CellPress

Review

Open Questions on the Origin of Eukaryotes

Purificación López-García[1,*] and David Moreira[1]

Despite recent progress, the origin of the eukaryotic cell remains enigmatic. It is now known that the last eukaryotic common ancestor was complex and that endosymbiosis played a crucial role in eukaryogenesis at least via the acquisition of the alphaproteobacterial ancestor of mitochondria. However, the nature of the mitochondrial host is controversial, although the recent discovery of an archaeal lineage phylogenetically close to eukaryotes reinforces models proposing archaea-derived hosts. We argue that, in addition to improved phylogenomic analyses with more comprehensive taxon sampling to pinpoint the closest prokaryotic relatives of eukaryotes, determining plausible mechanisms and selective forces at the origin of key eukaryotic features, such as the nucleus or the bacterial-like eukaryotic membrane system, is essential to constrain existing models.

Trends

Eukaryotes arose from the endosymbiosis of an alphaproteobacterium in an unknown host. Eukaryogenic models diverge in the hypothetical host proposed.

Recent phylogenomic analyses and the discovery of archaea with seemingly more eukaryotic-like genes suggest that the mitochondrial host was an archaeon or had a vital archaeal contribution, excluding a third hypothetical proto-eukaryotic lineage different from archaea.

As the phylogenetic origin of eukaryotes gets clearer, mechanistic questions remain open: the type of metabolic symbioses involved, the timing of mitochondrial acquisition and, most importantly, the origin of the eukaryotic nucleus and bacterial-like membranes.

Plausible driving forces and processes for the evolution of the eukaryotic nucleus are missing.

If the mitochondrial host was an archaeon, a difficult-to-explain archaeal-to-bacterial membrane transition is required.

A Long-Lasting Query

The origin of the eukaryotic cell was a major evolutionary event that led to a wide diversification of lineages displaying very different morphologies, several of which independently evolved towards multicellularity [1]. Compared to the average prokaryotic cell, the average early eukaryotic cell represented a considerable increase in structural complexity, typified by the presence of an endomembrane system delimiting a hallmark eukaryotic feature, the nucleus, and membrane-bound organelles, notably mitochondria. In the past decades progress in cell and molecular biology, microbial diversity studies, and, most of all, comparative genomics and molecular phylogeny have all helped to better constrain the nature of that transition. Several facts related to the early evolution of eukaryotes have been unambiguously established including, among others, a mixed **archaeal**–bacterial legacy in eukaryotic genomes (see Glossary; Box 1). It is now clear that the last eukaryotic common ancestor (LECA) was a fairly complex organism already possessing major idiosyncratic features associated with extant eukaryotes (Box 2). It is also recognized that **endosymbiosis** did play a crucial role in **eukaryogenesis** and that the evolution of the **alphaproteobacterial** endosymbiont at the origin of mitochondria contributed the basics of energy metabolism [2,3] and largely shaped the eukaryotic genome [4], leading to innovations (Box 3). While more detailed knowledge about the last common eukaryotic ancestor and the alphaproteobacterial ancestor of mitochondria is still needed, the most fundamental open query relates to the nature of the host that acquired the mitochondrial ancestor and the eukaryogenic process itself. This is the point upon which most models for the origin of eukaryotes have traditionally diverged [5–8]. The recent discovery of an archaeal lineage, the Lokiarchaeota, sharing more, and seemingly more closely related, genes with eukaryotes [9] represents a significant advance towards the understanding of eukaryotic origins. From the **phylogenomics** perspective, this observation supports an archaeon as (or an archaeal contribution to) the host of mitochondria [10]. However, from the mechanistic perspective, this solution renovates fundamental open questions that relate to the specific evolutionary process and the underlying selective drivers at the origin of key eukaryotic features. In the following we very briefly

[1]Unité d'Ecologie, Systématique et Evolution, Centre National de la Recherche Scientifique (CNRS), Université Paris-Sud/Paris-Saclay, AgroParisTech, 91405 Orsay, France

*Correspondence: puri.lopez@u-psud.fr (P. López-García).

Box 1. Established Facts About Eukaryotic Origins

Two major cell structural types, prokaryotic and eukaryotic, exist. The eukaryotic cell is, on average, structurally more complex, possessing an endomembrane system with Golgi apparatus, lysosomes or peroxisomes, and endoplasmic reticulum (continuous with the nuclear membrane). Transcription and translation are thought to be generally coupled in prokaryotes (while mRNA is being synthesized, ribosomes start protein synthesis [61]); however, whereas eukaryotic transcription occurs in the nucleus, translation takes place in the cytoplasm.

There are three classically recognized phylogenetic domains of life: Archaea and Bacteria (both prokaryotic), and Eucarya [62]. Established in the 1970s with the first universal molecular phylogenetic analyses [63], this tripartite division has been validated over the years with thousands of genes and genomes from cultured and environmental lineages [64]. Although the three domains share basic biochemistry, genetic code, and some universally conserved molecular complexes (ribosome, membrane ATPase) [65], each has distinctive characteristics: a different DNA replication machinery for bacteria [66], unique ether-linked isoprenoid membrane phospholipids in archaea [57], and various complex cellular components and processes in eukaryotes.

In addition to universal and domain-specific traits, eukaryotes specifically share exclusive characteristics with either bacteria or archaea, suggesting some type of chimeric heritage. Eukaryotic machineries involved in informational processes (replication, transcription, translation) are more similar to, or share homologs only with, archaea [51,67,68]. Genes involved in energy and carbon metabolism, and membrane phospholipids are bacterial-like [57].

Historically, prokaryotes pre-date eukaryotes, as supported by two independent sources of evidence: the fossil record and the universal occurrence of mitochondria (or derivatives) in extant eukaryotes. The oldest unambiguous eukaryotic microfossils date back to ~2 Ga [69], compatible with molecular-dating inferences for LECA [70,71]. By contrast, geochemical isotope records support the likely occurrence of bacterial and/or archaeal metabolisms (methanogenesis, sulfate reduction, nitrogen fixation) much earlier (3.2–3.4 Ga) [72–74], and the 2.4 Ga atmospheric oxygen hike clearly attests to the prior evolution of cyanobacterial oxygenic photosynthesis [75]. Consequently, microbial communities at 3.5 Ga [76] were exclusively prokaryotic; eukaryotes appeared >1 Ga later. Likewise, compelling evidence shows that parasitic and anaerobic protists lacking typical mitochondria lost them secondarily: they possess genes of mitochondrial origin in nuclear genomes and mitochondria-related organelles [77]. Therefore, LECA already possessed mitochondria. Because mitochondria evolved from alphaproteobacteria, itself a derived bacterial lineage, bacteria had diversified well before divergence of current eukaryotic lineages.

Finally, symbiosis did play a crucial role in the evolution of the eukaryotic cell, at least via the acquisition of mitochondria from an alphaproteobacterial endosymbiont [78].

Box 2. The Nature of the Last Common Eukaryotic Ancestor (LECA)

Ultrastructural and phylogenetic studies have provided overwhelming support to the idea that all eukaryotes are monophyletic: namely that they derive from a single LECA. Ultrastructural and genomic characters widespread in contemporary eukaryotic lineages are likely to have been inherited vertically and can be used as source of information to infer ancestral characteristics. Using this comparative approach, many studies have contributed elements to reconstruct a detailed portrait of LECA [52,79]. It possessed all the paradigmatic eukaryotic features, including the nucleus (with nuclear lamina and nuclear pores), a complex endomembrane system, and a sophisticated tubulin/actin-based cytoskeleton. In relation to the endomembrane system, LECA possessed developed endocytic and exocytic pathways as well as concomitant vesicle-trafficking networks (including Golgi apparatus, lysosomes, and autophagosomes) which also involved the cytoskeleton. The cytoskeleton was also essential for phagocytosis, and the presence of this mechanism indicates that LECA was most likely heterotrophic and fed on organic matter, perhaps as a predator of other cells. Its metabolism was most likely aerobic because it possessed oxygen-respiring mitochondria. The cytoskeleton had also a key role in mitosis, cell cytokinesis, and cell motility (probably by several flagella). Meiosis was likely present, opening the possibility for some form of sexual reproduction. The genome contained introns, making necessary a splicing system, likely integrated in a sophisticated gene regulation machinery that also included the activity of small non-coding RNAs and RNA interference. This list is not exhaustive because many other processes, such as ubiquitination and proteasome-mediated degradation, were also present. Coding for all those characters, some of them based on the participation of hundreds of different proteins, requires a very large number of genes. Conservative estimates suggest that the eukaryotic ancestor had a genome with at least 4000–5000 genes [52,79]. This implies that LECA was complex, fully comparable to many modern eukaryotes, and that the toolkit for eukaryotic cell components was established very early. The subsequent evolution of eukaryotic lineages mostly involved fine-tuning of those components rather than major evolutionary innovations (two major exceptions are the endosymbiotic acquisition of photosynthesis and the multiple origins of multicellularity, both at the origin of massive eukaryotic evolutionary radiations).

Glossary

Alphaproteobacteria: highly diversified and metabolically versatile class of bacteria within the phylum Proteobacteria from which the ancestor of mitochondria evolved.
Archaea: one of the three classically recognized domains of life and one of the two primary phylogenetic domains. Archaea exhibit prokaryotic cell structure. They are traditionally divided into two main branches, the Euryarchaeota and the TACK (Thaumarchaeota, Aigarchaeota, Crenarchaeota, Korarchaeota) superphylum or Proteoarchaeota.
Autogenous models: hypotheses postulating that the endomembrane system and the nucleus result from the invagination of the plasma membrane in a proto-eukaryotic or a prokaryotic ancestor. Historically, before the general acceptance of the endosymbiotic origin of mitochondria and chloroplasts, the term also referred to the internal development of these organelles from endogenous endomembranes.
Deltaproteobacteria: proteobacterial class encompassing predominantly anaerobic, sulfate-reducing bacteria, fermentative syntrophic bacteria (e.g., *Syntrophomonas*) and myxobacteria.
Endosymbiosis: symbiotic relationship in which one partner (endosymbiont) is within the other (host).
Endosymbiotic gene transfer: transfer of genes from the genome of the endosymbiont to that of the host; in some cases it can lead to complete genome extinction.
Eocyte: term historically applied to the archaeal clade that was more similar to eukaryotes based on ribosomal proteins (originally, the Crenarchaeota); by extension, some authors apply it today to the TACK superphylum.
Eukaryogenesis: process that led to the evolution of the eukaryotic cell from prokaryotic ancestors.
Hydrogenosome: hydrogen-producing organelle that evolved from mitochondria.
Mitochondria-related organelle (MRO): include hydrogenosomes, mitosomes, and, in general, organelles derived from mitochondria.
Mitosome: a type of MRO, a genome-lacking simple mitochondrial remnant where Fe–S clusters are assembled.