

Endosymbiotic selective pressure at the origin of eukaryotic cell biology

Parth K Raval¹, Sriram G Garg², Sven B Gould^{1*}

¹Institute for Molecular Evolution, Heinrich-Heine-University Düsseldorf, Dusseldorf, Germany; ²Evolutionary Biochemistry Group, Max-Planck Institute for Terrestrial Microbiology, Marburg, Germany

Abstract The dichotomy that separates prokaryotic from eukaryotic cells runs deep. The transition from pro- to eukaryote evolution is poorly understood due to a lack of reliable intermediate forms and definitions regarding the nature of the first host that could no longer be considered a prokaryote, the first eukaryotic common ancestor, FECA. The last eukaryotic common ancestor, LECA, was a complex cell that united all traits characterising eukaryotic biology including a mitochondrion. The role of the endosymbiotic organelle in this radical transition towards complex life forms is, however, sometimes questioned. In particular the discovery of the asgard archaea has stimulated discussions regarding the pre-endosymbiotic complexity of FECA. Here we review differences and similarities among models that view eukaryotic traits as isolated coincidental events in asgard archaeal evolution or, on the contrary, as a result of and in response to endosymbiosis. Inspecting eukaryotic traits from the perspective of the endosymbiont uncovers that eukaryotic cell biology can be explained as having evolved as a solution to housing a semi-autonomous organelle and why the addition of another endosymbiont, the plastid, added no extra compartments. Mitochondria provided the selective pressures for the origin (and continued maintenance) of eukaryotic cell complexity. Moreover, they also provided the energetic benefit throughout eukaryogenesis for evolving thousands of gene families unique to eukaryotes. Hence, a synthesis of the current data lets us conclude that traits such as the Golgi apparatus, the nucleus, autophagosomes, and meiosis and sex evolved as a response to the selective pressures an endosymbiont imposes.

*For correspondence: gould@hhu.de

Competing interest: The authors declare that no competing interests exist.

Funding: See page 14

Received: 14 June 2022

Accepted: 31 October 2022

Published: 10 November 2022

Reviewing Editor: Agnieszka Chacinska, IMol Polish Academy of Sciences, Poland

© Copyright Raval et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

Introduction

'A scientist in his laboratory is not a mere technician: he is also a child confronting natural phenomena that impress him [her] as though they were fairy tales' (Marie Curie).

In evolutionary biology, the transition from prokaryotic to eukaryotic life was a true game changer. Eukaryogenesis involves the origin of new cell biology, genetics, and an unprecedented emergence of morphological diversity. Historically, the prokaryote-eukaryote divide was based on observed differences in morphology and in turn defined this aboriginal branch in the tree of life by their lack of traits that eukaryotes possess (*Stanier and Van Niel, 1962*). Phylogeny and biochemistry separate prokaryotes into bacteria and archaea (*Fox et al., 1980; Koga et al., 1998; Woese et al., 1990*) and document the dichotomy of pro- and eukaryotes, which is further evident in the number of protein families (*Rebeaud et al., 2021*), average protein length (*Brocchieri and Karlin, 2005*), cellular and morphological complexity (*Stanier et al., 1963*), and the overall prevalent contribution to the planet's biomass (*Bar et al., 2018*).

For decades the field of eukaryogenesis speculated on the existence of a eukaryotic lineage with intermediate biology bridging the prokaryote-eukaryote divide, an elusive grade known as archezoa (*Cavalier-Smith, 1987*). For curious reasons (see *Martin et al., 2017a*, for details) this search focused on a eukaryotic phylum lacking a mitochondrion (*Cavalier-Smith, 1987; Speijer, 2020*), but not necessarily one lacking a nucleus or endoplasmic reticulum (ER). Varying models, but with the common

theme of promoting an autogenous origin of a last eukaryotic common ancestor (LECA) independent of a bacterial partner, were proposed (reviewed in [Martin et al., 2015](#)). Through the identification of hydrogenosomes and mitosomes (reduced mitochondria; [Tovar et al., 1999](#); [Williams et al., 2002](#)) and modern phylogenomics ([Burki et al., 2020](#)), we now understand that the biology of LECA matched that of extant garden variety protists. This might appear trivial from today's perspective, but reaching this consensus and settling on a mitochondrion-bearing LECA took decades. LECA evolved from an archaeal host cell and its endosymbiotic alphaproteobacterial partner ([Imachi et al., 2020](#); [Wu et al., 2022](#); [Zaremba-Niedzwiedzka et al., 2017](#)) and could have been syncytial and fungus-like, with the first gametes budding off as a selectable unit, in what one could describe to be a flagellated protist ([Garg and Martin, 2016](#); [Skejo et al., 2021](#)).

The field exploring eukaryogenesis has moved on to studying the nature and origin of the first eukaryotic common ancestor (FECA). This subtle change in terminology has far-reaching consequences. The term FECA only puts a label on the first lineage that we would no longer define as prokaryotic, but which had not yet evolved all traits characterizing the LECA. But at what point did prokaryotic evolution transition into eukaryotic origin? Was it upon the emergence of meiosis and sex? Or the ER and its specialized compartment the nucleus? Or after the transition from archaeal to eukaryotic (bacterial-type) membrane lipids? The transition between prokaryotes and eukaryotes was fluid in nature, with the emergence of the new traits occurring in a currently unknown order ([Gould et al., 2016](#); [López-García and Moreira, 2020](#); [Vosseberg et al., 2021](#)). The critical question is: what drove the emergence of eukaryotic traits and what fixed them in evolution?

Here, we discuss the scenarios of a morphologically simple FECA versus a complex one on the basis of reviewing models and data that emerged after the report of the asgard archaeal superphylum from which the eukaryotic host lineage stems. We explore key eukaryotic traits and the phylogenetic distribution of protein associated families, in light of housing an endosymbiont that differs by all other traits in that it represents a semi-autonomous living entity. This imposed unique challenges onto the

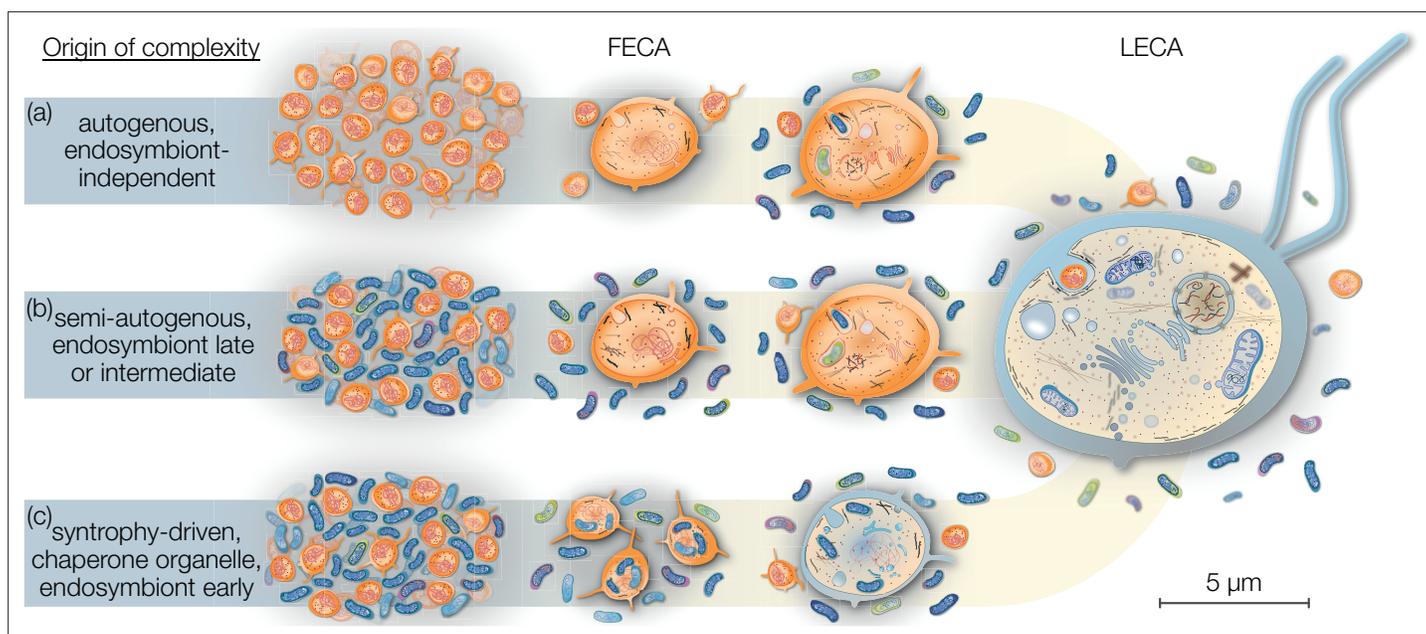


Figure 1. Eukaryogenesis scenarios with respect to the origin and nature of the first eukaryotic common ancestor (FECA). Scenarios summarized in **(a)** were based on the assumption that eukaryotic groups existed that lack mitochondria and which originated independently of endosymbiosis. Such models are no longer supported and we understand that the last eukaryotic common ancestor (LECA) was a mitochondrion-bearing lineage. Models in **(b)** speculate on a (semi-) autogenous origin of complexity. By their definition, the FECA was free of vertically transmittable endosymbionts. Genomic and cellular complexity of FECA grew through horizontally donated genes from diverse prokaryotes part of its habitat; to a degree FECA was eukaryote-like prior to endosymbiosis. Models falling under **(c)** differ from **(b)** in that both partners at the onset of eukaryogenesis were prokaryotic in nature and that complexity only began to evolve after endosymbiosis of the mitochondrial ancestor. By this definition, the FECA is characterized by the endosymbiotic event and describes the lineage that transitioned from prokaryotic to eukaryotic evolution. Note that the prokaryotic cells on the left and the LECA on the right are shown to scale on the basis of average cell diameters.

host throughout eukaryogenesis and whose solution, we argue, is witnessed in the form of compartmentalization, meiosis, and sex.

FECA and theories of eukaryogenesis in light of the asgard archaea

'It can be considered a relatively harmless habit, like eating peanuts, unless it assumes the form of an obsession; then it becomes a vice' (Roger Y Stanier).

The relatively harmless habit of tracing the origins of the eukaryotic cell has occupied scientists across several generations, a historical account of which is beyond the scope of this review but has been summarized elsewhere (Martin, 2017). Current models of eukaryogenesis differ above all in the relative placement and contribution of the endosymbiont and consequently the cellular complexity of the host archaeon prior to endosymbiosis. Briefly, mitochondria-early theories place endosymbiosis closer to or at FECA (Figure 1), suggesting that the traits that characterize LECA evolved after endosymbiosis from a prokaryotic-like host cell. On the contrary, mitochondria-late scenarios view endosymbiosis and mitochondrial origin as a finishing touch to the LECA (Figure 1). Intermediate models are gaining popularity, but are often vague on which traits evolved prior or after endosymbiosis.

The notion that eukaryote-like complexity was a prerequisite to phagocytosis for promoting mitochondrial origin appears mandatory to some, but the idea remains unsubstantiated (Leão et al., 2018; Martin et al., 2017b; Mills, 2020; Shiratori et al., 2019). Mitochondria-lacking but phagocytosing LECA models – such as the archezoa hypothesis (Cavalier-Smith, 1987) – lost support due to the now known universal presence of mitochondria across the diversity of all eukaryotic super groups (Hjort et al., 2010; Stairs et al., 2015), but variations of the archezoa hypothesis populate the literature, rekindled on the basis of inferred proteomes from asgard archaea.

Reports of a patchy distribution of homologues of the eukaryotic ESCRT-III, a ubiquitin modifying system, and eukaryote-like actin in the TACK superphylum, triggered thoughts about phagocytosing archaea (Guy and Ettema, 2011; Yutin et al., 2009). The identification of proteins with homology to ESCRT I and II, longin domains, sec23 and sec24 (Zaremba-Niedzwiedzka et al., 2017), Rab-like GTPase (Akil and Robinson, 2018; Klinger et al., 2016; Surkont and Pereira-Leal, 2016), or profilin that can inhibit (in vitro) rabbit actin polymerization (Akil and Robinson, 2018; Survery et al., 2021) quickly channelled into speculations that asgard archaea might have a dynamic cytoskeleton, intracellular membrane trafficking, and are morphologically complex (Klinger et al., 2016; Neveu et al., 2020; Zachar et al., 2018; Zaremba-Niedzwiedzka et al., 2017). These interpretations mirror a FECA that is reminiscent of the host lineage at the centre of the archezoa hypothesis (discussed in Martin et al., 2017b), culminating in the depiction of a mitochondrion-lacking eukaryote on the cover of Nature (Pittis and Gabaldón, 2016). When transmission electron microscopy revealed images of asgard archaea, that of *Prometheoarchaeon syntrophicum*, they uncovered tiny prokaryotes with no intracellular eukaryotic traits living in syntrophy with bacteria (Imachi et al., 2020). Such images contradict the narrative of complex asgard archaea, but resonate well with early warnings regarding overinterpretations of metagenome data (Dey et al., 2016).

Eukaryogenesis models rapidly adapted to the discovery of asgard archaea. They now focus on FECA with various speculations regarding the roles of the discovered genes in host biology prior to endosymbiosis. While the level of cellular complexity is not always explicitly declared, several cases can be made out that depict FECA without an endosymbiont (Baum and Baum, 2020; Dacks et al., 2016; Eme et al., 2017; Pittis and Gabaldón, 2016; Vosseberg et al., 2021). Some models can be interpreted one way or the other (Imachi et al., 2020), while some explicitly state that the host cell was a bona fide prokaryote and that eukaryotic traits and biology evolved after endosymbiosis (Gould et al., 2016; López-García and Moreira, 2020; Wu et al., 2022). Notably, the differences among these models rest upon a few dozen genes from the pan-asgard archaeal genome repertoire, whose overall unique contribution to the roots of eukaryotes was 0.3% or less (Knopp et al., 2021; Liu et al., 2021). Sources and timing of gene acquisition in the FECA to LECA transition are equally essential to correctly quantify as they remain hard to predict.

Entangled branches connecting kingdoms

Among eukaryotic genomes there are more genes of bacterial than of archaeal origin (Alvarez-Ponce et al., 2013; Brueckner and Martin, 2020; Makarova et al., 2005). An autogenous origin of cellular complexity on the basis of an archaeal (host) source alone would predict the opposite but prokaryotes

are characterized by mosaic genomes due to horizontal gene transfer (HGT) whose contribution to cellular complexity prior to endosymbiosis is debated (*Martin et al., 2017b; Pittis and Gabaldón, 2016*). Claims concerning differential loss of genes in extant archaea (*Koonin and Yutin, 2014; Eme et al., 2017*) are at odds with pangenomes that support a pan-asgard concept (*Knopp et al., 2021; Liu et al., 2021*). Dynamic genomes and the time passed since eukaryote origin challenge phylogenomic approaches and can skew interpretations including the timing of compartment origin. The estimated timing of gene duplications that depend on molecular clock techniques that are error-prone (*Graur and Martin, 2004; Rodríguez-Trelles et al., 2001; Tiley et al., 2020*) with respect to the origins of cellular complexity are also debated (*Tria et al., 2021; Vosseberg et al., 2021*). A reliance purely on relative branch lengths concluded that mitochondrial metabolism and the ER in eukaryogenesis ensued the origin of the nucleus (*Pittis and Gabaldón, 2016*). The method used has been questioned (*Martin et al., 2017a*), and the use of unspecific COG (cluster of orthologous genes) annotations in the study is problematic. The few universal proteins listed might operate in the present-day nucleus, but provide little to no evidence for the presence of one prior to endosymbiosis. Proteins of the nuclear pore complex, of which there are about three dozen (*Raices and D'Angelo, 2012*), were not identified or discussed, nor was the fact that the nucleus is a specialized compartment of the ER from which it forms during cell division (*Anderson and Hetzer, 2008*). Substitution rates that challenge molecular clock studies vary substantially across species (*Baer et al., 2007; Halligan and Keightley, 2009*) and the functional unit a protein is associated with (*Hartwell et al., 1999*). Considering that thousands of new protein families emerged at eukaryote origin that fall into such categories (*Preisner et al., 2018*) further highlights the caution with which we need to digest molecular clock studies on eukaryogenesis.

The distribution of protein families associated with eukaryotic traits across the domains of life is always telling. Seventy percent or more protein families involved in major eukaryotic traits (such as cell cycle, meiosis, autophagy, nucleus) are specific to eukaryotes, 10–15% (e.g. kinases) are universal across all domains of life, 10–15% are bacterial (e.g. aminopeptidases, mTOR interacting proteins, glycosyltransferases), and a small fraction appear to originate from archaea (DNA licensing proteins of cell cycle, ARG GTPases, N-glycan biosynthesis). The distribution of protein families across prokaryotes and eukaryotes (*Figure 2*) confirms that eukaryotes acquired genes from bacterial or archaeal sources and co-opted them to suit new eukaryotic traits evolving in the FECA to LECA transition, but the majority of protein families involved in eukaryotic cellular complexity are absent across the entire realm of prokaryotic diversity (*Brunk and Martin, 2019; Dell et al., 2010; Knopp et al., 2021; Liu et al., 2021; Lombard, 2016*). Hence, HGT falls short at explaining the pro- to eukaryote transition with respect to the origin of thousands of eukaryote-unique gene families and a reason for their positive selection in the absence of an endosymbiont. Beyond question, HGT fed into eukaryogenesis – after all, the eukaryotic cell is the product of two prokaryotes – but endosymbiotic partners bring along thousands of genes and many were integrated into the host genome (*Timmis et al., 2004*). They can explain the pronounced non-alphaproteobacterial signal among proteins supporting eukaryotic traits, especially if we place mitochondrial origin at the root of the FECA and accept HGT to be prevalent.

Phylogenetic trees built using concatenated gene sequences boost phylogenetic signals, but under the premise that the individual genes used recapitulate the evolutionary history of the species (*Robinson and Foulds, 1981*). For incomplete and contaminated metagenomes (including early releases of asgard archaeal ones), the individual ribosomal gene trees were incongruent (*Garg et al., 2021*). Similar to simulated chimeric genomes containing genes from different species, metagenome assembled genomes are prone to assembly and binning artefacts. The frequent use of automated pipelines and poorly fitting phylogenetic models exacerbates the risk of drawing false conclusions from metagenome data (*Williams and Embley, 2014*). For instance, the presence of glycerol-3-phosphate lipids in asgard archaea was claimed (with far-reaching implications on the lipid transition during eukaryogenesis) based on the predicted presence of enzymes involved in the synthesis of ester-linked fatty acids (*Villanueva et al., 2017*). No evidence of such lipids, however, was found in the biochemical analysis of a cultured asgard archaeon (*Imachi et al., 2020*) and the presence of the set of required enzymes in asgard archaea has yet to be identified. Better assembly methods result in more complete circular genomes from both axenic culture and metagenomic approaches that mitigate issues of tree congruence (*Garg et al., 2021*), albeit leaving the same room for interpretations.

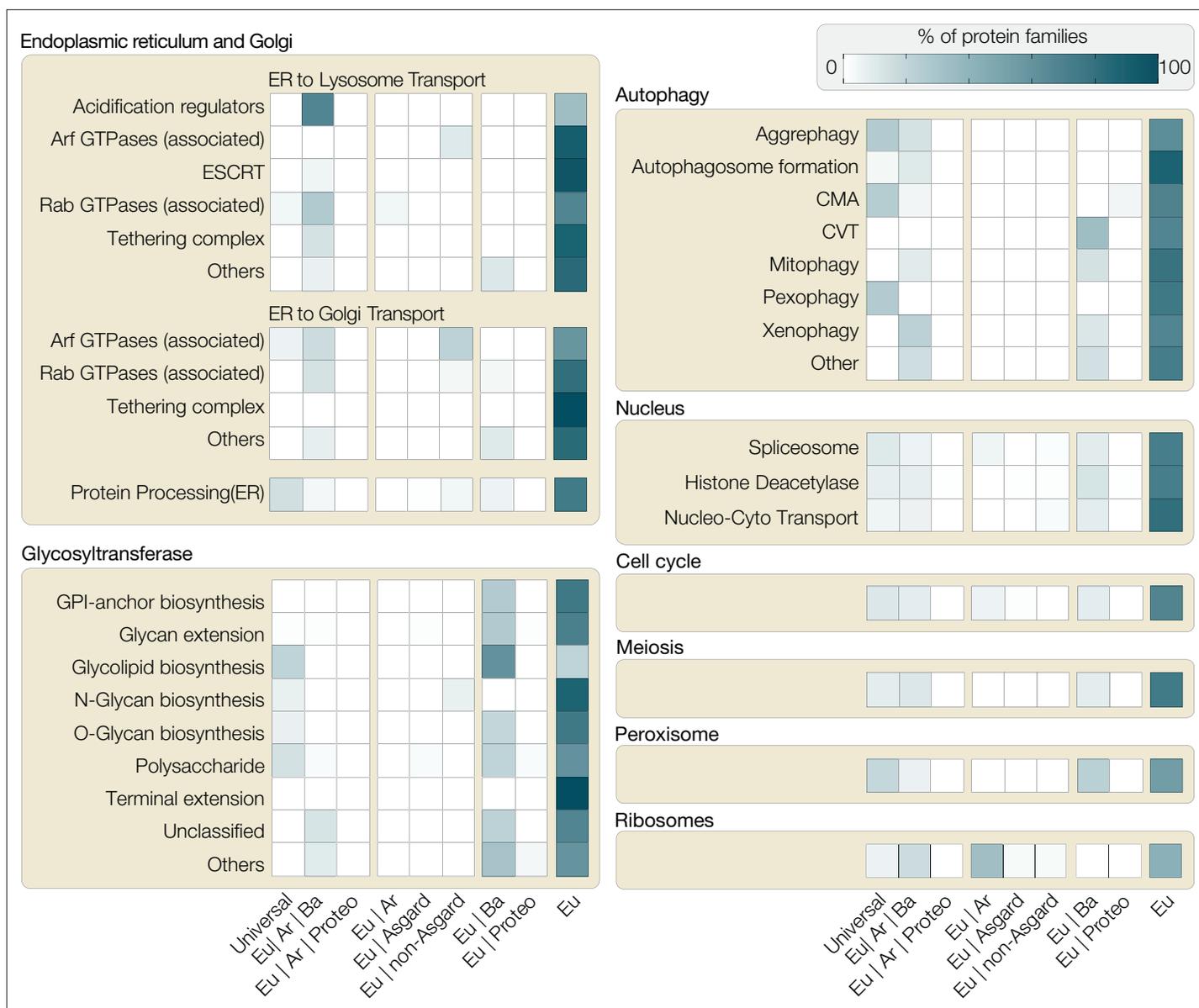


Figure 2. Distribution of protein families involved in eukaryotic traits. All eukaryotic protein sequences associated with different functions in KEGG (Kyoto Encyclopedia of Genes and Genomes, as of November 2021) were used to build representative hidden Markov models (HMMER/3.1), which were then searched against all prokaryotic genomes available on KEGG. Based on hits found across bacterial phyla, the following categories were assigned: Universal (shared with >50% bacterial phyla and archaea), Eu|Ar|Ba (shared with archaea and <50% bacterial phyla), Eu|Ar|Pr (shared with archaea and proteobacteria only), Eu|Ar (shared with Ar), Eu|Asgard (shared only with asgard archaea), Eu|non-Asgard (shared only with non-asgard archaea), Eu|Ba (shared with bacteria), Eu|Proteo (shared only with proteobacteria), Eu (not shared with any prokaryotes). Each box shows the percentage of protein families across these categories (x-axis) for the pathways analysed. For example, a substantial percentage of ribosomal proteins (bottom right) are shared between Eukaryote-Archaea or Eukaryote-Archaea-Bacteria, highlighting the host's ribosomal contribution to eukaryotes. Proteins for other categories, however, are either predominantly eukaryote-specific or shared between eukaryotes and bacteria (e.g. glycotransferases or some proteins of autophagy). CMA, chaperone-mediated autophagy; CVT, cytoplasm to vacuolar targeting pathway.

Underpinning studies of evolutionary history are phylogenetic trees and theories behind constructing and interpreting them. While it is well beyond the scope of this manuscript to discuss all the vagaries of the field of cladistics and modern phylogenies, it is increasingly evident that many phylogenetic studies have moved from a field that requires expertise in biology to a field that requires expertise in computation (*Fitzhugh, 2016*) – hypotheses generated from DNA sequences run the risk of taking precedence over morphological evidence (*Wheeler et al., 2013*). This is less of a critique

than a realization. Although sequencing and computational techniques have made significant progress over the years, for the timescales dealt with in early evolution, most issues and challenges remain. It is critical to remember that phylogenetic trees are hypotheses on the evolutionary relationship between organisms and not an observation on itself (*Hennig, 2000*). No phylogenetic tree is perfect, few are for eternity, and no tree alone will ever satisfy the need for empirical evidence.

Eukaryotic traits in light of accommodating a prokaryotic endosymbiont

'In the case of living machinery, the 'designer' is unconscious natural selection, the blind watchmaker' (Richard Dawkins).

Evolution is typically understood to progress gradually and randomly through mutations and the selection of beneficial traits vertically across generations (*Darwin and Murray, 1859; Futuyma, 1986*). Endosymbiosis adds a massive horizontal component to evolution that is, however, still subject to the basic selection and fixation process. In other words, while the emergence of eukaryotic traits was gradual, the selective pressure that demanded their emergence was more radical. It is this duality that stands between eukaryogenesis theories like a firewall. Any hypothesis that pictures an archaeal lineage transitioning from prokaryotic to eukaryotic cell biology – even of an intermediate type – in the absence of an endosymbiont needs to explain why it was a singularity. Microbial syntrophy is the norm and so is the selective pressure to optimize it. Why are intermediate cell types not observed among the many syntrophic prokaryotes studied, if it was not for the lack of an endosymbiotic event? A mitochondrion-lacking but complex FECA explains eukaryotic traits solely from a host perspective and misses to provide a plausible reason for selection and the emergence and fixation of traits we here discuss in more detail (*Figure 3*).

The ER and Golgi apparatus

Glycosyltransferases are promiscuous enzymes and it has been suggested they are separated through ER-Golgi compartmentalization for that reason (*Biswas and Thattai, 2020*). N- and O-glycosylation are ubiquitous in eukaryotic cells, but not so in prokaryotes. Eukaryotic N-glycosylation is likely derived from the archaeal ancestor, while O-glycosylation is more prevalent among bacteria (*Abu-Qarn et al., 2008; Dell et al., 2010; Jarrell et al., 2014*). Hence, if each pathway stems from one of the prokaryotic partners, natural selection would foster a spatial separation only upon and not prior to endosymbiosis. The ER lumen and mitochondrial intramembrane space (the former bacterial periplasm) share notable homologies. This includes calcium storage (*Dominguez, 2004; Raffaello et al., 2016*), disulfide relay systems (*Backes et al., 2019*), and redox balance (*Cardenas-Rodriguez and Tokatlidis, 2017*). The contact sites of the ER and mitochondrion are cornerstones for the synthesis and regulation of lipids and a plethora of cellular roles (*Booth et al., 2016; Flis and Daum, 2013; Friedman et al., 2011; Hamasaki et al., 2013*). This could be a consequence of the ER stemming from mitochondrial-derived vesicles (MDVs) (*Gould et al., 2016*). MDVs could have provided the necessary endomembrane material for compartmentalization and remain the most plausible source for the lipid transition from ether-linked, archaeal head groups to ester-linked (bacterial) eukaryotic head groups. Much on the origin of the endomembrane system remains a speculation, but not so the existence of MDVs, their role in eukaryotic biology, and how they induce compartment formation (*Schuler et al., 2021; Sugiura et al., 2017; Sugiura et al., 2014; Yamashita et al., 2016*). Eukaryogenesis models failing to acknowledge their existence miss a biological fact with significant explanatory power.

Vesicle trafficking

Vesicle secretion from the plasma membrane into the environment is a common trait of all cells. Unique to eukaryotes are the many ways with which they can internalize membrane vesicles of various sizes, ranging from clathrin-mediated endocytosis (~100 nm) to phagocytosis (>750 nm), using different molecular machineries. Intracellular vesicle trafficking connects the plasma membrane with the endomembrane system and the compartments thereof among each other. All compartments that define the endomembrane system – with the ER at its core – as well as the majority of regulatory and structural proteins are conserved across eukaryotes and absent in prokaryotes (*Klinger et al., 2016; Kontou et al., 2022*).

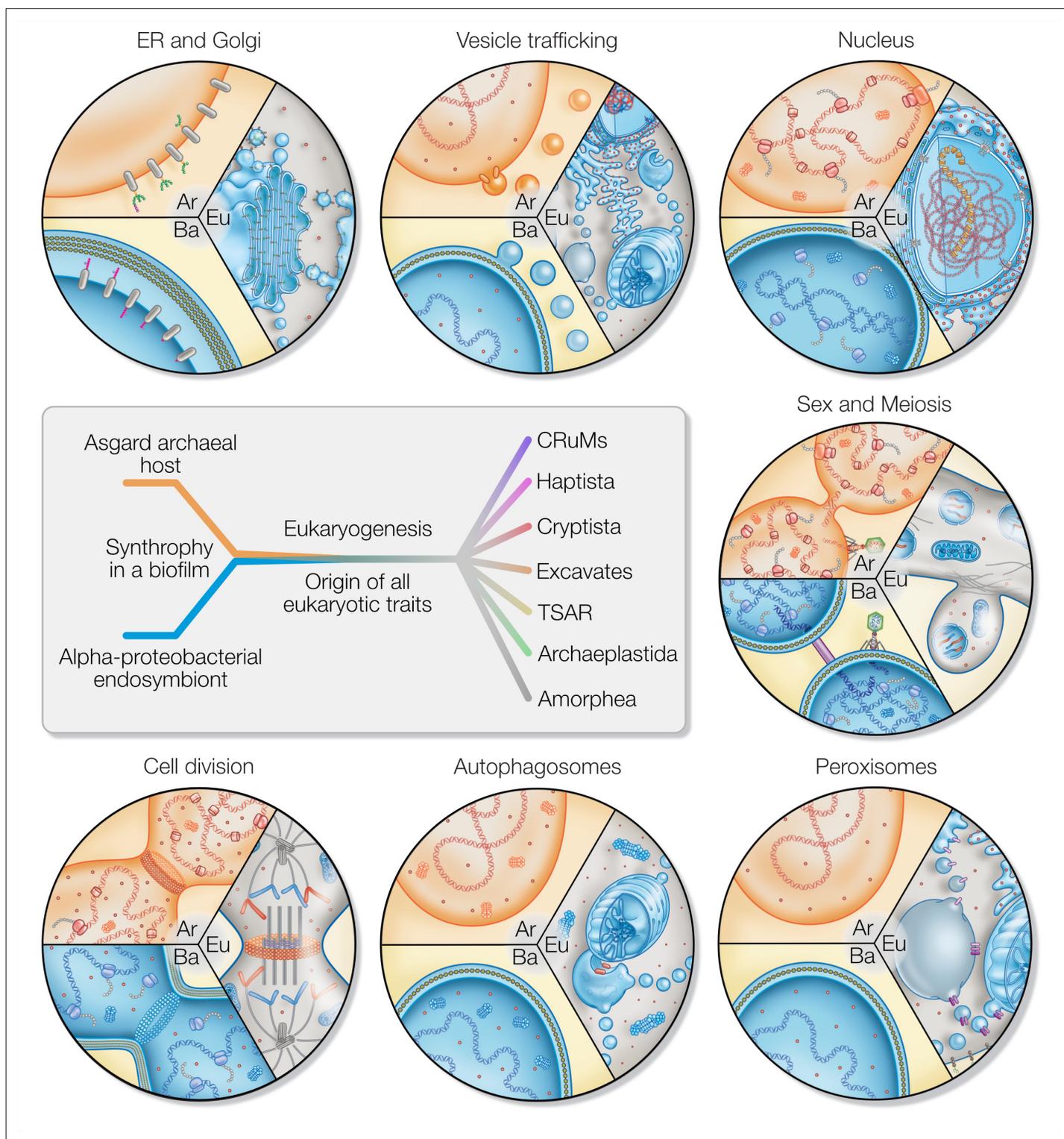


Figure 3. Eukaryotic traits in light of housing an endosymbiont. Each segment highlights an eukaryotic (Eu) trait and the comparable, if present, situation in bacteria (Ba) and archaea (Ar). (i) Prokaryotes secrete outer membrane vesicles (OMVs) and an endosymbiont (mitochondrion) secreting OMVs could have (ii) given rise to a dynamic endomembrane system within the archaeal host and explaining the transition from archaeal to bacterial lipids. *N*-glycosylation has been identified in all domains of life, but the eukaryotic *N*-glycosylation pathway is homologous to that of archaea. (iii) A specialized extension of the endoplasmic reticulum (ER), the nucleus prevents co-transcriptional translation of proteins – as is the rule in prokaryotes – to allow for the splicing of introns. (iv) Prokaryotes constantly shed and acquire DNA from the environment, and often promiscuously by transformation,

Figure 3 continued on next page

Figure 3 continued

transduction, and conjugation. In the absence of such dedicated mechanisms, eukaryotes avoid Muller's ratchet through sex and meiosis whose origin might be linked to coordinating the merging of two genomes and synchronizing nuclear and mitochondrial division. (v) Peroxisomes also form through mitochondria-derived vesicles and house enzymes of alphaproteobacterial origin. (vi) Eukaryotes perform autophagy using membranes and proteins of the ESCRT machinery to surround and digest internal membrane compartments including the mitochondrion. (vii) While bacteria use homologs of tubulin to perform fission, eukaryotic fission utilizes actin and components of the ESCRT machinery similar to archaea, whereas the tubulin in eukaryotes is for instance used to separate chromatin and intracellular compartments.

The nucleus

The nucleus is a distinctive extension of the ER and forms from the latter after mitosis (**Anderson and Hetzer, 2008**) using homologs of ESCRT complex (**Olmos et al., 2015**). It separates transcription from translation and is the site of pre-ribosome assembly (**Peña et al., 2017**). As with any trait, a selective reason for its presence must outweigh the costs for its maintenance; consider, for example, the exchange of mRNA and effectors with the cytoplasm (**Nerurkar et al., 2015; Warner, 1999**). A plausible selection could have been imposed by the transfer of group II introns from the endosymbiont that drove the origin of eukaryotic splicing and need for separating transcription slowed by the spliceosome from fast translation (**Martin and Koonin, 2006**). Mitochondria-early scenarios provide both the problem (group II introns that need to be spliced) and the solution (MDVs that might have given rise to the ER) (**Gould et al., 2016**).

Sex and meiosis

The prokaryotic solution to prevent mutational overload through Muller's ratchet is HGT (**Muller, 1964**). The nucleus renders the eukaryotic cytoplasm almost sterile of DNA (preventing HGT), wherein it plays a regulatory immune function (**Abe et al., 2019; Paludan and Bowie, 2013**). The eukaryotic solution was ploidy, multinucleated cells and reciprocal recombination through meiosis (**Garg and Martin, 2016**). A multinucleated state is readily achieved by decoupling nuclear from cell division, a mechanism commonly observed in prokaryotes wherein the DNA replicates independently of the cell before portioning into daughter cells (**Haeusser and Levin, 2008**). The syncytial theory for eukaryotic origin (**Garg and Martin, 2016**) posits that by virtue of multinucleated cells within a singular archaeal host, multiplying bacterial symbionts are free to lose genes via endosymbiotic gene transfer to the multiple copies of the host nucleus/nuclear material in the cytoplasm to explore various configurations under the constant onslaught of group II introns, yet retaining fitness by compensating viable mRNA in-trans within the same shared cytoplasm. The explanatory power of this model is twofold: (i) it explains how homologous recombination – which subsequently evolved to meiosis as we understand it today – was necessary to maintain viable copies of undisrupted genes, while simultaneously maintaining the presence of bacterial transferred genes, and (ii), it explains the monophyly of eukaryotes. As long as the FECA to LECA transition continued, the multitude of host nuclei remained within a single confined cytoplasm until the fittest version was optimized via various rounds of endo-meiosis and homologous recombination. Any origin of cell division and/or cell cycle might have given rise to gamete-like spores that separated off the original syncytium. In cases where a successful combination was released through ESCRT-driven scission (see below), a similar process applies for further optimization. In scenarios in which the budded off cell (gamete) was fitter than the syncytium, it would outcompete the original syncytium or alternatively would be outcompeted when it contained aberrant genomes. In either case, the singularity of LECA is well explained by the syncytial model of the FECA to LECA transition (**Garg and Martin, 2016; Skejo et al., 2021**).

Meiosis in itself is ancient, ubiquitous, and the central process that imparts an advantage to sex in eukaryotes (**Colnaghi et al., 2022; Colnaghi et al., 2020; Malik et al., 2008; Speijer et al., 2015**). Several theories place mito-nuclear interactions, heteroplasmy, and mitochondrial ROS as drivers of eukaryotic sex (**Colnaghi et al., 2020; Hörandl and Speijer, 2018; Radzvilavicius and Blackstone, 2015**). HGT alone was insufficient for LECA to escape Muller's ratchet in the absence of homologous recombination (**Colnaghi et al., 2022**), when considering expanding genome size and repeat sequence frequency. Everything points to an origin of sex and meiosis necessitated by the presence of mitochondria. Moreover, sex, as a trait, restricts the number of potential mating partners (by 1/number of sex types), and it is hence less surprising that it did not evolve in groups outside of eukaryotes, but had to in the FECA.

Peroxisomes

The majority of enzymes of peroxisomal beta-oxidation are of alphaproteobacterial origin (**Boite et al., 2015**) and peroxisomes might have evolved to compartmentalize ROS-producing beta-oxidation and protect the mitochondrial genome (**Speijer, 2017**). De novo biosynthesis of peroxisomes involves MDVs with integrated Pex3 and Pex14 that fuse with ER-derived vesicles containing Pex16 (**Sugiura et al., 2017**), and the compartment for beta-oxidation appears absent in species lacking respiring mitochondria (**Le et al., 2020**). Peroxisomes not only make sense in the presence of a mitochondrion, they are also partly a product thereof (**Mohanty and McBride, 2013; Sugiura et al., 2017**).

Autophagy

Cytosolic protein homeostasis in prokaryotes is performed by proteases and proteasomes, which are common to both archaea and bacteria. Defective membrane proteins and membranes are shed by mechanisms similar to bacterial outer membrane vesicle secretion (**Schwechheimer and Kuehn, 2015**). Eukaryotes utilize membrane-bound compartments in the form of autophagosomes, also for recycling membranes including their proteins (**Nakatogawa, 2020**). Mitophagy removes damaged mitochondria and is initiated by ER-mitochondrial contact sites (**Hamasaki et al., 2013**). It is a trait needed in the presence of large intracellular compartments and the occasional yet inevitable breakdown of organelles that require immediate containment (**Anding and Baehrecke, 2017**).

Cell division

The eukaryotic cell cycle is a series of choreographed steps that leads to the correct portioning of genetic material, endomembrane, and organelles to both daughter cells (**Harashima et al., 2013**). The presence of a nuclear compartment and an endosymbiont are incompatible with binary fission in the absence of orchestrated replication and organelle and compartment division, and cytokinesis. As mentioned in the previous section, in prokaryotes the nuclear material (genome) replicates independently of cell division, which would have facilitated the formation of syncytial populations (**Haeusser and Levin, 2008**). During this time, however, we speculate that mitochondrial metabolism started playing a more significant role in controlling cell division, given the role of nutrient availability in coordination of cell division. The G1 phase of the eukaryotic cell cycle results in the mitochondria as the master regulator for S/G2 progression (**Antico Arciuch et al., 2012; Mitra et al., 2009**), suggestive of a deep link between mitochondria and the cell cycle and one that would have been difficult to integrate into a pre-existing one. Eukaryotic cell division employs the use of ESCRT homologs and actin in contrast to bacterial division mechanisms involving FtsZs from which also tubulin evolved (**Christ et al., 2016; Goliand et al., 2014; Stoten and Carlton, 2018**). This suggests the evolution of an independent pathway for cell division involving ESCRT proteins consistent with their role in archaeal cell division (**Tarrason Risa et al., 2020**), one that was based on outer membrane vesicle secretion, but this time packaging mitochondria and the nucleus/genetic material, the latter similar to a role of prokaryotic OMVs (**Schwechheimer and Kuehn, 2015**). Understanding how these pre-existing mechanisms were leveraged in an elaborate checkpoint system of the eukaryotic cell cycle remains to be elucidated.

One might consider the cytoskeleton another eukaryotic trait, but this is more involved. The eukaryotic cytoskeleton rests on three main pillars: (i) actin and associated proteins, (ii) tubulin and associated proteins, and (iii) the utterly diverse intermediate filament (IF) proteins. Components of each pillar, sometimes also in combination, can be found in archaea and bacteria alike (**Duggin et al., 2015; Larsen et al., 2007; Preisner et al., 2018; van den Ent et al., 2001; Wickstead and Gull, 2011; Zaremba-Niedzwiedzka et al., 2017**). As with many things in eukaryogenesis, it is the intricate combination and universal presence of all three cytoskeletal pillars and the dynamic nature which they are used in eukaryotes that is characteristic. The latter is best demonstrated by the rapid switch in motility between actin-based gliding and tubulin-based flagella-driven swimming in many protists, likely also a feature of the LECA (**Fritz-Laylin et al., 2010; Kusdian et al., 2013; Preisner et al., 2018**). Basic components were derived from the host cell, such as gelsolin-regulated actin filaments (**Akil et al., 2020**) and evolution co-opted such mechanisms en route to LECA. It is conceivable that with expanding cell size, increased intracellular complexity and the need of an orchestrated cell cycle, the selection for a dynamic but simultaneously in parts rigid cytoskeleton increased, which triggered

the expansion of the IF protein family required for mechanical support, and the origin of additional accessory proteins and regulatory mechanisms that are absent in prokaryotes.

The identification and subsequent characterizations of asgard archaea have done the following for eukaryogenesis: (i) They underpin the syntrophic origin of eukaryotes involving two prokaryotic partners and (ii) provide support for a 2D tree of life (i.e. two domains of life, bacteria and archaea, evolved from the origin of life and eukaryotes emerged from within archaea after endosymbiosis of an alphaproteobacterial partner). (iii) They provide no evidence for the presence of bacterial-type ester-linked lipids in asgard archaea, (iv) reject a complex archaeal ancestor necessary to explain the patchy distribution of eukaryogenesis-relevant gene families (Wu et al., 2022), and (v) show that the asgard archaeal set of genes before unique to eukaryotes closes the gap to the number of gene families encoded by eukaryotes by only 0.3% (Knopp et al., 2021) or less (Liu et al., 2021). Hence, with respect to explaining the origin of eukaryotic traits and a rationale for their universal presence in eukaryotes, the asgard archaea and their syntrophic bacterial partners support and place us at scenarios that were submitted some 25 years ago (Martin and Müller, 1998; Moreira, 1998; Vellai and Vida, 1999).

Considering syntrophy as a key ecological parameter in eukaryogenesis was an early notion that has stood the test of time (Imachi et al., 2020; López-García and Moreira, 2020; López-García and Moreira, 2019; Sousa et al., 2016; Spang et al., 2019; Wu et al., 2022). Ever since, observations from the field of microbial ecology, genomics, and geology continue to encourage us to picture eukaryogenesis to have occurred within a microbial mat, where multiple species thrive in close proximity and ample syntrophies exchanging substrates such as H₂/electrons under limited or no oxygen (López-García and Moreira, 2020). Recent advances in geochemistry added new support to the proposal that eukaryogenesis occurred in anoxic niches with a preferred shift towards aerobic metabolism being a secondarily derived state (Mills et al., 2022) and so do the culturing conditions of Promethoarchaeon (Imachi et al., 2020). Such prokaryotic consortia can source genes through HGT from pangenomes of other bacteria and (asgard) archaea and the virosphere that also contributed to the birth of the eukaryotic genome (Spang et al., 2022; Wu et al., 2022). Evidently, however, most eukaryotic protein families evolved during the FECA to LECA transition and selective pressures due to endosymbiosis were likely key.

To conclude, physiological and phylogenomic studies support a mitochondria-early scenario and so does cell biology (Figure 3). Claiming mitochondria were of little importance in eukaryogenesis contradicts the simultaneous claim that intermediates lacking mitochondria all went extinct – an oxymoron that suggests that the mitochondrial endosymbiont contributed little in the FECA to LECA transition, while its presence was vital for the survival of FECA during eukaryogenesis.

On being the right size in eukaryogenesis

'The most obvious differences between different animals are differences of size, but for some reason the zoologists have paid singularly little attention to them' (John BS Haldane).

Haldane began his influential essay by addressing a lack of scale bars in zoology books. One can point to a similar issue regarding eukaryogenesis, in which models often depict cells not changing in size up to scale in the course of the FECA to LECA transition (Baum and Baum, 2020; Gould et al., 2016; Spang et al., 2019). Though not intentional, this is important. In eukaryogenesis we are dealing with at least a 10 times increase in cell diameter, with known consequences regarding cell volume, morphology, and molecular diffusion limits among other factors (Young, 2006). Engulfing

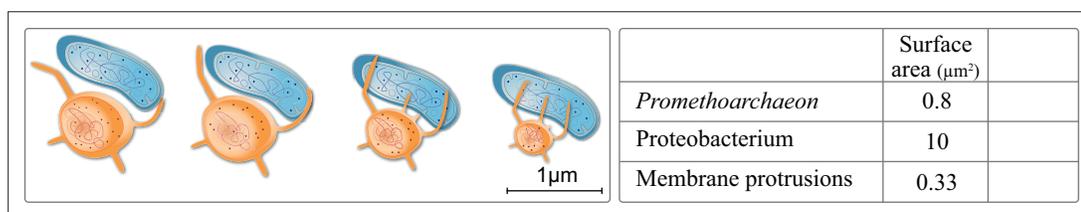


Figure 4. Extracellular membrane protrusions from asgard archaea. Potential trade-off between cytosol investment between cell and protrusion. Each 1 μm increase in protrusion takes 12% cytoplasm, likely limiting total protrusion and reducing cell volume. Values in the table for the asgard archaea were calculated based on Eme et al., 2017: average asgard cell radius: 0.25 μm, average protrusion radius 50 nm, and length 1 μm. Average alphaproteobacterium radius and length were taken to be 0.5 and 3 μm, respectively.

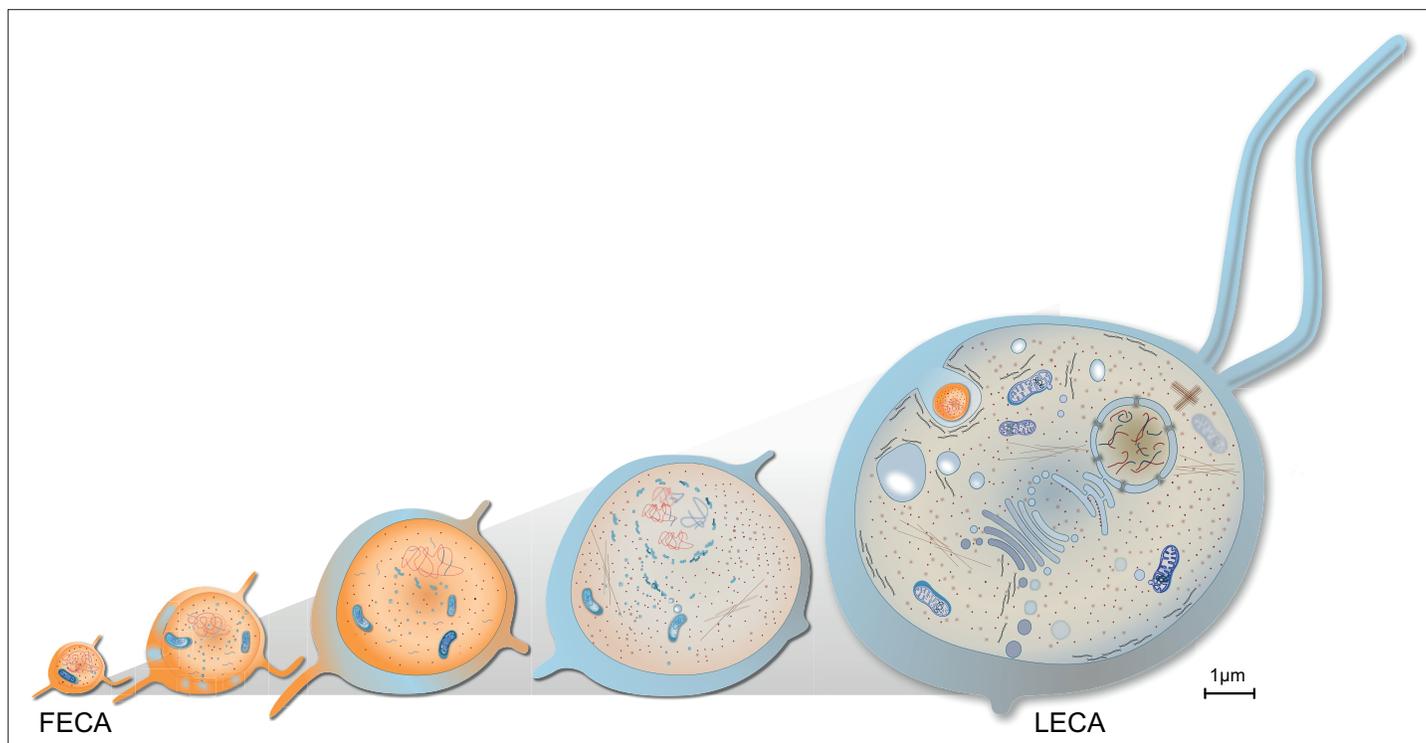


Figure 5. Evolution of last eukaryotic common ancestor (LECA) after the symbiotic event. Vesicles from the proteobacterium accumulated in cytosol giving rise to permeable barrier around the host DNA [Brueckner and Martin, 2020](#), selected positively due to parallel bombardment of genes ([Land et al., 2015](#)). As a result, the genome of bacterium shrunk, whereas that of the host expanded, along with increase in cell size. New combinations of genes, powered by energy from the endosymbiont, led to emergence of novel trait and LECA with eukaryotic cell biology.

a proteobacterium with a surface area of $10 \mu\text{m}^2$ requires 10 times the surface area of the asgard archaeon *Prometheoarchaeon*. For a typical protist it is only 1% of its surface area. Putting scales on a recent model, the entangle-engulf-endogenize mode ([Imachi et al., 2020](#)) brings forth details worthy considering: an observed tubular protrusion requires 12% of cytoplasm for a 50% increase in surface area interacting with syntrophic partners ([Figure 4](#)). Four to six protrusions approximately result in the doubling of cytosolic volume, maybe explaining why *Prometheoarchaeon* has not been observed to produce more than six protrusion per cell.

Such protrusions might have been relevant for the uptake of the symbiont, but the surface area of such a protrusion is 3% that of a proteobacterium. Entangling a proteobacterium entirely would take at minimum 50 protrusions, the cost of which is six times the cytoplasmic volume and not considering the multitude of proteins needed for recognition, surface attachment, and the processes thereof. We also note that we know neither of a case in nature where tubular-like extensions (allowing nutrient exchange) fuse to sheets (allowing cell engulfment), nor can we imagine how this would work on a molecular level, respecting membrane biology, and in 3D. Considering scale bars or tubular versus sheet-like membrane biology is not intended to disprove any model, but it highlights potential issues and also reminds us of the question of when and how did the size increase in the FECA to LECA transition.

A feed forward loop supporting an increase in cell size

'The higher animals are not larger than the lower because they are more complicated. They are more complicated because they are larger' (John BS Haldane).

Haldane noted that an elephant has to be as complicated as an elephant, because it is as large as an elephant. Across eukaryotes, an increase in cell size (1000–10,000 times) and morphological complexity is common, matched by a comparable increase in genome size. The upper end of bacterial genomes is 15 Mb ([Land et al., 2015](#)), that of (haploid) eukaryote is around 130 Bb ([Pellicer et al., 2010](#)). What drove this increase in cell and genome size during eukaryogenesis ([Figure 5](#))?

In prokaryotes ATP production is limited by the available cell surface. This limits the replication rate and imposes negative selection on genome expansion (Lane, 2007). Conversely, in mitochondria powered eukaryotes, energetic efficiency increases with cell size (Hochachka et al., 2003; Lane, 2007), imposing a positive selection. Increased cell size in eukaryotes means increased DNA content to maintain the karyoplasmic ratio (Cavalier-Smith, 2005; Cavalier-Smith, 1985), which is positively selected for (Lane, 2007). Through this, and remembering that the endosymbiont provided the host cell with both problems and solutions (main text), one can speculate on mechanisms and a selective pressures for the emergence of eukaryotic cell biology, cell and genome size during the FECA to LECA trajectory.

Endosymbiosis provided an influx of endosymbiotic genes and membrane material. An early endomembrane system with minimal protein content, for example, proteins that are likely to be packed for secretion via OMVs, etc., emerged and the nucleus formed for reasons discussed in the main text. Constant fusion of endosymbiont-derived vesicles with the archaeal host provided a mechanism for the lipid shift and compartment origin (Gould et al., 2016), which might have fostered an increase in cell size. Integration of endosymbiotic DNA provided one early mechanism for why the genome size increased, together with duplication events (Kelly, 2021; Tria et al., 2021; Van de Peer et al., 2009). The ongoing concentration of ATP synthesis to mitochondria imposed positive selection on cell size, which allowed for a further increase in genome size. A feed-forward process of increased cell size stipulating increased genome size and vice versa commenced that was supported by an emerging endomembrane system and intracellular transport that counteracted the molecular diffusion limit (Figure 5). It provided ground for new combinations of genes to emerge and expressed (Lane and Martin, 2010) and increased cell size accommodated experimental expression of new proteins in the cytosol (Dill et al., 2011). This presented an opportunity – almost unlimited in theory – for the origin of new protein families and complex traits. Considering that most of these new inventions revolve around the endosymbiont (Figure 3), suggests it drove the selection for their emergence and fixation. Haldane might have put it this way: the LECA got larger because it was complex and it became complex because it was larger.

Did eukaryogenesis come with a price tag and if so, who paid?

'I have not failed. I've just found 10,000 ways that won't work' (Thomas Edison).

The cost of innovation is significantly higher than the manufacturing of the final product – the COVID vaccine serves as a topical example (Light and Lexchin, 2021). Eukaryotes have innovated several folds higher number of protein families than archaea as evident from genome (Brueckner and Martin, 2020) and proteome data (Müller et al., 2020) alike, underpinning the complexity across the eukaryotic tree of life. Since mutations lack foresight and are more likely to be deleterious than advantageous (Eyre-Walker and Keightley, 2007), inventing new proteins takes a considerable amount of trial and error. Ribosome production and protein biosynthesis consumes the majority of a cell's energy budget (Harold, 1986; Kafri et al., 2016) and the energy budget of trial and error would be orders of magnitude higher. Mitochondria were key by providing eukaryogenesis with an energetic freedom that supported this unparalleled level of innovative protein evolution and expression (Lane and Martin, 2010). While challenged (Lynch and Marinov, 2017; Schavemaker and Muñoz-Gómez, 2022), we do not see it disproven (Gerlitz et al., 2018; Lane, 2020). Calculations questioning the bioenergetic contribution of mitochondria do not account for the cost of evolving novel proteins.

The acquisition of a respiratory electron transport chain through excessive HGT does not make a cell complex (Nelson-Sathi et al., 2012), because the location of the bioenergetic membrane matters. The ratio of bioenergetic membrane (=energy generation) to genome size is high when harbouring an endosymbiont with internalized energetic membranes and a reduced genome (Lane and Martin, 2010). Also, a bioenergetic plasma membrane is incompatible with phagocytosis and the internalization of the bioenergetic membrane was a prerequisite to evolve phagocytosis (Martin et al., 2017b). A physiological observation that puts a timing on events in the FECA to LECA transition. Eukaryotes that maintain complexity in the absence of respiring mitochondria has prompted some to question the importance of mitochondria and a surplus of ATP at eukaryote origin (Hampl et al., 2019), while missing a critical detail: the examples listed stem from species that are either parasites or commensals of eukaryotes and who are energetically dependent on canonical mitochondria. The same holds true for the only eukaryotic taxon not possessing mitochondria, *Monocercomonoides*. They secondarily lost mitochondria and can only thrive in the gut of some animals (Karnkowska et al.,

2016). Such parasites or commensals have engaged an evolutionary path characterized by energetic dependency. Their complexity might diminish over evolutionary timescale, should they not go extinct with their hosts first. The issue is the origin of eukaryote complexity from prokaryotic ancestors, not the maintenance of eukaryotic complexity from eukaryotic ancestors.

An alternative to the energetics argument in explaining the ubiquity of mitochondria and its role in eukaryogenesis is missing, and the papers that question it are no exception. The internalization of energetic membrane – energy production from only 10% of cell volume – decoupled from the genome, as is the case in mitochondria (Fenchel, 2014), provides an optimum for protein innovation and a selection towards a complexity that can maintain a 200 ton blue whale.

A successful endosymbiosis and origin of a new domain: chance or necessity?

'Everything existing in the universe is the fruit of chance and of necessity' (Democritus).

Evolution is random and selects for fitness. Extinction is the rule and so is the common principle of use it or lose it (Lahti et al., 2009). Traits that remain unchanged across various organisms and through a billion years of evolution are indicative of the fact they are fundamental. Endosymbiosis is absent among prokaryotes (apart from isolated exceptions) and so is morphological complexity comparable to that of eukaryotes (reviewed in Martin et al., 2017b). Is it by chance or necessity? It is unlikely that the endosymbiosis leading into the origin of the eukaryotic domain was the first and only attempt throughout now 4 billion years of prokaryotic evolution.

The set of challenges posed by an endosymbiont are generic in nature: (i) there is a constant influx of endosymbiotic DNA which, after integration into the host genome (Allen, 2015; Portugez et al., 2018), is also exposed to an increased accumulation of deleterious mutations (Eyre-Walker and Keightley, 2007). The secretion of outer membrane vesicles by the endosymbiont is inevitable (Deatherage and Cookson, 2012), as well as the removal of irreversibly damaged organelles or the need to supply the endosymbiont with substrate from ions to peptides. Dividing endosymbionts need to be integrated into the division cycle of an archaeal host itself relying on simple binary fission.

A lot of solutions are associated with compartmentalization and this is a good time to remember that the mitochondrial endosymbiont not only provided the challenges, but maybe also the material to solve some (Gould et al., 2016). Any attempts at eukaryogenesis are prone to fail, if such challenges are not met by solutions that furthermore require correct timing (Barbrook et al., 2006). An influx of genes via HGT alone does not translate into complexity. Despite the metabolic transformation of haloarchaea via a chunk of some 1000 genes of bacterial origin (Nelson-Sathi et al., 2012) – maybe through a syntrophic partner and failed endosymbiosis – haloarchaea show no intracellular complexity. So while the encounter of the mitochondrial ancestor with an archaeal host occurred by chance, the emergence of a complex cell biology upon endosymbiosis was a necessity. Once a cell biology that can chaperone an endosymbiont is established, however, additional endosymbionts may follow without noticeable changes to the host.

The subsequent acquisition of the plastid added no extra compartments to the heterotrophic host that gave rise to the Archaeplastida, despite adding thousands of cyanobacterial genes to the host genome (Timmis et al., 2004). The same is true for an independent plastid acquisition by a rhizarian protist (Lhee et al., 2019) and likely many other endosymbiont-bearing protists (Husnik et al., 2021). Ever since eukaryogenesis, the cellular framework required for housing another prokaryote was in place. Some compartments have experienced physiological remodelling, such as the peroxisome (Islinger et al., 2010), but many components that evolved to service mitochondria during eukaryogenesis were recycled for the plastid: dynamins for fission (Fujimoto and Tsutsumi, 2014), redox balance through thioredoxins (Thormählen et al., 2017), and organelle digestion and recycling through the autophagosome (Ishida et al., 2014). One could add secondary endosymbioses, in which the acquisition of algae by eukaryotic hosts can lead to the stripping of all eukaryotic compartments of the endosymbionts (including their mitochondria), but that otherwise add no additional compartment or complexity to the host.

Conclusion

Morphologically complex life on Earth has a singular origin: eukaryogenesis. The LECA had evolved all canonical traits that we understand separates prokaryotic from eukaryotic life. Closing the gap

between a simple FECA and a complex LECA by presupposing a complex FECA opens an equally wide gap between a simple and a complex archaeal host. Picturing FECA without an endosymbiont offers little explanation for the existence or emergence of eukaryotic traits and the lack thereof in prokaryotes (including asgard archaea), apart from the inevitable that all eukaryogenesis models face: the need to script the blueprint of a eukaryotic cell. And should then not all (asgard) archaea with a syntrophic partner be considered FECA in the sense that in principle they have the potential to become eukaryotic? For 4 billion years, prokaryotes have overall remained the same in terms of cellular complexity with some rare exceptions having evolved a single compartment type, but nothing vaguely similar to the conserved nature of the eukaryotic endomembrane system. Reflecting on eukaryotic traits and their cell biological connection to mitochondrial origin lets us conclude they are better understood as being selected for to service an endosymbiont and less so as means of acquiring one. Phylogeny guided models should connect interpretations to a physiological and cell biological rationale, while facing the challenge of resolving the fluid nature of the pangenomes of both host and endosymbiont genome throughout eukaryogenesis – we need to talk to phylogenetic trees, not only about them. Physiology (*Imachi et al., 2020; Martin and Müller, 1998; Moreira, 1998; Spang et al., 2019*), geochemistry (*Mills et al., 2022*), phylogenetics (*Wu et al., 2022*), and culturing and imaging (*Imachi et al., 2020*) all point to a syntrophic origin of eukaryotes involving two prokaryotic partners. The data suggests that first steps towards endosymbiosis in eukaryogenesis were of prokaryotic nature, that eukaryogenesis likely only solidified upon endosymbiosis, and that hence the definition of FECA should include an endosymbiont.

Acknowledgements

We gratefully acknowledge funding by the VolkswagenStiftung (Life) and the DFG (SFB 1208-2672 05415). SGG also acknowledges and thanks the support by the Moore-Simons Project on the Origin of the Eukaryotic Cell (GBMF9743).

We acknowledge support from the high-performance computing cluster (HILBERT) at the Heinrich Heine University Düsseldorf.

Additional information

Funding

| Funder | Grant reference number | Author |
|---------------------------------|------------------------|---------------|
| Deutsche Forschungsgemeinschaft | SFB 1208-2672 05415 | Parth K Raval |
| Moore-Simons Foundation | GBMF9743 | Sriram G Garg |
| VolkswagenStiftung | Life | Sven B Gould |

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Parth K Raval, Sriram G Garg, Writing – original draft; Sven B Gould, Funding acquisition, Writing – original draft, Project administration

Author ORCIDs

Parth K Raval  <http://orcid.org/0000-0001-6151-437X>
Sriram G Garg  <http://orcid.org/0000-0003-4160-5228>
Sven B Gould  <http://orcid.org/0000-0002-2038-8474>

References

Abe T, Marutani Y, Shoji I. 2019. Cytosolic DNA-sensing immune response and viral infection. *Microbiology and Immunology* **63**:51–64. DOI: <https://doi.org/10.1111/1348-0421.12669>, PMID: 30677166

- Abu-Qarn M**, Giordano A, Battaglia F, Trauner A, Hitchen PG, Morris HR, Dell A, Eichler J. 2008. Identification of a second glycosyltransferase involved in N glycosylation of the *Haloferax volcanii* S-layer glycoprotein. *Journal of Bacteriology* **190**:3140–3146. DOI: <https://doi.org/10.1128/JB.00056-08>, PMID: 18310347
- Akil C**, Robinson RC. 2018. Genomes of asgard archaea encode Profilins that regulate actin. *Nature* **562**:439–443. DOI: <https://doi.org/10.1038/s41586-018-0548-6>, PMID: 30283132
- Akil C**, Tran LT, Orhant-Prioux M, Baskaran Y, Manser E, Blanchoin L, Robinson RC. 2020. Insights into the evolution of regulated actin dynamics via characterization of primitive gelsolin/cofilin proteins from asgard archaea. *PNAS* **117**:19904–19913. DOI: <https://doi.org/10.1073/pnas.2009167117>, PMID: 32747565
- Allen JF**. 2015. Why chloroplasts and mitochondria retain their own genomes and genetic systems: collocation for redox regulation of gene expression. *PNAS* **112**:10231–10238. DOI: <https://doi.org/10.1073/pnas.1500012112>, PMID: 26286985
- Alvarez-Ponce D**, Lopez P, Baptiste E, McInerney JO. 2013. Gene similarity networks provide tools for understanding eukaryote origins and evolution. *PNAS* **110**:E1594–E1603. DOI: <https://doi.org/10.1073/pnas.1211371110>, PMID: 23576716
- Anderson DJ**, Hetzer MW. 2008. Shaping the endoplasmic reticulum into the nuclear envelope. *Journal of Cell Science* **121**:137–142. DOI: <https://doi.org/10.1242/jcs.005777>, PMID: 18187447
- Anding AL**, Baehrecke EH. 2017. Cleaning house: selective autophagy of organelles. *Developmental Cell* **41**:10–22. DOI: <https://doi.org/10.1016/j.devcel.2017.02.016>, PMID: 28399394
- Antico Arciuch VG**, Elguero ME, Poderoso JJ, Carreras MC. 2012. Mitochondrial regulation of cell cycle and proliferation. *Antioxidants & Redox Signaling* **16**:1150–1180. DOI: <https://doi.org/10.1089/ars.2011.4085>, PMID: 21967640
- Backes S**, Garg SG, Becker L, Peleh V, Glockshuber R, Gould SB, Herrmann JM. 2019. Development of the mitochondrial intermembrane space disulfide relay represents a critical step in eukaryotic evolution. *Molecular Biology and Evolution* **36**:742–756. DOI: <https://doi.org/10.1093/molbev/msz011>, PMID: 30668797
- Baer CF**, Miyamoto MM, Denver DR. 2007. Mutation rate variation in multicellular eukaryotes: causes and consequences. *Nature Reviews. Genetics* **8**:619–631. DOI: <https://doi.org/10.1038/nrg2158>, PMID: 17637734
- Bar YM**, Phillips R, Milo R. 2018. The biomass distribution on earth. *PNAS* **115**:6506–6511. DOI: <https://doi.org/10.1073/pnas.1711842115>, PMID: 29784790
- Barbrook AC**, Howe CJ, Purton S. 2006. Why are plastid genomes retained in non-photosynthetic organisms? *Trends in Plant Science* **11**:101–108. DOI: <https://doi.org/10.1016/j.tplants.2005.12.004>, PMID: 16406301
- Baum B**, Baum DA. 2020. The merger that made us. *BMC Biology* **18**:1–4. DOI: <https://doi.org/10.1186/s12915-020-00806-3>, PMID: 32580772
- Biswas A**, Thattai M. 2020. Promiscuity and specificity of eukaryotic glycosyltransferases. *Biochemical Society Transactions* **48**:891–900. DOI: <https://doi.org/10.1042/BST20190651>, PMID: 32539082
- Bolte K**, Rensing SA, Maier UG. 2015. The evolution of eukaryotic cells from the perspective of peroxisomes: phylogenetic analyses of peroxisomal beta-oxidation enzymes support mitochondria-first models of eukaryotic cell evolution. *BioEssays* **37**:195–203. DOI: <https://doi.org/10.1002/bies.201400151>, PMID: 25394329
- Booth DM**, Enyedi B, Geiszt M, Várnai P, Hajnóczky G. 2016. Redox nanodomains are induced by and control calcium signaling at the ER-mitochondrial interface. *Molecular Cell* **63**:240–248. DOI: <https://doi.org/10.1016/j.molcel.2016.05.040>, PMID: 27397688
- Brocchieri L**, Karlin S. 2005. Protein length in eukaryotic and prokaryotic proteomes. *Nucleic Acids Research* **33**:3390–3400. DOI: <https://doi.org/10.1093/nar/gki615>, PMID: 15951512
- Brueckner J**, Martin WF. 2020. Bacterial genes outnumber archaeal genes in eukaryotic genomes. *Genome Biology and Evolution* **12**:282–292. DOI: <https://doi.org/10.1093/gbe/evaa047>, PMID: 32142116
- Brunk CF**, Martin WF. 2019. Archaeal histone contributions to the origin of eukaryotes. *Trends in Microbiology* **27**:703–714. DOI: <https://doi.org/10.1016/j.tim.2019.04.002>, PMID: 31076245
- Burki F**, Roger AJ, Brown MW, Simpson AGB. 2020. The new tree of eukaryotes. *Trends in Ecology & Evolution* **35**:43–55. DOI: <https://doi.org/10.1016/j.tree.2019.08.008>, PMID: 31606140
- Cardenas-Rodriguez M**, Tokatlidis K. 2017. Cytosolic redox components regulate protein homeostasis via additional localisation in the mitochondrial intermembrane space. *FEBS Letters* **591**:2661–2670. DOI: <https://doi.org/10.1002/1873-3468.12766>, PMID: 28746987
- Cavalier-Smith T**. 1985. *The Evolution of Genome Size*. Chichester: John Wiley.
- Cavalier-Smith T**. 1987. Eukaryotes with no mitochondria. *Nature* **326**:332–333. DOI: <https://doi.org/10.1038/326332a0>, PMID: 3561476
- Cavalier-Smith T**. 2005. Economy, speed and size matter: evolutionary forces driving nuclear genome miniaturization and expansion. *Annals of Botany* **95**:147–175. DOI: <https://doi.org/10.1093/aob/mci010>
- Christ L**, Wenzel EM, Liestøl K, Raiborg C, Campsteijn C, Stenmark H. 2016. Alix and ESCRT-III function as parallel ESCRT-III recruiters in cytotkinetic abscission. *The Journal of Cell Biology* **212**:499–513. DOI: <https://doi.org/10.1083/jcb.201507009>, PMID: 26929449
- Colnaghi M**, Lane N, Pomiankowski A. 2020. Genome expansion in early eukaryotes drove the transition from lateral gene transfer to meiotic sex. *eLife* **9**:e58873. DOI: <https://doi.org/10.7554/eLife.58873>, PMID: 32990598
- Colnaghi M**, Lane N, Pomiankowski A. 2022. Repeat sequences limit the effectiveness of lateral gene transfer and favored the evolution of meiotic sex in early eukaryotes. *PNAS* **119**:1–7. DOI: <https://doi.org/10.1073/pnas.2205041119>, PMID: 35994648

- Dacks JB**, Field MC, Buick R, Eme L, Gribaldo S, Roger AJ, Brochier-Armanet C, Devos DP. 2016. The changing view of eukaryogenesis-fossils, cells, lineages and how they all come together. *Journal of Cell Science* **129**:3695–3703. DOI: <https://doi.org/10.1242/jcs.178566>, PMID: 27672020
- Darwin C**, Murray J. 1859. On the Origin of Species by Means of Natural Selection, or, The Preservation of Favoured Races in the Struggle for Life. London: William Clowes and Sons. DOI: <https://doi.org/10.5962/bhl.title.82303>
- Deatherae BL**, Cookson BT. 2012. Membrane vesicle release in bacteria, eukaryotes, and archaea: a conserved yet underappreciated aspect of microbial life. *Infection and Immunity* **80**:1948–1957. DOI: <https://doi.org/10.1128/IAI.06014-11>, PMID: 22409932
- Dell A**, Galadari A, Sastre F, Hitchen P. 2010. Similarities and differences in the glycosylation mechanisms in prokaryotes and eukaryotes. *International Journal of Microbiology* **2010**:148178. DOI: <https://doi.org/10.1155/2010/148178>, PMID: 21490701
- Dey G**, Thattai M, Baum B. 2016. On the archaeal origins of eukaryotes and the challenges of inferring phenotype from genotype. *Trends in Cell Biology* **26**:476–485. DOI: <https://doi.org/10.1016/j.tcb.2016.03.009>, PMID: 27319280
- Dill KA**, Ghosh K, Schmit JD. 2011. Physical limits of cells and proteomes. *PNAS* **108**:17876–17882. DOI: <https://doi.org/10.1073/pnas.1114477108>, PMID: 22006304
- Dominguez DC**. 2004. Calcium signalling in bacteria. *Molecular Microbiology* **54**:291–297. DOI: <https://doi.org/10.1111/j.1365-2958.2004.04276.x>, PMID: 15469503
- Duggin IG**, Aylett CHS, Walsh JC, Michie KA, Wang Q, Turnbull L, Dawson EM, Harry EJ, Whitchurch CB, Amos LA, Löwe J. 2015. Cetz tubulin-like proteins control archaeal cell shape. *Nature* **519**:362–365. DOI: <https://doi.org/10.1038/nature13983>, PMID: 25533961
- Eme L**, Spang A, Lombard J, Stairs CW, Ettema TJG. 2017. Archaea and the origin of eukaryotes. *Nature Reviews. Microbiology* **15**:711–723. DOI: <https://doi.org/10.1038/nrmicro.2017.133>, PMID: 29123225
- Eyre-Walker A**, Keightley PD. 2007. The distribution of fitness effects of new mutations. *Nature Reviews. Genetics* **8**:610–618. DOI: <https://doi.org/10.1038/nrg2146>, PMID: 17637733
- Fenchel T**. 2014. Respiration in heterotrophic unicellular eukaryotic organisms. *Protist* **165**:485–492. DOI: <https://doi.org/10.1016/j.protis.2014.05.006>, PMID: 24995585
- Fitzhugh K**. 2016. Sequence data, phylogenetic inference, and implications of downward causation. *Acta Biotheoretica* **64**:133–160. DOI: <https://doi.org/10.1007/s10441-016-9277-0>, PMID: 26961079
- FlisV**, Daum G. 2013. Lipid transport between the endoplasmic. *Cold Spring Harbor Perspectives in Biology* **5**:1–22. DOI: <https://doi.org/10.1101/cshperspect.a013235>
- Fox GE**, Stackebrandt E, Hespell RB, Gibson J, Maniloff J, Dyer TA, Wolfe RS, Balch WE, Tanner RS, Magrum LJ, Zablen LB, Blakemore R, Gupta R, Bonen L, Lewis BJ, Stahl DA, Luehrsen KR, Chen KN, Woese CR. 1980. The phylogeny of prokaryotes. *Science* **209**:457–463. DOI: <https://doi.org/10.1126/science.6771870>, PMID: 6771870
- Friedman JR**, Lackner LL, West M, DiBenedetto JR, Nunnari J, Voeltz GK. 2011. Er tubules mark sites of mitochondrial division. *Science* **334**:358–362. DOI: <https://doi.org/10.1126/science.1207385>, PMID: 21885730
- Fritz-Laylin LK**, Prochnik SE, Ginger ML, Dacks JB, Carpenter ML, Field MC, Kuo A, Paredez A, Chapman J, Pham J, Shu S, Neupane R, Cipriano M, Mancuso J, Tu H, Salamov A, Lindquist E, Shapiro H, Lucas S, Grigoriev IV, et al. 2010. The genome of *Naegleria gruberi* illuminates early eukaryotic versatility. *Cell* **140**:631–642. DOI: <https://doi.org/10.1016/j.cell.2010.01.032>, PMID: 20211133
- Fujimoto M**, Tsutsumi N. 2014. Dynamin-Related proteins in plant post-Golgi traffic. *Frontiers in Plant Science* **5**:408. DOI: <https://doi.org/10.3389/fpls.2014.00408>, PMID: 25237312
- Futuyma DJ**. 1986. *Evolutionary Biology*. Springer.
- Garg SG**, Martin WF. 2016. Mitochondria, the cell cycle, and the origin of sex via a syncytial eukaryote common ancestor. *Genome Biology and Evolution* **8**:1950–1970. DOI: <https://doi.org/10.1093/gbe/evw136>, PMID: 27345956
- Garg SG**, Kapust N, Lin W, Knopp M, Tria FDK, Nelson-Sathi S, Gould SB, Fan L, Zhu R, Zhang C, Martin WF. 2021. Anomalous phylogenetic behavior of ribosomal proteins in metagenome-assembled asgard archaea. *Genome Biology and Evolution* **13**:1–12. DOI: <https://doi.org/10.1093/gbe/evaa238>, PMID: 33462601
- Gerlitz M**, Knopp M, Kapust N, Xavier JC, Martin WF. 2018. Elusive data underlying debate at the prokaryote-eukaryote divide. *Biology Direct* **13**:1–12. DOI: <https://doi.org/10.1186/s13062-018-0221-x>, PMID: 31196150
- Goliand I**, Nachmias D, Gershony O, Elia N. 2014. Inhibition of ESCRT-II-CHMP6 interactions impedes cytokinetic abscission and leads to cell death. *Molecular Biology of the Cell* **25**:3740–3748. DOI: <https://doi.org/10.1091/mbc.E14-08-1317>, PMID: 25232011
- Gould SB**, Garg SG, Martin WF. 2016. Bacterial vesicle secretion and the evolutionary origin of the eukaryotic endomembrane system. *Trends in Microbiology* **24**:525–534. DOI: <https://doi.org/10.1016/j.tim.2016.03.005>, PMID: 27040918
- Graur D**, Martin W. 2004. Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. *Trends in Genetics* **20**:80–86. DOI: <https://doi.org/10.1016/j.tig.2003.12.003>, PMID: 14746989
- Guy L**, Ettema TJG. 2011. The archaeal “tack” superphylum and the origin of eukaryotes. *Trends in Microbiology* **19**:580–587. DOI: <https://doi.org/10.1016/j.tim.2011.09.002>, PMID: 22018741
- Haeusser DP**, Levin PA. 2008. The great divide: coordinating cell cycle events during bacterial growth and division. *Current Opinion in Microbiology* **11**:94–99. DOI: <https://doi.org/10.1016/j.mib.2008.02.008>, PMID: 18396093

- Halligan DL, Keightley PD. 2009. Spontaneous mutation accumulation studies in evolutionary genetics. *Annual Review of Ecology, Evolution, and Systematics* **40**:151–172. DOI: <https://doi.org/10.1146/annurev.ecolsys.39.110707.173437>
- Hamasaki M, Furuta N, Matsuda A, Nezu A, Yamamoto A, Fujita N, Oomori H, Noda T, Haraguchi T, Hiraoka Y, Amano A, Yoshimori T. 2013. Autophagosomes form at ER-mitochondria contact sites. *Nature* **495**:389–393. DOI: <https://doi.org/10.1038/nature11910>, PMID: 23455425
- Hampfl V, Čepička I, Eliáš M. 2019. Was the mitochondrion necessary to start eukaryogenesis? *Trends in Microbiology* **27**:96–104. DOI: <https://doi.org/10.1016/j.tim.2018.10.005>, PMID: 30466901
- Harashima H, Dissmeyer N, Schnittger A. 2013. Cell cycle control across the eukaryotic Kingdom. *Trends in Cell Biology* **23**:345–356. DOI: <https://doi.org/10.1016/j.tcb.2013.03.002>, PMID: 23566594
- Harold FM. 1986. *The Vital Force: A Study of Bioenergetics*. W.H. Freeman & Co Ltd.
- Hartwell LH, Hopfield JJ, Leibler S, Murray AW. 1999. Hartwell et al 199. *Nature* **402**:C47–C52. DOI: <https://doi.org/10.1038/35011540>, PMID: 10591225
- Hennig W. 2000. *Phylogenetic Systematics*. University of Illinois Press.
- Hjort K, Goldberg AV, Tsaousis AD, Hirt RP, Embley TM. 2010. Diversity and reductive evolution of mitochondria among microbial eukaryotes. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **365**:713–727. DOI: <https://doi.org/10.1098/rstb.2009.0224>, PMID: 20124340
- Hochachka PW, Darveau CA, Andrews RD, Suarez RK. 2003. Allometric cascade: a model for resolving body mass effects on metabolism. *Comparative Biochemistry and Physiology. Part A, Molecular & Integrative Physiology* **134**:675–691. DOI: [https://doi.org/10.1016/s1095-6433\(02\)00364-1](https://doi.org/10.1016/s1095-6433(02)00364-1), PMID: 12814777
- Hörandl E, Speijer D. 2018. How oxygen gave rise to eukaryotic sex. *Proceedings. Biological Sciences* **285**:20172706. DOI: <https://doi.org/10.1098/rspb.2017.2706>, PMID: 29436502
- Husnik F, Tashyreva D, Boscaro V, George EE, Lukeš J, Keeling PJ. 2021. Bacterial and archaeal symbioses with protists. *Current Biology* **31**:R862–R877. DOI: <https://doi.org/10.1016/j.cub.2021.05.049>, PMID: 34256922
- Imachi H, Nobu MK, Nakahara N, Morono Y, Ogawara M, Takaki Y, Takano Y, Uematsu K, Ikuta T, Ito M, Matsui Y, Miyazaki M, Murata K, Saito Y, Sakai S, Song C, Tasumi E, Yamanaka Y, Yamaguchi T, Kamagata Y, et al. 2020. Isolation of an archaeon at the prokaryote-eukaryote interface. *Nature* **577**:519–525. DOI: <https://doi.org/10.1038/s41586-019-1916-6>, PMID: 31942073
- Ishida H, Izumi M, Wada S, Makino A. 2014. Biochimica et biophysica acta roles of autophagy in chloroplast recycling ☆. *BBA - Bioenergetics* **1837**:512–521. DOI: <https://doi.org/10.1016/j.bbabi.2013.11.009>, PMID: 24269172
- Islinger M, Cardoso MJR, Schrader M. 2010. Be different—the diversity of peroxisomes in the animal kingdom. *Biochimica et Biophysica Acta* **1803**:881–897. DOI: <https://doi.org/10.1016/j.bbamcr.2010.03.013>, PMID: 20347886
- Jarrell KF, Ding Y, Meyer BH, Albers SV, Kaminski L, Eichler J. 2014. N-Linked glycosylation in Archaea: a structural, functional, and genetic analysis. *Microbiology and Molecular Biology Reviews* **78**:304–341. DOI: <https://doi.org/10.1128/MMBR.00052-13>, PMID: 24847024
- Kafri M, Metzl-Raz E, Jona G, Barkai N. 2016. The cost of protein production. *Cell Reports* **14**:22–31. DOI: <https://doi.org/10.1016/j.celrep.2015.12.015>, PMID: 26725116
- Karnkowska A, Vacek V, Zubáčová Z, Treitli SC, Petřelková R, Eme L, Novák L, Žárský V, Barlow LD, Herman EK, Soukal P, Hroudová M, Doležal P, Stairs CW, Roger AJ, Eliáš M, Dacks JB, Vlček Č, Hampfl V. 2016. A eukaryote without a mitochondrial organelle. *Current Biology* **26**:1274–1284. DOI: <https://doi.org/10.1016/j.cub.2016.03.053>, PMID: 27185558
- Kelly S. 2021. The economics of organellar gene loss and endosymbiotic gene transfer. *Genome Biology* **22**:1–22. DOI: <https://doi.org/10.1186/s13059-021-02567-w>, PMID: 34930424
- Klinger CM, Spang A, Dacks JB, Ettema TJG. 2016. Tracing the archaeal origins of eukaryotic membrane-traffic system building blocks. *Molecular Biology and Evolution* **33**:1528–1541. DOI: <https://doi.org/10.1093/molbev/msw034>, PMID: 26893300
- Knopp M, Stockhorst S, van der Giezen M, Garg SG, Gould SB. 2021. The asgard archaeal-unique contribution to protein families of the eukaryotic common ancestor was 0.3. *Genome Biology and Evolution* **13**:1–10. DOI: <https://doi.org/10.1093/gbe/evab085>, PMID: 33892498
- Koga Y, Kyuragi T, Nishihara M, Sone N. 1998. Did archaeal and bacterial cells arise independently from noncellular precursors? A hypothesis stating that the advent of membrane phospholipid with enantiomeric glycerophosphate backbones caused the separation of the two lines of descent. *Journal of Molecular Evolution* **46**:54–63. DOI: <https://doi.org/10.1007/pl00006283>, PMID: 9419225
- Kontou A, Herman EK, Field MC, Dacks JB, Koumandou VL. 2022. Evolution of factors shaping the endoplasmic reticulum. *Traffic* **23**:462–473. DOI: <https://doi.org/10.1111/tra.12863>, PMID: 36040076
- Koonin EV, Yutin N. 2014. The dispersed archaeal eukaryome and the complex archaeal ancestor of eukaryotes. *Cold Spring Harbor Perspectives in Biology* **6**:a016188. DOI: <https://doi.org/10.1101/cshperspect.a016188>, PMID: 24691961
- Kusdian G, Woehle C, Martin WF, Gould SB. 2013. The actin-based machinery of trichomonas vaginalis mediates flagellate-amoeboid. *Cellular Microbiology* **15**:1707–1721. DOI: <https://doi.org/10.1111/cmi.12144>, PMID: 23530917
- Lahti DC, Johnson NA, Ajie BC, Otto SP, Hendry AP, Blumstein DT, Coss RG, Donohue K, Foster SA. 2009. Relaxed selection in the wild. *Trends in Ecology & Evolution* **24**:487–496. DOI: <https://doi.org/10.1016/j.tree.2009.03.010>, PMID: 19500875

- Land M**, Hauser L, Jun SR, Nookaew I, Leuze MR, Ahn TH, Karpinetis T, Lund O, Kora G, Wassenaar T, Poudel S, Ussery DW. 2015. Insights from 20 years of bacterial genome sequencing. *Functional & Integrative Genomics* **15**:141–161. DOI: <https://doi.org/10.1007/s10142-015-0433-4>, PMID: 25722247
- Lane N**. 2007. Mitochondria: key to complexity. Andersson SG (Ed). *Origin of Mitochondria and Hydrogenosomes*. Springer. p. 13–38. DOI: https://doi.org/10.1007/978-3-540-38502-8_2
- Lane N**, Martin W. 2010. The energetics of genome complexity. *Nature* **467**:929–934. DOI: <https://doi.org/10.1038/nature09486>, PMID: 20962839
- Lane N**. 2020. How energy flow shapes cell evolution. *Current Biology* **30**:R471–R476. DOI: <https://doi.org/10.1016/j.cub.2020.03.055>, PMID: 32428484
- Larsen RA**, Cusumano C, Fujioka A, Lim-Fong G, Patterson P, Pogliano J. 2007. Treadmilling of a prokaryotic tubulin-like protein, TubZ, required for plasmid stability in *Bacillus thuringiensis*. *Genes & Development* **21**:1340–1352. DOI: <https://doi.org/10.1101/gad.154610>, PMID: 17510284
- Le T**, Žárský V, Nýlvtová E, Rada P, Harant K, Vancová M, Verner Z, Hrdý I, Tachezy J. 2020. Anaerobic peroxisomes in *mastigamoeba balamuthi*. *PNAS* **117**:2065–2075. DOI: <https://doi.org/10.1073/pnas.1909755117>
- Leão P**, Gueiros-Filho FJ, Bazylinski DA, Lins U, Abreu F. 2018. Association of magnetotactic multicellular prokaryotes with pseudoalteromonas species in a natural lagoon environment. *Antonie van Leeuwenhoek* **111**:2213–2223. DOI: <https://doi.org/10.1007/s10482-018-1113-x>, PMID: 29915893
- Lhee D**, Ha J-S, Kim S, Park MG, Bhattacharya D, Yoon HS. 2019. Evolutionary dynamics of the chromatophore genome in three photosynthetic paulinella species. *Scientific Reports* **9**:1–11. DOI: <https://doi.org/10.1038/s41598-019-38621-8>, PMID: 30796245
- Light DW**, Lexchin J. 2021. The costs of coronavirus vaccines and their pricing. *Journal of the Royal Society of Medicine* **114**:502–504. DOI: <https://doi.org/10.1177/01410768211053006>, PMID: 34732097
- Liu Y**, Makarova KS, Huang W-C, Wolf YI, Nikolskaya AN, Zhang X, Cai M, Zhang C-J, Xu W, Luo Z, Cheng L, Koonin EV, Li M. 2021. Expanded diversity of asgard archaea and their relationships with eukaryotes. *Nature* **593**:553–557. DOI: <https://doi.org/10.1038/s41586-021-03494-3>, PMID: 33911286
- Lombard J**. 2016. The multiple evolutionary origins of the eukaryotic N-glycosylation pathway. *Biology Direct* **11**:36. DOI: <https://doi.org/10.1186/s13062-016-0137-2>, PMID: 27492357
- López-García P**, Moreira D. 2019. Eukaryogenesis, a syntrophy affair. *Nature Microbiology* **4**:1068–1070. DOI: <https://doi.org/10.1038/s41564-019-0495-5>, PMID: 31222170
- López-García P**, Moreira D. 2020. The syntrophy hypothesis for the origin of eukaryotes revisited. *Nature Microbiology* **5**:655–667. DOI: <https://doi.org/10.1038/s41564-020-0710-4>, PMID: 32341569
- Lynch M**, Marinov GK. 2017. Membranes, energetics, and evolution across the prokaryote-eukaryote divide. *eLife* **6**:e20437. DOI: <https://doi.org/10.7554/eLife.20437>, PMID: 28300533
- Makarova KS**, Wolf YI, Mekhedov SL, Mirkin BG, Koonin EV. 2005. Ancestral paralogs and pseudoparalogs and their role in the emergence of the eukaryotic cell. *Nucleic Acids Research* **33**:4626–4638. DOI: <https://doi.org/10.1093/nar/gki775>, PMID: 16106042
- Malik SB**, Pightling AW, Stefaniak LM, Schurko AM, Logsdon JM. 2008. An expanded inventory of conserved meiotic genes provides evidence for sex in *trichomonas vaginalis*. *PLOS ONE* **3**:e2879. DOI: <https://doi.org/10.1371/journal.pone.0002879>, PMID: 18663385
- Martin W**, Müller M. 1998. The hydrogen hypothesis for the first eukaryote. *Nature* **392**:37–41. DOI: <https://doi.org/10.1038/32096>, PMID: 9510246
- Martin W**, Koonin EV. 2006. Introns and the origin of nucleus-cytosol compartmentalization. *Nature* **440**:41–45. DOI: <https://doi.org/10.1038/nature04531>, PMID: 16511485
- Martin WF**, Garg S, Zimorski V. 2015. Endosymbiotic theories for eukaryote origin. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **370**:20140330. DOI: <https://doi.org/10.1098/rstb.2014.0330>, PMID: 26323761
- Martin WF**. 2017. Symbiogenesis, gradualism, and mitochondrial energy in eukaryote origin. *Periodicum Biologorum* **119**:141–158. DOI: <https://doi.org/10.18054/pb.v119i3.5694>
- Martin WF**, Roettger M, Ku C, Garg SG, Nelson-Sathi S, Landan G. 2017a. Late mitochondrial origin is an artifact. *Genome Biology and Evolution* **9**:373–379. DOI: <https://doi.org/10.1093/gbe/evx027>, PMID: 28199635
- Martin WF**, Tielens AGM, Mentel M, Garg SG, Gould SB. 2017b. The physiology of phagocytosis in the context of mitochondrial origin. *Microbiology and Molecular Biology Reviews* **81**:1–36. DOI: <https://doi.org/10.1128/MMBR.00008-17>, PMID: 28615286
- Mills DB**. 2020. The origin of phagocytosis in earth history. *Interface Focus* **10**:0–2. DOI: <https://doi.org/10.1098/rsfs.2020.0019rsfs20200019>
- Mills DB**, Boyle RA, Daines SJ, Sperling EA, Pisani D, Donoghue PCJ, Lenton TM. 2022. Eukaryogenesis and oxygen in earth history. *Nature Ecology & Evolution* **6**:520–532. DOI: <https://doi.org/10.1038/s41559-022-01733-y>, PMID: 35449457
- Mitra K**, Wunder C, Roysam B, Lin G, Lippincott-Schwartz J. 2009. A hyperfused mitochondrial state achieved at G1-S regulates cyclin E buildup and entry into S phase. *PNAS* **106**:11960–11965. DOI: <https://doi.org/10.1073/pnas.0904875106>, PMID: 19617534
- Mohanty A**, McBride HM. 2013. Emerging roles of mitochondria in the evolution, biogenesis, and function of peroxisomes. *Frontiers in Physiology* **4**:1–12. DOI: <https://doi.org/10.3389/fphys.2013.00268>, PMID: 24133452
- Moreira D**. 1998. Symbiosis between methanogenic archaea and delta-proteobacteria as the origin of eukaryotes: the syntrophic hypothesis. *Journal of Molecular Evolution* **47**:517–530. DOI: <https://doi.org/10.1007/pl00006408>, PMID: 9797402

- Muller HJ. 1964. The relation of recombination to mutational advance. *Mutation Research* **106**:2–9. DOI: [https://doi.org/10.1016/0027-5107\(64\)90047-8](https://doi.org/10.1016/0027-5107(64)90047-8), PMID: 14195748
- Müller JB, Geyer PE, Colaço AR, Treit PV, Strauss MT, Oroshi M, Doll S, Virreira Winter S, Bader JM, Köhler N, Theis F, Santos A, Mann M. 2020. The proteome landscape of the kingdoms of life. *Nature* **582**:592–596. DOI: <https://doi.org/10.1038/s41586-020-2402-x>, PMID: 32555458
- Nakatogawa H. 2020. Mechanisms governing autophagosome biogenesis. *Nature Reviews. Molecular Cell Biology* **21**:439–458. DOI: <https://doi.org/10.1038/s41580-020-0241-0>, PMID: 32372019
- Nelson-Sathi S, Dagan T, Landan G, Janssen A, Steel M, Mclnerney JO, Deppenmeier U, Martin WF. 2012. Acquisition of 1,000 eubacterial genes physiologically transformed a methanogen at the origin of haloarchaea. *PNAS* **109**:20537–20542. DOI: <https://doi.org/10.1073/pnas.1209119109>, PMID: 23184964
- Nerurkar P, Altvater M, Gerhardy S, Schütz S, Fischer U, Weirich C, Panse VG. 2015. Eukaryotic ribosome assembly and nuclear export. *International Review of Cell and Molecular Biology* **319**:107–140. DOI: <https://doi.org/10.1016/bs.ircmb.2015.07.002>, PMID: 26404467
- Neveu E, Khalifeh D, Salamin N, Fasshauer D. 2020. Prototypic SNARE proteins are encoded in the genomes of heimdallarchaeota, potentially bridging the gap between the prokaryotes and eukaryotes. *Current Biology* **30**:2468–2480. DOI: <https://doi.org/10.1016/j.cub.2020.04.060>, PMID: 32442459
- Olmos Y, Hodgson L, Mantell J, Verkade P, Carlton JG. 2015. ESCRT-III controls nuclear envelope reformation. *Nature* **522**:236–239. DOI: <https://doi.org/10.1038/nature14503>, PMID: 26040713
- Paludan SR, Bowie AG. 2013. Immune sensing of DNA. *Immunity* **38**:870–880. DOI: <https://doi.org/10.1016/j.immuni.2013.05.004>, PMID: 23706668
- Pellicer J, Fay MF, Leitch IJ. 2010. The largest eukaryotic genome of them all? *Botanical Journal of the Linnean Society* **164**:10–15. DOI: <https://doi.org/10.1111/j.1095-8339.2010.01072.x>
- Peña C, Hurt E, Panse VG. 2017. Eukaryotic ribosome assembly, transport and quality control. *Nature Structural & Molecular Biology* **24**:689–699. DOI: <https://doi.org/10.1038/nsmb.3454>, PMID: 28880863
- Pittis AA, Gabaldón T. 2016. Late acquisition of mitochondria by a host with chimaeric prokaryotic ancestry. *Nature* **531**:101–104. DOI: <https://doi.org/10.1038/nature16941>, PMID: 26840490
- Portugez S, Martin WF, Hazkani-Covo E. 2018. Mosaic mitochondrial-plastid insertions into the nuclear genome show evidence of both non-homologous end joining and homologous recombination. *BMC Evolutionary Biology* **18**:1–10. DOI: <https://doi.org/10.1186/s12862-018-1279-x>, PMID: 30390623
- Preisner H, Habicht J, Garg SG, Gould SB. 2018. Intermediate filament protein evolution and protists. *Cytoskeleton* **75**:231–243. DOI: <https://doi.org/10.1002/cm.21443>, PMID: 29573204
- Radzvilavicius AL, Blackstone NW. 2015. Conflict and cooperation in eukaryogenesis: implications for the timing of endosymbiosis and the evolution of sex. *Journal of the Royal Society, Interface* **12**:20150584. DOI: <https://doi.org/10.1098/rsif.2015.0584>, PMID: 26468067
- Raffaello A, Mammucari C, Gherardi G, Rizzuto R. 2016. Calcium at the center of cell signaling: interplay between endoplasmic reticulum, mitochondria, and lysosomes. *Trends in Biochemical Sciences* **41**:1035–1049. DOI: <https://doi.org/10.1016/j.tibs.2016.09.001>, PMID: 27692849
- Raices M, D'Angelo MA. 2012. Nuclear pore complex composition: a new regulator of tissue-specific and developmental functions. *Nature Reviews. Molecular Cell Biology* **13**:687–699. DOI: <https://doi.org/10.1038/nrm3461>, PMID: 23090414
- Rebeaud ME, Mallik S, Goloubinoff P, Tawfik DS. 2021. On the evolution of chaperones and cochaperones and the expansion of proteomes across the tree of life. *PNAS* **118**:e2020885118. DOI: <https://doi.org/10.1073/pnas.2020885118>, PMID: 34001607
- Robinson DF, Foulds LR. 1981. Comparison of phylogenetic trees. *Mathematical Biosciences* **53**:131–147. DOI: [https://doi.org/10.1016/0025-5564\(81\)90043-2](https://doi.org/10.1016/0025-5564(81)90043-2)
- Rodríguez-Trelles F, Tarrío R, Ayala FJ. 2001. Erratic overdispersion of three molecular clocks: gpdh, sod, and xdh. *PNAS* **98**:11405–11410. DOI: <https://doi.org/10.1073/pnas.201392198>, PMID: 11553790
- Schavemaker PE, Muñoz-Gómez SA. 2022. The role of mitochondrial energetics in the origin and diversification of eukaryotes. *Nature Ecology & Evolution* **6**:1307–1317. DOI: <https://doi.org/10.1038/s41559-022-01833-9>, PMID: 35915152
- Schuler MH, English AM, Xiao T, Campbell TJ, Shaw JM, Hughes AL. 2021. Mitochondrial-Derived compartments facilitate cellular adaptation to amino acid stress. *Molecular Cell* **81**:3786–3802. DOI: <https://doi.org/10.1016/j.molcel.2021.08.021>, PMID: 34547239
- Schwechheimer C, Kuehn MJ. 2015. Outer-Membrane vesicles from gram-negative bacteria: biogenesis and functions. *Nature Reviews. Microbiology* **13**:605–619. DOI: <https://doi.org/10.1038/nrmicro3525>, PMID: 26373371
- Shiratori T, Suzuki S, Kakizawa Y, Ishida KI. 2019. Phagocytosis-like cell engulfment by a planctomycete bacterium. *Nature Communications* **10**:5529. DOI: <https://doi.org/10.1038/s41467-019-13499-2>, PMID: 31827088
- Skejo J, Garg SG, Gould SB, Hendriksen M, Tria FDK, Bremer N, Franjević D, Blackstone NW, Martin WF. 2021. Evidence for a syncytial origin of eukaryotes from ancestral state reconstruction. *Genome Biology and Evolution* **13**:1–14. DOI: <https://doi.org/10.1093/gbe/evab096>, PMID: 33963405
- Sousa FL, Neukirchen S, Allen JF, Lane N, Martin WF. 2016. Lokiarchaeon is hydrogen dependent. *Nature Microbiology* **1**:16034. DOI: <https://doi.org/10.1038/nmicrobiol.2016.34>, PMID: 27572645
- Spang A, Stairs CW, Dombrowski N, Eme L, Lombard J, Caceres EF, Greening C, Baker BJ, Ettema TJG. 2019. Proposal of the reverse flow model for the origin of the eukaryotic cell based on comparative analyses of

- asgard archaeal metabolism. *Nature Microbiology* **4**:1138–1148. DOI: <https://doi.org/10.1038/s41564-019-0406-9>, PMID: 30936488
- Spang A**, Mahendrarajah TA, Offre P, Stairs CW. 2022. Evolving perspective on the origin and diversification of cellular life and the virosphere. *Genome Biology and Evolution* **14**:1–19. DOI: <https://doi.org/10.1093/gbe/evac034>, PMID: 35218347
- Speijer D**, Lukeš J, Eliáš M. 2015. Sex is a ubiquitous, ancient, and inherent attribute of eukaryotic life. *PNAS* **112**:8827–8834. DOI: <https://doi.org/10.1073/pnas.1501725112>, PMID: 26195746
- Speijer D**. 2017. Evolution of peroxisomes illustrates symbiogenesis. *BioEssays* **39**:1–8. DOI: <https://doi.org/10.1002/bies.201700050>, PMID: 28782202
- Speijer D**. 2020. Debating eukaryogenesis-part 2: how anachronistic Reasoning can lure us into inventing intermediates. *BioEssays* **42**:e1900153. DOI: <https://doi.org/10.1002/bies.201900153>, PMID: 32157725
- Stairs CW**, Leger MM, Roger AJ. 2015. Diversity and origins of anaerobic metabolism in mitochondria and related organelles. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **370**:20140326. DOI: <https://doi.org/10.1098/rstb.2014.0326>, PMID: 26323757
- Stanier RY**, Van Niel CB. 1962. The concept of a bacterium. *Archiv Fur Mikrobiologie* **42**:17–35. DOI: <https://doi.org/10.1007/BF00425185>, PMID: 13916221
- Stanier RY**, Doudoroff M, Adelberg EA. 1963. *The Microbial World*. Prentice-Hall.
- Stoten CL**, Carlton JG. 2018. ESCRT-dependent control of membrane remodelling during cell division. *Seminars in Cell & Developmental Biology* **74**:50–65. DOI: <https://doi.org/10.1016/j.semcdb.2017.08.035>, PMID: 28843980
- Sugiura A**, McLelland G, Fon EA, McBride HM. 2014. A new pathway for mitochondrial quality control: mitochondrial-derived vesicles. *The EMBO Journal* **33**:2142–2156. DOI: <https://doi.org/10.15252/embj.201488104>, PMID: 25107473
- Sugiura A**, Mattie S, Prudent J, McBride HM. 2017. Newly born peroxisomes are a hybrid of mitochondrial and ER-derived pre-peroxisomes. *Nature* **542**:251–254. DOI: <https://doi.org/10.1038/nature21375>, PMID: 28146471
- Surkont J**, Pereira-Leal JB. 2016. Are there rab gtpases in archaea? *Molecular Biology and Evolution* **33**:1833–1842. DOI: <https://doi.org/10.1093/molbev/msw061>, PMID: 27034425
- Survery S**, Hurtig F, Haq SR, Eriksson J, Guy L, Rosengren KJ, Lindås AC, Chi CN. 2021. Heimdallarchaea encodes profilin with eukaryotic-like actin regulation and polyproline binding. *Communications Biology* **4**:1024. DOI: <https://doi.org/10.1038/s42003-021-02543-x>, PMID: 34471213
- Tarrason Risa G**, Hurtig F, Bray S, Hafner AE, Harker-Kirschneck L, Faull P, Davis C, Papatziadou D, Mutavchiev DR, Fan C, Meneguello L, Arashiro Pulschen A, Dey G, Culley S, Kilkenny M, Souza DP, Pellegrini L, de Bruin RAM, Henriques R, Snijders AP, et al. 2020. The proteasome controls ESCRT-III-mediated cell division in an archaeon. *Science* **369**:eaaz2532. DOI: <https://doi.org/10.1126/science.aaz2532>, PMID: 32764038
- Thormählen I**, Zupok A, Rescher J, Leger J, Weissenberger S, Groysman J, Orwat A, Chatel-Innocenti G, Issakidis-Bourguet E, Armbruster U, Geigenberger P. 2017. Thioredoxins play a crucial role in dynamic acclimation of photosynthesis in fluctuating light. *Molecular Plant* **10**:168–182. DOI: <https://doi.org/10.1016/j.molp.2016.11.012>, PMID: 27940305
- Tiley GP**, Poelstra JW, Dos Reis M, Yang Z, Yoder AD. 2020. Molecular clocks without rocks: new solutions for old problems. *Trends in Genetics* **36**:845–856. DOI: <https://doi.org/10.1016/j.tig.2020.06.002>, PMID: 32709458
- Timmis JN**, Ayliffe MA, Huang CY, Martin W. 2004. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nature Reviews. Genetics* **5**:123–135. DOI: <https://doi.org/10.1038/nrg1271>, PMID: 14735123
- Tovar J**, Fischer A, Clark CG. 1999. The mitosome, a novel organelle related to mitochondria in the amitochondrial parasite *Entamoeba histolytica*. *Molecular Microbiology* **32**:1013–1021. DOI: <https://doi.org/10.1046/j.1365-2958.1999.01414.x>, PMID: 10361303
- Tria FDK**, Brueckner J, Skejo J, Xavier JC, Kapust N, Knopp M, Wimmer JLE, Nagies FSP, Zimorski V, Gould SB, Garg SG, Martin WF. 2021. Gene duplications trace mitochondria to the onset of eukaryote complexity. *Genome Biology and Evolution* **13**:1–17. DOI: <https://doi.org/10.1093/gbe/evab055>, PMID: 33739376
- Van de Peer Y**, Maere S, Meyer A. 2009. The evolutionary significance of ancient genome duplications. *Nature Reviews. Genetics* **10**:725–732. DOI: <https://doi.org/10.1038/nrg2600>, PMID: 19652647
- van den Ent F**, Amos LA, Löwe J. 2001. Prokaryotic origin of the actin cytoskeleton. *Nature* **413**:39–44. DOI: <https://doi.org/10.1038/35092500>, PMID: 11544518
- Vellai T**, Vida G. 1999. The origin of eukaryotes: the difference between prokaryotic and eukaryotic cells. *Proceedings. Biological Sciences* **266**:1571–1577. DOI: <https://doi.org/10.1098/rspb.1999.0817>, PMID: 10467746
- Villanueva L**, Schouten S, Damsté JSS. 2017. Phylogenomic analysis of lipid biosynthetic genes of archaea shed light on the “lipid divide.” *Environmental Microbiology* **19**:54–69. DOI: <https://doi.org/10.1111/1462-2920.13361>, PMID: 27112361
- Vosseberg J**, van Hooff JJE, Marcet-Houben M, van Vlimmeren A, van Wijk LM, Gabaldón T, Snel B. 2021. Timing the origin of eukaryotic cellular complexity with ancient duplications. *Nature Ecology & Evolution* **5**:92–100. DOI: <https://doi.org/10.1038/s41559-020-01320-z>, PMID: 33106602
- Warner JR**. 1999. The economics of ribosome biosynthesis in yeast. *Trends in Biochemical Sciences* **24**:437–440. DOI: [https://doi.org/10.1016/s0968-0004\(99\)01460-7](https://doi.org/10.1016/s0968-0004(99)01460-7), PMID: 10542411
- Wheeler Q**, Assis L, Rieppel O. 2013. Phylogenetics: heed the father of cladistics. *Nature* **496**:295–296. DOI: <https://doi.org/10.1038/496295a>, PMID: 23598324

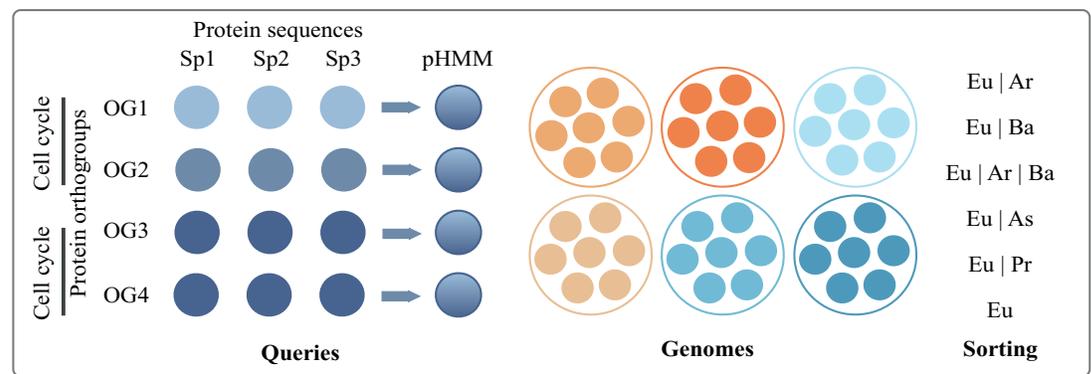
- Wickstead B**, Gull K. 2011. The evolution of the cytoskeleton. *The Journal of Cell Biology* **194**:513–525. DOI: <https://doi.org/10.1083/jcb.201102065>, PMID: 21859859
- Williams BAP**, Hirt RP, Lucocq JM, Embley TM. 2002. A mitochondrial remnant in the microsporidian *Trachipleistophora hominis*. *Nature* **418**:865–869. DOI: <https://doi.org/10.1038/nature00949>, PMID: 12192407
- Williams TA**, Embley TM. 2014. Archaeal “dark matter” and the origin of eukaryotes. *Genome Biology and Evolution* **6**:474–481. DOI: <https://doi.org/10.1093/gbe/evu031>, PMID: 24532674
- Woese CR**, Kandler O, Wheelis ML. 1990. Towards a natural system of organisms: proposal for the domains archaea, bacteria, and eucarya. *PNAS* **87**:4576–4579. DOI: <https://doi.org/10.1073/pnas.87.12.4576>, PMID: 2112744
- Wu F**, Speth DR, Philosof A, Crémière A, Narayanan A, Barco RA, Connon SA, Amend JP, Antoshechkin IA, Orphan VJ. 2022. Unique mobile elements and scalable gene flow at the prokaryote-eukaryote boundary revealed by circularized asgard archaea genomes. *Nature Microbiology* **7**:200–212. DOI: <https://doi.org/10.1038/s41564-021-01039-y>, PMID: 35027677
- Yamashita SI**, Jin X, Furukawa K, Hamasaki M, Nezu A, Otera H, Saigusa T, Yoshimori T, Sakai Y, Mihara K, Kanki T. 2016. Mitochondrial division occurs concurrently with autophagosome formation but independently of Drp1 during mitophagy. *The Journal of Cell Biology* **215**:649–665. DOI: <https://doi.org/10.1083/jcb.201605093>, PMID: 27903607
- Young KD**. 2006. The selective value of bacterial shape. *Microbiology and Molecular Biology Reviews* **70**:660–703. DOI: <https://doi.org/10.1128/MMBR.00001-06>, PMID: 16959965
- Yutin N**, Wolf MY, Wolf YI, Koonin EV. 2009. The origins of phagocytosis and eukaryogenesis. *Biology Direct* **4**:9. DOI: <https://doi.org/10.1186/1745-6150-4-9>, PMID: 19245710
- Zachar I**, Szilágyi A, Számadó S, Szathmáry E. 2018. Farming the mitochondrial ancestor as a model of endosymbiotic establishment by natural selection. *PNAS* **115**:E1504–E1510. DOI: <https://doi.org/10.1073/pnas.1718707115>, PMID: 29382768
- Zaremba-Niedzwiedzka K**, Caceres EF, Saw JH, Bäckström D, Juzokaite L, Vancaester E, Seitz KW, Anantharaman K, Starnawski P, Kjeldsen KU, Stott MB, Nunoura T, Banfield JF, Schramm A, Baker BJ, Spang A, Ettema TJG. 2017. Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* **541**:353–358. DOI: <https://doi.org/10.1038/nature21031>, PMID: 28077874

Appendix 1

Brief methods and criteria for sorting eukaryotic protein families (Figure 2 of the main text)

To screen for protein families involved in eukaryotic complexity in the prokaryotic world, we compartmentalized ‘eukaryotic complexity’ into: ER and Golgi (ER to lysosome transport and ER to Golgi transport), glycotransferases, autophagy, nucleus, cell cycle, meiosis, and peroxisome.

Protein orthogroups (a collection of protein sequences that are likely to perform a close to identical function across species) annotated under each of these categories were taken from the Kyoto Encyclopedia of Genes and Genomes (KEGG). Protein sequences under each of these orthogroups (annotated by unique KEGG Orthology or KO, ID) were obtained from all eukaryotic species available as of November 2021. From multiple protein sequences, hidden Markov models (HMM) were built, each representing a single protein orthogroup. These HMM profiles were used as query against all prokaryotic species available on the database. Searches were conducted using HMMER/3.1.



Appendix 1—figure 1. Overview of the pHMM screening pipeline to identify proteins (orthogroups) shared between eukaryotes and varying prokaryotic groups.

Bacterial species were categorized into phyla (e.g. Acetobacteria, Cyanobacteria) as per NCBI taxonomy, available also at KEGG. Archaeal species were categorized only into asgard and non-asgard. Each phyla had variable numbers of species and to account for this, we used % of species thresholds instead of number of species threshold. If a match was found in 5% of the species of a given phyla, we called it a match for that phylum to keep our estimates conservative. Protein families that had no hits in >5% species of any bacterial phylum (and asgard and non-asgard archaea) were considered unique to eukaryotes. Protein families present in >50% of bacterial phyla, asgard archaea, and non-asgard archaea were labelled as universal. Between the two extremes of universal and unique to eukaryotes, protein families with hits in a prokaryotic domain were further sorted into categories of relevance (shown in **Figure 2**, main text) as follows:

Appendix 1—table 1. Key to the different groups compared among each other.

| Category | Present in |
|---------------|---|
| Universal | >50% bacterial phyla, asgard and non-asgard archaea |
| Eu Ar Ba | <50% bacterial phyla, asgard and non-asgard archaea |
| Eu Ar Proteo | Asgard and non-asgard archaea, exclusively proteobacteria (no other bacterial phylum) |
| Eu Ar | Asgard and non-asgard archaea, no bacterial phylum |
| Eu Asgard | Asgard only (no other archaea or bacteria) |
| Eu non-Asgard | Non-asgard only (no other bacteria or asgard archaea) |

Appendix 1—table 1 Continued on next page

Appendix 1—table 1 Continued

| Category | Present in |
|-----------------|--|
| Eu Ba | Bacterial phylum other than proteobacteria |
| Eu Proteo | Only proteobacterial phyla |
| Eukaryotes (Eu) | No hits in >5% of any prokaryotic phylum |