryotes are known. Intracellular bacterial endosymbionts exist within endosymbiotic bacteria in mealybugs (von Dohlen et al., 2001), and the symbiont can be replaced in this nested symbiosis (Husnik and McCutcheon, 2016). A similar nested symbiosis involves intracellular alphaproteobacteria residing within tick mitochondria (Sassera et al., 2006).
3. Symbiosis in eukaryotic evolution

If Margulis was well aware of the importance of symbiosis in nature, her major contribution was to propose a symbiogenetic origin of chloroplasts, mitochondria and the eukaryotic cell itself. Except for details regarding the partners and the specific mechanisms, which still remain to be elucidated, she was essentially right.

3.1. Symbiosis and the origin of plastids

Photosynthesis originated very early during the evolution of bacteria, as attested by the phototactic behavior observed in the oldest fossil stromatolites, more than 3.4 Gya (Awramik, 1992) and, especially, by the oxygenation of Earth’s atmosphere that started at least 2.4 Gya (Lyons et al., 2014). Whereas the organisms that built the oldest stromatolites were most likely anoxygenic photosynthesizers, the enormous amount of oxygen that was necessary to oxidize the atmosphere was undoubtedly produced by cyanobacteria, the only bacterial lineage able to carry out oxygenic photosynthesis. Thus, for more than 2 billion years, photosynthetic primary production was exclusively assured by bacteria, as photosynthetic eukaryotes only evolved ~1 Gya (Eme et al., 2014) (see below). Interestingly, eukaryotes did not evolve a brand new oxygenic photosynthetic metabolism but, as Mereschkowsky (Mereschkowsky, 1905, 1910) and, later on, Margulis (Sagan, 1967) posited, they just adopted it from bacteria by the endosymbiosis of a cyanobacterium that became the first photosynthetic plastid. The number of endosymbiotic events and the identity of the partners involved have been the matter of debate for decades (Moreira and Philippe, 2001). Nevertheless, phylogenetic analysis of plastid-encoded genes points to a single origin of all eukaryotic plastids, namely, a single initial cyanobacterial endosymbiosis within a heterotrophic eukaryotic host (an evolutionary event that is known as primary endosymbiosis) (Archibald, 2009; Keeling, 2013; Moreira et al., 2009; Rodriguez-Ezpeleta et al., 2009). Recently, it has been shown that the cyanobacterial endosymbiont most likely belonged to the Gloeomargaritales, a group of unicellular cyanobacteria widely distributed in freshwater systems (Ponce-Toledo et al., 2017). Although virtually all eukaryotic plastids derive from this endosymbiont, a second case of primary endosymbiosis has been identified, involving a cyanobacterium from a different group (the Synechococcus-Prochlorococcus clad) that established a primary symbiosis with testate amoebae of the genus Paulinella (Nakayama and Ishida, 2009). In all these symbioses, a recurrent observation is that genes from the cyanobacterial endosymbiont are massively transferred into the nucleus of the eukaryotic host, a phenomenon called endosymbiotic gene transfer –EGT- (Martin et al., 1998). Since several proteins encoded by those transferred genes remain necessary for plastid function, this implies that a system to address them into the plastid has to evolve.

The initial primary endosymbiosis gave rise to three groups of photosynthetic eukaryotes: red algae, green algae and plants, and glaucophyte algae. They form a monophyletic supergroup called Archaeplastida (Adl et al., 2005) or Plantae (Cavalier-Smith, 1982). However, photosynthetic plastids can be noticed in many other eukaryotic lineages. In contrast with the initial cyanobacterial symbiosis, these plastids originated by the endosymbiosis of red or green algae in other eukaryotic hosts. These eukaryote-within-eukaryote secondary endosymbioses are at the origin of a vast diversity of algae, including the euglenids and chlorarachniophytes (with green plastids) and the cryptophytes, haptophytes, dinoflagellates, stramenopiles (diatoms, brown and golden algae, etc.) and several other groups containing red plastids. In some cases, plastids have lost their photosynthetic activity, but they remain in the cell to carry out other functions, a notorious case being that of the apicoplast, a non-photosynthetic organelle found in the Apicomplexa, a major group of parasites (Keeling, 2013). Several lineages have even more complex histories as they may have acquired secondary plastids by tertiary endosymbioses (e.g., haptophyte and diatom plastids found in some dinoflagellates) or may have replaced a secondary plastid by another one by secondary plastid replacement (e.g., green algal plastids in the dinoflagellate Lepidodinium, which replaced the original red algal one).

Whereas the monophyly of the Archaeplastida is widely accepted, the evolutionary relationships among the lineages containing red algal secondary plastids remain controversial (in the case of euglenids and chlorarachniophytes it is clear that they derive from two independent endosymbioses with two different green algae). As they account for a substantial fraction of the whole eukaryotic diversity, the phylogenetic uncertainty around the secondary red lineages represents a major open question in eukaryotic evolution. It was initially thought that all eukaryotes with red algal plastids were monophyletic (the ‘Chromalveola hypothesis’ (Cavalier-Smith, 1999)), as supported by the phylogeny of plastid-encoded genes where all red algal-derived secondary plastids form a monophyletic group. However, the phylogeny of nucleus-encoded genes does not support the expected monophyly of the different hosts. To reconcile both observations, it has been proposed that red secondary plastids were acquired by one eukaryotic lineage and this foundational event was followed by an undetermined number of tertiary endosymbioses that spread the red plastids across a variety of lineages (Baurain et al., 2010). As in the case of the primary endosymbiosis, the secondary endosymbioses have been accompanied by massive EGT from the nucleus of the red or green algal secondary plastid to the nuclei of the hosts, which are therefore highly chimeric since they contain genes with very different evolutionary origins. This complicates the study of these organisms as the phylogenies of the different genes may be discordant and difficult to interpret (Moreira and Deschamps, 2014). At any rate, the contemporary diversity and ecological significance of photosynthetic eukaryotes demonstrates the power of symbiosis to generate evolutionary novelty in eukaryotes. Moreover, this is an ongoing phenomenon as many eukaryotes may acquire facultative photosynthetic abilities through the endosymbiosis of a variety of algae. This is very common in animals (e.g., zooxanthellae in corals and other invertebrates) but also in single-celled eukaryotes (e.g., endosymbiotic green algae in some ciliates). Some of these endosymbioses may become obligatory, especially by EGT if essential genes are transferred from the endosymbiont to the host, and enlarge the list of eukaryotic photosynthetic lineages.

3.2. Symbiosis and the origin of mitochondria

In aerobic eukaryotes, mitochondria are the energy factories of the cell, the organelles were oxygen respiration is used to convert biochemical energy from nutrients into ATP. The presence of a small genome within these organelles prompted biologists to speculate that they derive from endosymbiotic organisms, though the diversity of their genome size and physical organization and some other peculiarities (e.g., variations in the genetic code in mitochondria of some eukaryotic lineages) casted doubts on their monophyly and the nature of the endosymbiotic organisms at their origin (Gray and Doolittle, 1982). However, sequence similarities between mitochondrial and bacterial rRNAs (Bonen et al., 1977) and, especially, phylogenetic analysis of e-type cytochrome sequences (Schwartz and Dayhoff, 1978) unequivocally demonstrated the bacterial origin of mitochondria, closely related to the nonsulfur purple bacteria, known today as Proteobacteria. Subsequent analyses showed that all eukaryotic mitochondria originated from a single ancestor. Much more debated was the timing of the endosymbiosis that gave rise to mitochondria during eukaryotic evolution. In fact, a variety of anaerobic eukaryotes lack conventional mitochondria and, in some cases, no trace of a mitochondrial genome was detectable. Moreover, some of these eukaryotes (microsporidia, trichomonads, diplomonads) branched at the base of the first eukaryotic phylogenies reconstructed using 18S rRNA sequences (Sogin, 1991), which led to propose that these organisms were
some kind of living fossils that predated the mitochondrial endosymbiosis: the ‘Archezoa hypothesis’ (Cavalier-Smith, 1989).

This hypothesis did not survive the passing of time when genes of clear mitochondrial origin were discovered in the nuclear genomes of ‘amitochondriate’ eukaryotes and when alternative phylogenetic markers replaced these lineages far from the base of the eukaryotic tree (Roger, 1999). In fact, as in the case of plastids described above, the mitochondrial endosymbiosis has been followed by massive EGT from the mitochondrial ancestor to the nuclear host genome and several of the proteins encoded by the transferred genes are targeted back into the mitochondria by a specialized translocation mechanism. Thus, even in the case of complete loss of the mitochondrial genome (Dyall et al., 2004; Gray, 2012; Timmis et al., 2004) (Dyall et al., 2004; Gray, 2012; Muller et al., 2012; Schwartz and Dayhoff, 1978), some mitochondrial activities can be maintained thanks to the genes transferred into the nucleus. This is the case for many anaerobic eukaryotes, which do not carry out aerobic respiration any more, but retain mitochondria-derived organelles that may have energy-related activities (such as ATP synthesis in hydrogenosomes (Müller, 1975)) or other functions (such as the synthesis of Fe–S clusters (Gill et al., 2007)). The presence of conventional mitochondria or of mitochondria-related organelles (MROs) in all major eukaryotic lineages strongly advocates for a mitochondrial endosymbiosis that predated the diversification of all contemporary eukaryotes. Thus, not only photosynthetic eukaryotes possess chimeric genomes because of the EGT of cyanobacterial genes, but all eukaryotes do as they contain genes of mitochondrial origin in their nuclear genomes. A possible outstanding exception has come with the characterization of the anaerobic oxymonad Monocercomonoides sp., which appears to have completely lost all typical mitochondrial genes (Karnowska et al., 2016). This whole absence of mitochondrial genes reflects a secondary loss, as the phylogenetic position of this species confirms that it derives from mitochondriate ancestors. Nonetheless, this discovery has interesting evolutionary implications as it demonstrates that mitochondria (or any form of MRO) are not a requirement for eukaryotic cell existence despite the fact that, historically, the mitochondrial symbiosis was indissolubly linked to eukaryotic origins.

3.3. Eukaryogenesis by symbiogenesis

As explained above, there is overwhelming evidence supporting that the mitochondrial endosymbiosis occurred before the diversification of all contemporary eukaryotic lineages. Thus, if we define eukaryogenesis as the whole evolutionary process leading from prokaryotic ancestors to the last common ancestor of all eukaryotes, this process is necessarily symbiogenic at the very least because of the ubiquity of mitochondria in eukaryotes. Mitochondria did not only provide an efficient energy metabolism, oxygen respiration, but genes of mitochondrial origin seem to be involved in a variety of cell activities in eukaryotes, such as DNA repair (Lin et al., 2007) and diverse metabolic functions (Thierngart et al., 2012). Therefore, mitochondrial acquisition not only represents an ancient event in eukaryotic evolution, but it probably entailed important changes at various functional levels of the host cell. Some authors have suggested that the mitochondrial endosymbiosis was the key event that triggered the diversification of contemporary eukaryotes (Philippe et al., 2000) and others have even proposed that it triggered the very origin of the eukaryotic cell itself (see below). At any rate, there is no doubt that contemporary eukaryotes are chimeras as they are formed by the integration of at least two types of cells: an ancient bacterium that transformed into mitochondria plus the host that acquired it. This fact is significant as it allowed to discard a series of classical hypotheses for the origin of eukaryotes that posited that all eukaryotic hallmarks features, including mitochondria and the other organelles, evolved by transformation of pre-existing structures of a prokaryotic cell. These hypotheses, collectively called ‘autogenous’ (Gray and Doolittle, 1982), were instrumental in the early opposition to Margulis’ symbiogenetic ideas but were abandoned when molecular biology and phylogenetics demonstrated the endosymbiotic origin of mitochondria and chloroplasts. Nevertheless, the term autogenous is still used by several authors to refer to the origin of the most idiosyncratic eukaryotic cellular structure: the nucleus (see below).

Whereas the nature of the mitochondrial ancestor was easy to determine by using phylogenetic analysis, which placed mitochondria within Proteobacteria more specifically, within Alphaproteobacteria (Andersson et al., 1998), the nature of the host and its nucleus have remained much more elusive. Analysis of the macromolecular structure of ribosomes of a variety of organisms revealed interesting similarities between eukaryotes and a group of ‘sulfur dependent bacteria’, which were called ‘eocytes’ (Lake et al., 1984). These organisms were later recognized to actually be a major subgroup of Archaea, the Crenarchaeota (Woese et al., 1990). Subsequent phylogenetic analyses of diverse proteins, especially when the first complete genome sequences became available, led to recognize that eukaryotic genomes appeared to be composed of two fractions with different functions and evolutionary origins: eukaryotic genes which function in translation, transcription, and replication (called ‘informational genes’) are closely related to archaeal homologs, whereas genes involved in energy and intermediary metabolism and in the synthesis of cell components, including amino acids, cofactors, the cell envelope, fatty acids, phospholipids, and nucleotides (referred to as ‘operational genes’) are more closely related to bacterial homologs (Rivera et al., 1998). The presence of bacterial-like genes in eukaryotic genomes was easy to explain as genes acquired by EGT from the mitochondrial endosymbiont. However, the presence of archaean-like genes was more intriguing and has been the focus of most debates on eukaryogenesis in recent decades. The hypothesis initially favored was that a third lineage different from the two prokaryotic ones (Archaea and Bacteria) but phylogenetically sister to the Archaea was at the origin of the eukaryotic nucleocyttoplasm and, therefore, the host of the mitochondrial endosymbiont. This hypothetical lineage would have evolved ‘autogenously’ many of the hallmark eukaryotic features (notably a complex cytoskeleton and endomembrane system, including the nucleus). This idea fitted well the rRNA-based Woeseian tree of life that supports the division of living beings into three major lineages (or Domains): Archaea, Bacteria and Eucarya (Woese et al., 1990) and by this reason it was called the ‘3D’ hypothesis (Gribaldo et al., 2010). Despite some contradictory results (Lake, 1987), phylogenetic analysis of protein markers also appeared to support this view, as the archaean-like component of eukaryotes tended to branch deep, before the diversification of contemporary archaean lineages (Yutin et al., 2008).

This situation has radically changed in recent years thanks to significant improvements in phylogenetic reconstruction methods and in the taxon sampling available. The first ones are mostly related to the implementation of site-heterogeneous mixture models of sequence evolution, which better account for the substitution patterns observed in sequence datasets (Lartillot and Philippe, 2004). The application of those methods to the reconstruction of universal trees has yielded phylogenies where eukaryotes do not form an independent branch but emerge within the Archaea (Williams et al., 2013, 2012). On the other hand, taxon sampling improvement has been made possible by the exploration of a variety of environments using new tools, which have led to the discovery of a vast diversity of archaean lineages (Pester et al., 2011), some of which are phylogenetically related to the Crenarchaeota (the ‘eocytes’) and form a major archaean supergroup, the TACK or Proteoarchaeota (Guy and Ettema, 2011). Among them, a recently discovered lineage has had a major impact on our view of eukaryogenesis: the Asgard archaean clades (Thor-, Odin-, Heimdall and Lokiarchaeota) (Spang et al., 2015; Zaremba-Niedzwiedzka et al., 2017). Although not yet cultured in the laboratory, composite genome sequences reconstructed from sediment metagenomes coming from a variety of sources, from deep sea to hot springs, have shown that...
Asgard lineages appear to contain more genes shared with eukaryotes than any other archaea and, even more relevantly, phylogenetic analysis of conserved markers retrieves the monophyly of eukaryotes and Asgard archaea (Spang et al., 2015; Zaremba-Niedzwiedzka et al., 2017). These results provide strong support to the so-called ‘2D models’, which posit that there are only two primary domains (Bacteria and Archaea) and that eukaryotes constitute a derived secondary domain, originated from some sort of mix of lineages of the two prokaryotic domains (Fig. 1). Thus, the archaeal-like component of eukaryotic genomes, namely the informational genes, seems to have originated from within the archaea. These discoveries have led not only to discard the 3D models mentioned above, but also to further discredit other models that proposed that these informational genes, and even the nucleus itself, were acquired from giant viruses, a miscellaneous group of viruses belonging to the nucleocytoplasmic large DNA virus (NCLDV) superfamily (Bell, 2009; Boyer et al., 2010; Forterre and Gaia, 2016; Raoult et al., 2004). Indeed, even before the discovery of Lokiarchaeota and other Asgard archaea, it had already been shown that eukaryotic-like genes found in NCLDV had been acquired by the viruses from their hosts via horizontal gene transfer (HGT) and not the other way round. This can be masked by the high evolutionary rate of viruses that can distort phylogenetic reconstruction, but when accurate phylogenetic methods are applied to retrace the evolutionary history of these genes, their eukaryotic origin becomes clear (Lopez-Garcia and Moreira, 2015; Moreira and Lopez-Garcia, 2015; Williams et al., 2011). Therefore, unless there are specific features that definitively associate the fossils/biomarkers with particular crown eukaryotic group, it is not possible to distinguish whether they represent stem or crown eukaryotes. Additionally, identifying with confidence microfossils as eukaryotic rather than bacterial is not a trivial task and only fossils combining large size, ornamented walls, complex ultrastructure, and a preservable composition are regarded as probable eukaryotes (Knoll, 2015). As such, while the oldest of microfossils of possible eukaryotic origin were described from 3200 Ma shales, their simple ultrastructure prevents from any definitive conclusion (Javaux et al., 2010). The geological record preserves part of life history, including evidence for early (chemical and morphological) traces of the existence of eukaryotes. When considering this evidence, it may be important to distinguish between ‘stem’ and ‘crown’ lineages. Briefly, all lineages that descend from LECA are often referred to as crown eukaryotes (Fig. 1). But before a full-fledged LECA evolved, a more or less long period of evolution occurred from the time when, depending on the model, the archaean ancestor of eukaryotes started to acquire eukaryotic features (Fig. 2A) or when the first eukaryogenetic archaea-bacteria consortium started to co-evolve together (Fig. 2B-D). Lineages diverging during that period are considered stem lineages and are now all extinct (but the one that successfully led to ‘crown’ eukaryotes). Any characteristic feature of eukaryotes evident in ancient microfossils can, in principle, be a property of either stem or crown organisms. Therefore, unless there are specific features that definitively associate the fossils/biomarkers with particular crown eukaryotic group, it is not possible to distinguish whether they represent stem or crown eukaryotes.

4. Eukaryogenesis: open questions

It has become clear now that eukaryogenesis occurred by symbiogenesis of archaea and bacteria (Fig. 1), supporting Margulis’ visionary idea, also adopted by later symbiogenetic models, that symbiosis was crucial in eukaryotic evolution, leading to an increase in average cell complexity. Symbiogenesis was an implicit conclusion from recent gene content and phylogenomic analyses suggesting that archaea of the TACK superphylum were more similar to eukaryotes (Martijn and Ettema, 2013; McInerney et al., 2014; Williams and Embley, 2014), which eliminated the possibility that a third, proto-eukaryotic lineage different from archaea ever existed, and became much more explicit when Lokiarchaeota and other Asgard archaea were discovered (Koonin, 2015; Lopez-Garcia and Moreira, 2015; Spang et al., 2015; Zaremba-Niedzwiedzka et al., 2017). However, despite the confirmation of the symbiogenetic origin of eukaryotes, many crucial questions remain open (reviewed in (Lopez-Garcia and Moreira, 2015)).

4.1. When did eukaryotes evolve?

The geological record preserves part of life history, including evidence for early (chemical and morphological) traces of the existence of eukaryotes. When considering this evidence, it may be important to distinguish between ‘stem’ and ‘crown’ lineages. Briefly, all lineages that descend from LECA are often referred to as crown eukaryotes (Fig. 1). But before a full-fledged LECA evolved, a more or less long period of evolution occurred from the time when, depending on the model, the archaean ancestor of eukaryotes started to acquire eukaryotic features (Fig. 2A) or when the first eukaryogenetic archaea-bacteria consortium started to co-evolve together (Fig. 2B-D). Lineages diverging during that period are considered stem lineages and are now all extinct (but the one that successfully lead to ‘crown’ eukaryotes). Any characteristic feature of eukaryotes evident in ancient microfossils can, in principle, be a property of either stem or crown organisms. Therefore, unless there are specific features that definitively associate the fossils/biomarkers with particular crown eukaryotic group, it is not possible to distinguish whether they represent stem or crown eukaryotes. Additionally, identifying with confidence microfossils as eukaryotic rather than bacterial is not a trivial task and only fossils combining large size, ornamented walls, complex ultrastructure, and a preservable composition are regarded as probable eukaryotes (Knoll, 2015). As such, while the oldest of microfossils of possible eukaryotic origin were described from 3200 Ma shales, their simple ultrastructure prevents from any definitive conclusion (Javaux et al., 2010). The oldest fossil showing robust structural evidence of eukaryotic affiliation

Fig. 1. Schematic representation of the tree of life and the origin of eukaryotes in an approximate historical framework.
discrepancies can be explained by a myriad of sources of variability and error (Eme et al., 2014; Kumar and Hedges, 2016; Roger and Hug, 2006) such as the relaxed molecular clock models and methods used, and the nature and treatment of fossil calibrations. The most recent analyses provide estimates for the age of LECA in the range of 1000–1600 Ma (Eme et al., 2014).

From these results, estimating the age of the mitochondrial endosymbiosis intrinsically depends on the model favored for eukaryogenesis. In the case of ‘mitochondria-early’ scenarios, which consider the mitochondrial endosymbiosis to be the initial triggering event of eukaryogenesis (Fig. 2B), the timing of the endosymbiosis can be inferred to be older than the oldest evidence for the occurrence of eukaryotes (i.e., 1.7 Ga). It is worth noting that the first organisms belonging to the eukaryotic lineage might have been morphologically indistinguishable from their prokaryotic ancestors. Consequently, any microfossil clearly distinguishable as eukaryotic (i.e., combining all the morphological features listed above) would be considerably more recent than the origin of the eukaryotic lineage itself, and thus, than the mitochondrial endosymbiosis. By contrast, ‘mitochondria-late’ hypotheses postulate that a significant number of specific eukaryotic-like features predated the acquisition of the mitochondrion (Fig. 2A, C–D). In this case, the timing of the mitochondrial endosymbiosis would be considerably closer to the age of LECA.

The timing of the primary plastid endosymbiosis can be more easily bracketed since it had to occur after LECA and before the last common ancestor of Archaeplastida. Most recent analyses estimate the latter to be 900–1300 Ma (Eme et al., 2014). It is worth noting that the Bangiomorpha fossils seem to be much older than any of the dates estimated from molecular clock data (Berney and Pawlowski, 2006; Eme et al., 2014; Parfrey et al., 2011; Sharpe et al., 2015) and at odds with all other microfossil calibrations for crown eukaryotes. This discrepancy can stem from a number of reasons. For example, either the taxonomic identification of Bangiomorpha fossils, or the estimated age of the rocks in which they were found, could be erroneous, although the latter is thought to be unlikely (see discussion in (Knoll, 2014; Parfrey et al., 2011)). Alternatively, it is possible that currently available molecular clock models do not capture properly the evolutionary process occurring in some lineages, leading to incorrect date estimates. The second case of primary stable endosymbiosis identified in Paulinella is considerably younger and has been estimated to have occurred ~60 Mya (Delaye et al., 2016). As to the time when the lineages containing secondary plastids evolved, their confused evolutionary relationships among their heterotrophic hosts currently hampers any serious attempt to date secondary and tertiary endosymbioses. Future phylogenomics progress on the reconstruction of the eukaryotic tree should help constraining the evolution of plastids derived from green and red algae.

4.2. Where did eukaryotes evolve?

One key question concerns the environmental setting where eukaryotes evolved and the metabolic nature of the symbiosis that gave rise to eukaryotes. Metabolic endosymbioses can drive evolution and have explanatory power in evolution (Lopez-Garcia and Moreira, 2015; O’Malley, 2015). Because prokaryotic symbioses essentially involve syntrophy, the eukaryogenetic symbiogenesis was most likely based on metabolic cooperation. Some insight can be obtained from the kind of ecosystems and metabolic networks established in ecosystems where Asgard archaean lineages occur. Lokiarchaeota, previously known as Deep Sea Archaeal Group (DSAG) or marine benthic group B, have been classically detected in sediments and microbial mats (Biddle et al., 2006; Jorgensen et al., 2013, 2012; Spang et al., 2015). Because their native habitats are anoxic, they must be anaerobic, which would be consistent with a putative dependency on hydrogen deduced from their gene content (Sousa et al., 2016); they have indeed been suggested to be involved in syntrophic interactions leading to dissim-
illy CH₄ oxidation (Biddle et al., 2006) or iron or manganese reduction (Jorgensen et al., 2012). Likewise, the rest of recently identified Asgard archaea are associated to sediments from various environmental settings (Zaremba-Niedzwiedzka et al., 2017). Obviously, these Asgard archaea constitute a collection of derived lineages that may be quite different from thearchaeal ancestor that established the eukaryogenic symbiosis with bacteria. Also, it might well be that there are otherarchaeal lineages that are even closer to eukaryotes, sincearchaeal diversity is far from being completely explored. Actually, a large diversity of deeply divergentarchaeal clades occurs precisely in anoxic settings, including the deep subsurface, sediments and different kinds of microbial mats (Baker et al., 2016; Biddle et al., 2006; Harris et al., 2013; Hedlund et al., 2013; Jorgensen et al., 2013, 2012; Kozubal et al., 2013; Lloyd et al., 2013). However, the fact that Asgard and many other deep-branching archaea are systematically found in anoxic settings where syntrophic interactions are the rule supports the idea that eukaryotes emerged from metabolically cooperative consortia in anoxic settings likely close to redox boundaries.

Indeed, symbiotic models that have advanced detailed mechanisms, including metabolic interactions, for the origin of eukaryotes, such as the hydrogen (Martin and Muller, 1998) and the syntrophy (López-García and Moreira, 2006; Moreira and López-García, 1998) hypotheses implied that eukaryotes evolved in anoxic settings where redox gradients occurred and that the mitochondrial ancestor was a facultative anaerobe (López-García and Moreira, 1999). This would offer the possibility, among others, to establish a symbiosis with a facultative aerobicalphaproteobacterium endowed with oxygen respiration but able to thrive in oxygen-depleted environments. A facultative anaerobic alphaproteobacterium, possibly microaerophilic (adapted to low-oxygen concentrations) as mitochondrial ancestor would make sense at a time when oxygen started to accumulate in the atmosphere and most of the oceanic water column was still anoxic (Scott et al., 2008). Both, hydrogen and syntrophy hypotheses proposed hydrogen transfer between symbiotic partners, in the former between a fermentative alphaproteobacterium that established as endosymbiont within a methanogenic archaeon (Martin and Muller, 1998), in the second between an ancestral sulfate-reducing fermentative myxobacterium (Deltaproteobacteria) and an endosymbiotic methanogenic archaeon (future nucleus) to which a versatile alphaproteobacterial methanotroph would have joined in a tripartite symbiosis (López-García and Moreira, 2006; Moreira and López-García, 1998). Thought to be exclusive of the Euryarchaeota branch, methanogenesis has recently been inferred in the Batharchaeota, within the TACK superphylum, making methanogenesis likely ancestral in archaea (Evans et al., 2015; Lever, 2016). Nonetheless, as mentioned in Section 2.3, the number and type of syntrophic interactions mediated by hydrogen among archaea and bacteria in sediments and other oxygen-depleted environments are far from fully understood. Multiple syntrophic interactions appear to be the rule in natural conditions, so that multiple symbioses at the origin of eukaryotes cannot be excluded on the basis of parsimony criteria because microbial ecology contradicts those arguments. At any rate, better understanding metabolic interactions between archaea and bacteria in this kind of ecosystems will greatly help proposing plausible metabolic interactions at the origin of the eukaryotic cell.

4.3. How did eukaryotes evolve?

Many models proposing symbiotic merging between archaea and bacteria at the origin of eukaryotes avoid providing detailed mechanistic scenarios explaining the origin of the most characteristic features of the eukaryotic cell. However, the most open and contentious questions in eukaryogenesis precisely relate to the mechanisms by which the different eukaryotic features appeared (reviewed in Lopez-García and Moreira, 2015), such that exposing the details is important to discriminate between existing models, refine them or propose more realistic ones. The different types of models that can currently be considered are schematically shown in Fig. 2. The most popular (Fig. 2, A) would correspond to the transposition of old ‘3D’ models to a ‘2D’ situation. Here, an archaeon would develop an endomembrane system by invagination of the plasma membrane and the nucleus in an ‘autogenous’ way together with the capacity of phagocytosis, which would allow it to engulf the mitochondrial ancestor (phagocytosing archaeon model, PhAT (Martijn and Ettema, 2013)). Many of these models are mechanistically naive but could be nurtured by much more developed character evolution scenarios from previous 3D phagocytosis-based models, including particular variants such as the Neomura hypothesis (Cavalier-Smith, 2014). A second model would correspond to the initial endosymbiosis of the alphaproteobacterial ancestor of mitochondria within an archaeon, the rest of archaeal features would evolve after the symbiosis established (Fig. 2, B; e.g. hydrogen hypothesis (Martin and Muller, 1998)). A third type of models would correspond to the endosymbiotic origin of the nucleus within one bacterium, the mitochondrion deriving for a second endosymbiosis with an alphaproteobacterium (Fig. 2, C; e.g. syntrophy hypothesis (López-García and Moreira, 2006; Moreira and López-García, 1998)). Finally, a fourth type of models would imply multiple serial endosymbioses (either within a bacterium or within an archaeon) with the participation of several bacteria that would have left an imprint in the eukaryotic nucleus by successive waves of EGT (Fig. 2, D) (Pittis and Gabaldon, 2016).

Although there is some debate about the specific identity of the mitochondrial ancestor within the Alphaproteobacteria and as to whether it was or not a facultative anaerobe, the most controversial point refers to whether mitochondria arrived early or not (Fig. 2). In the hydrogen hypothesis, mitochondria arrive early because the mitochondrial endosymbiosis is proposed to be the cause of eukaryogenesis (Martin and Muller, 1998). By contrast, in the phagocytosing archaeon model, the mitochondria arrive late because they are, in a way, the consequence of an autogenous eukaryogenic process that led to phagocytosis and the possibility to engulf exogenous bacteria (Martijn and Ettema, 2013). In the syntrophy hypothesis, the mitochondrial symbiosis is the last, highly successful, endosymbiotic event that determines the eukaryogenic closure, providing a clear selective metabolic advantage related to the oxygenic respiration (López-García and Moreira, 2006). Recent phylogenomic analyses exploring the relative timing of prokaryotic gene acquisition in eukaryotes indeed suggest that mitochondria arrived late and might additionally support successive, multiple symbioses (Pittis and Gabaldon, 2016).

Another crucial problem relates to the origin of the eukaryotic nucleus, which many eukaryogenetic models leave unexplained. In addition to the selective forces that led to the evolution of this idiosyncratic structure, which remain elusive (López-García and Moreira, 2015), the process by which the nuclear membrane evolved, like the timing of mitochondrial acquisition, also discriminates the hydrogen hypothesis from the rest of symbiogenetic models. The latter propose an autogenous origin of the endoplasmatic reticulum and the nuclear membrane (López-García and Moreira, 2006; Martijn and Ettema, 2013), which is consistent with cell biology knowledge and with the presence of many proteins involved in membrane remodeling and vesicle trafficking in archaea including, notably, Asgard archaea, but also other prokaryotes (Devos et al., 2004; Klinger et al., 2016; Spang et al., 2015; Surkont and Pereira-Leal, 2016). By contrast, the hydrogen hypothesis, speculates that the nuclear membrane and the endomembrane system would derive from the formation and progressive fusion of vesicles budding off the alphaproteobacterial ancestor of mitochondria around what would be the future nucleus (Gould et al., 2016). The production of lipid vesicles from that alphaproteobacterial endosymbiont would also explain, according to these authors, another fundamental problem in eukaryogenesis: the bacterial nature of eukaryotic membrane phospholipids.
Indeed, if the host that acquired mitochondria was an archaeon, then the archaean membranes should have undergone a profound transformation, substituting their membrane phospholipids, very different from bacterial ones, by bacterial phospholipids (Lombard et al., 2012), but also adapting all the integral membrane proteins to a very different physicochemical environment (López-García and Moreira, 2015). In Gould’s et al.’s view, the archaean lipids of the plasma membrane would have been substituted by mitochondria-derived lipids via vesicles protruding from the mitochondria that would fuse with the archaean membrane and progressively replace archaean lipids. This still leaves unexplained the selective forces that drove that putative membrane transition, as membrane proteins would still be adapted to archaean lipids. Likewise, bacteria-to-archaean horizontal gene transfer of genes involved in, e.g. fatty acid biosynthesis, has been evoked to explain a putative archaean-to-bacteria phospholipid transition, but some archaean containing those genes (e.g. haloarchaea) have typical archaean membranes (López-García et al., 2015; Nelson-Sathi et al., 2012), indicating that fatty acids are used for other purpose in the cell. Likewise, the absence of the glycerol-1-phosphate dehydrogenase (G1PDH) gene, which is responsible for the G1P isomer characteristically used to synthesize archaean phospholipids, seems to be absent from Lokiarchaeota and marine Euryarchaeota groups II/III, leading to propose that these archaean might have chimeric archaea-bacteria or even bacterial-like membrane phospholipids (Villanueva et al., 2016). Failure to identify G1PDH is intriguing. However, this observation cannot be taken for proof in the absence of direct access to the membrane lipids of these organisms. First, the genomes of these organisms have been reconstructed from metagenomes and are not necessarily complete (Deschamps et al., 2014; Iverson et al., 2012; Spang et al., 2015). Second, the gene might have significantly diverged in these organisms and, in any case, the absence of a classical G1PDH does not necessarily imply that G1P is not synthesized in an alternative way. It is indeed worth noting that archaean lipids were dominant in sediments with abundant Lokiarchaeota (DSAG) (Biddle et al., 2006). Finally, unusual butane- and pentanetriol-based tetraether lipids replacing glycolcerol-based backbones have been recently detected in a group of methanosamia sister to Group II/III archaean and Thermoplasmatales, and these special lipids seem also abundant in sediments (Becker et al., 2016), suggesting a modified, but not bacterial, nature for the lipids of this clade. Clearly, having access to the real biochemistry of the membrane lipids for these novel archaean groups is required.

5. Conclusion

analysis of Chlorobium chlorohromatii Ca3D reveals clues of the symbiosis in 'Chlorochromatium aggregatum'. ISME J. 8, 991–998.


