Mitochondria, bioenergetics and the origins of eukaryotic complexity

Georgi K. Marinov[1] and Michael Lynch[2]

[1] *Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305, United States*
[2] *The Biodesign Center for Mechanisms of Evolution, Arizona State University, Tempe, AZ 85287, United States*

Abstract

Eukaryogenesis is thought to be one of the major events in the history of life on Earth, as the appearance of eukaryotes is often considered a prerequisite for the eventual evolution of complex multicellularity. Accordingly, figuring out how it happened has generated considerable interest over the decades, as has the more general question of what factors enabled/drove the evolution of complexity once eukaryotes had emerged. Phylogenomic studies over the last few decades have helped tremendously with the clarification of the phylogenetic relationships of eukaryotes, which are now understood to be the product of an endosymbiosis between an archaeal host and an α-proteobacterial symbiont. However, debate continues regarding the cellular complexity of the host, the role the protomitochondrion played in the process, and the factors that steered eukaryote evolution in the direction of further increases in complexity. The hypothesis that the mitochondrion provided an enormous boost in available "power per gene" relative to prokaryotes, allowing cellular and genomic complexity to expand, has generated significant controversy. Here, we examine this hypothesis in the light of our current understanding of the *process* of eukaryogenesis, discuss why a mitochondrial energy boost cannot explain the origin of complex eukaryotic features, and consider alternative explanations for their appearance.

Cellular life on Earth is divided in two fundamentally different categories in terms of organization – prokaryotes and eukaryotes. While a few examples of simple multicellularity can be found in prokaryotes [1–4], multicellularity has evolved on a large number of occasions in eukaryotes [5, 6], and complex multicellular organisms can only be found among them. The key lineages include land plants, fungi and metazoans, the latter featuring human beings, which are usually perceived, deservedly or not, as the pinnacle of life's complexity. For these reasons, the appearance of a eukaryote-level cellular organization is typically seen as a prerequisite for the evolution of "truly complex" and eventually sentient organisms, even if it is not possible to strictly prove the validity of such a claim. Eukaryogenesis is therefore often viewed as one of the key evolutionary transitions in the history of life on our planet, thereby eliciting substantial attention from theoretical and organismal biologists over the decades.

A robust scenario for eukaryogenesis needs to explain both the mechanistic details of how the process unfolded and the evolutionary forces that drove and shaped it. On a broader level, there is also the question of why it was eukaryotes to which the path towards greater and greater complexification was open, in contrast to prokaryotes, which never progressed beyond very simple multicellularity. To be noted, it is perhaps not necessarily warranted to claim that prokaryotes are incapable of evolving complex multicellularity, as only a very small number of groups have done so in eukaryotes [96]. Thus is therefore a conclusion based on an extremely limited set of examples. Still, it is a common assumption shared by many, and it can be accepted for the sake of the argument.

These are questions that may turn out to be in principle unanswerable, as they involve speculations regarding what did not happen in prokaryotes and why that might have been. They will nevertheless persist in the minds of researchers interested in the subject, even more so given that examples of many of the hallmark features of the eukaryote cell can in fact be found in some form in at least some prokaryotes.

Eukaryotes are distinguished from prokaryotes by numerous traits, among them the enclosure of genetic material in a double membraned nuclear compartment and its organization into multiple linear chromosomes consisting of histone-based nucleosomal chromatin; the existence of an intracellular membrane network, with sophisticated mechanisms for internal membrane transport between compartments; the ancestral presence of mitochondria (some-

times reduced to mitochondria-related organelles, or MROs, and lost completely on only one known occasion [7]); the presence of a complex cytoskeleton; mitotic cell division; generally larger cellular size; and an overall higher tendency to contain/tolerate endosymbionts. However, examples of many of these features can be found in prokaryotes too [8], e.g. cells with internal membrane-bound compartments, in some cases even enveloping the nucleoid [9–13] (though not necessarily completely enclosing it, as the eukaryotic nucleus does), giant size [14–17], multiple chromosomes/replicons [18], linear chromosomes [19], presence of endosymbionts [20], and others. Yet these are generally isolated examples, and the full complement of eukaryotic features is indeed unique to eukaryotes.

The chimeric origin of eukaryotes

Given the long history of heated debates surrounding eukaryogenesis and its very distant timing (ensuring that little solid evidence is available to constrain theories), it is no surprise that there is an enormous variety of hypotheses regarding how it happened. Our goal is not to comprehensively review all of them, but to specifically examine the role of the mitochondrion in the process. We will therefore primarily focus on the consensus view of how eukaryotes evolved that has recently emerged [21–31], together with some of the prior hypotheses most relevant to the discussion.

It is now almost universally accepted that the mitochondrion originated as a result of an endosymbiotic event involving a member of the α-proteobacterial clade. The first endosymbiotic proposals were advanced in the early 20th century by Mereschkowski regarding the origin of chloroplasts [32, 33]. That the mitochondrion also originated that way became a popular view in the 1960s, thanks to the endosymbiotic theory advanced by Margulis [34]. The relationship between α-proteobacteria and mitochondria was later confirmed by sequencing studies [35–37], and has been robustly reproduced in subsequent phylogenomic studies [38–40], although the exact positioning of mitochondria relative to the extant α-proteobacterial diversity is still a subject of debate [41–44].

The next major step in piecing together the puzzle of eukaryote origins came from the discovery by Woese, during pioneering rRNA sequencing efforts in the 1970s, that archaea constitute an independent domain of cellular life [45]. Based on early phylogenetic analyses, a three-domain system was proposed, in which bacteria, archaea, and eukaryotes form distinct domains, with eukaryotes being more closely related to archaea than to bacteria [45, 46]. A relationship between archaea and eukaryotes was also supported by observations of commonalities between the information processing machineries in the two groups, e.g. RNA polymerases and other components of the transcriptional and translation apparatuses [47, 48], the presence of nucleosomal histone proteins in archaea but not in bacteria [49–52], and others.

In the 1980s, the eocyte hypothesis was proposed as an alternative of the three-domain system, initially based on phylogenetic analysis of ribosomal structure and later supported by studies of individual highly conserved genes [53–57]. This hypothesis has eukaryotes emerging from within archaea rather than being a sister lineage that diverged early from a common ancestor of both.

Although for a long time it was difficult to establish the exact phylogenetic relationship between the various diverse archaeal lineages and eukaryotes, the eocyte view was bolstered by early phylogenomic studies, once whole-genome sequences for a sufficiently large sampling of microbial diversity became available [22–24, 26, 27, 29–31], More and more aspects of cellular biology thought to be specific to eukaryotes have been also found in at least some archaeal lineages [58–60], suggesting a rather complex eukaryote ancestor. However, the relevant genes are dispersed around the archaeal tree and not all found in a single lineage [61, 62].

A major recent breakthrough was the discovery of first, the Lokiarchaeota archaeal lineage [63], and then, the broader Asgard archaea grouping [64–66] (combining Lokiarcheaota and the related Thor-, Odin-, Heimdallarchaeota and other clades). Some caveats are needed when discussing Asgard archaea. First, the interpretation of a direct relationship between them and eukaryotes has been challenged [67–69]. Second, until very recently they were not isolated, cultured or observed directly, but were instead only known from metagenomic sequence assemblies leaving uncertain their cellular organization. The latter issue was partly resolved shortly prior to the finishing of this text with the isolation and culturing of the first Asgard archaeon, Candidatus *Prometheoarchaeum syntrophicum* [70] (to be discussed in more detail further below). However, it remains true that we still have only one cultured Asgard archaeon, and that we do not know with certainty what exactly the actual ancestor of eukaryotes was like. Nevertheless, these archaeal lineages do appear to be the most eukaryote-like archaea based on the gene content of their genomes, with the presence of numerous genes homologous to those for components of the eukaryote membrane dynamics and trafficking machinery being of particular interest.

The emerging consensus view therefore is that eukaryotes arose from within the archaea. Still open, however, are (among many others) the following questions:

1. When in the process of eukaryogenesis did the mitochondrial endosymbiosis become established?

2. What was the level of cellular organization of the primordial host?

3. What evolutionary forces drove the evolution of the hallmarks of the eukaryotic cell, in particular its distinct complexity relative to prokaryotes?

4. What influence did the mitochondrion have on this process?

As the mitochondrion itself is one of many hallmark complex features of eukaryotes, there are three broad classes of possibilities for the relative timing of its acquisition in the context of eukaryogenesis (Figure 1):

1. The host was already a fully fledged but amitochondriate eukaryotic cell (i.e. with a conventional eukaryotic nucleus, endomembrane system and cytoskeleton) that then acquired an α-proteobacterial endosymbiont, which in turn evolved into the modern mitochondrion
2. The host was a complex archaeon, (i.e. one that had already developed many features of the eukaryotic endomembrane and cytoskeleton systems, but did not yet have a eukaryotic nucleus), possibly more complex than any known so far in some detail extant archaeal lineages.
3. The endosymbiotic event coincided with and possibly triggered the process of eukaryogenesis.

The first possibility, traditionally referred to as the "Archezoan" scenario, enjoyed widespread popularity in the 1980s and 1990s [74, 75] based on the observation that numerous eukaryotes lack mitochondria and tended to group together as early diverging branches in early phylogenies. However, such topologies eventually proved to be long-branch attraction artifacts, while each of the examples of amitochondriate eukaryotes turned out be a case of secondary reduction of mitochondria into MROs, such as hydrogenosomes and mitosomes [24, 76–82]. No primitively amitochondriate eukaryotes are known (of note, recently a completely amitochondriate eukaryote, *Monocercomonoides* sp., was reported [7], but it is still clearly a product of secondary loss as it belongs to the oxymonads and not to a new early branching eukaryote lineage). Thus the Archezoan scenario has been largely rejected, and it is inferred that the Last Eukaryotic Common Ancestor (LECA) had fully functioning mitochondria.

Deciding between the remaining two classes of hypotheses is considerably more challenging, at least in part because the concept of a "complex archaeon" is not strictly defined. There have been proposals that the archaeal host was already capable of phagocytosis ("phagocytosing archaeon theory" [83], and phylogenomic analyses have been published suggesting a late acquisition of the mitochondrion [84]. However, the latter have been challenged on technical grounds [85, 86], and the former hypothesis seems unlikely to be true. The gene content of Asgard archaea genomes, and recent biochemical results showing the presence of profilins capable of regulating the polymerization of actin filaments [87], have provided support for eukaryote-like cytoskeleton and membrane dynamics being present in these organisms. However, bona fide phagocytosis might have evolved independently in eukaryotes on multiple occasions (based on the observation that very few of the relevant components are conserved across all eukaryotes, unlike what is observed for most ancestral eukaryotic features [88]). It is also likely that for phagocytosis to evolve a functioning mitochondrion must have been present in the cell (elaborated in more detail below). Such a view is supported, first, by recently published analyses of the Asgard archaea genomes from the perspective of the phagocytosis machinery as found in various eukaryotes, which predict that Asgard archaea are not capable of phagocytosis [89], and second and most importantly, by direct observations of *Prometheoarchaeum* cells, which have long membrane protrusions but no endomembrane system [70].

Undoubtedly, future direct studies of the cell biology of a wider sampling of the Asgard archaea diversity, plus the possible finding of even more eukaryote-like lineages will shed more light on the state of the eukaryotic ancestor.

It is still unlikely, however, that the archaeal host that gave rise to protoeukaryotes approached the ground state of complexity for eukaryotes. LECA seems to already have possessed a long list of complex features shared between all eukaryotes [90]. While precursors of some of these are found in some form in archaea, their totality and complexity is not approached in the known archeal diversity and likely developed between the First Eukaryotic Common Ancestor (FECA) and LECA. We have to then answer the question how and why these features, as well as additional complexities restricted to certain branches of the eukaryotes, such as massive increases in genome size and gene number as well as multicellularity, arose. Our remaining discussion will focus on the role that the endosymbiont played in the process of eukaryogenesis and the evolution of eukaryote complexity.

Energy and mitochondria

What prompted us to write this review is the argument that the bioenergetic influence of the mitochondrion drove the evolution of complexity in eukaryotes and the debate that has resulted from it [8, 91–105]. Both much confusion in need of clarification and an unnecessary amount of vitriol unrelated to the purely scientific issues have been generated in the resulting exchanges. It is the scientific questions that we want to focus on here.

The modern mitochondrion (when not secondarily reduced) functions as a dedicated cellular organelle in which aerobic respiration occurs (among some additional functions). Because aerobic respiration provides much more energy to the cell than anaerobic alternatives, the idea that the energetic boost provided by the mitochondrion enabled the later complexification of eukaryotes is appealing on the surface. However, the proper comparison is not between aerobic eukaryotes and anaerobic prokaryotes, but between aerobic eukaryotes and aerobic prokaryotes, and quantitative support of any evolutionary scenario is much preferable to verbal arguments.

A semi-quantitative version of the position that the bioenergetic effects of the mitochondrion were decisive for the emergence of eukaryote complexity was formulated by Lane and Martin in 2010 [8, 91, 92] (hereafter referred to as "power per gene" hypothesis). Underlying Lane and Martin's argument is the concept of the amount of power that

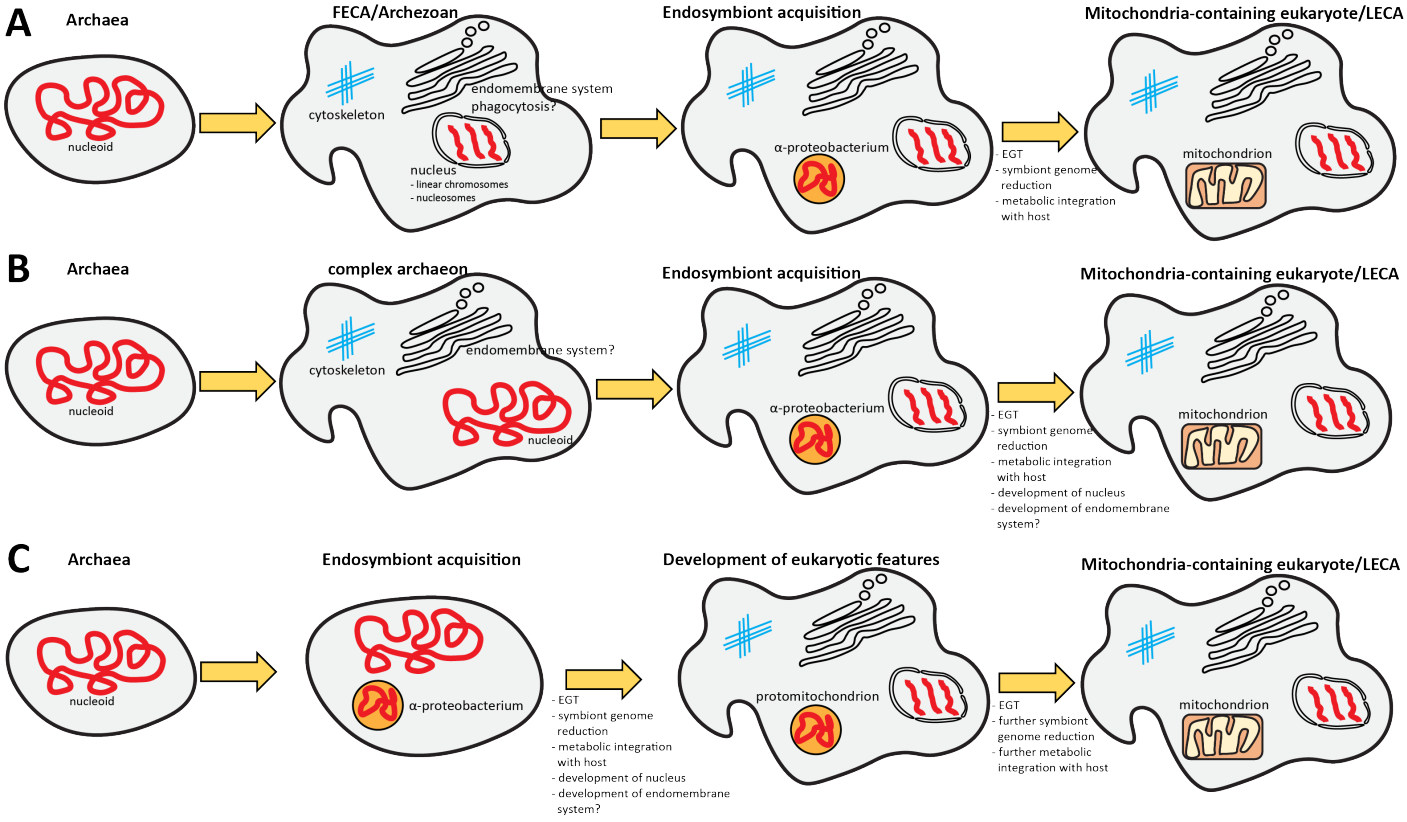

Figure 1: Different scenarios for eukaryogenesis and the relative placement and timing of symbiont acquisition, symbiont reduction, and host complexification. (A) Archezoan scenario. An archaeal ancestor develops into an organism with most typical eukaryotic features, with the exception of the presence of mitochondria. An α-proteobacterial endosymbiont is acquired subsequently. As a result of endosymbiotic gene transfer (EGT), endosymbiont genome reduction, and metabolic integration with the host, the α-proteobacterial endosymbiont/protomitochondrion eventually turns into the highly reduced mitochondrion of the Last Eukaryotic Common Ancestor (LECA). (B) The complex archaeal ancestor scenario. An archaeal ancestor develops certain eukaryote-like features, i.e. some combination of a complex cytoskeleton and an endomembrane system but no true nucleus. An α-proteobacterial endosymbiont is acquired. EGT, endosymbiont genome reduction, and metabolic integration with the host follow, resulting in the highly reduced mitochondrion of the Last Eukaryotic Common Ancestor (LECA). In parallel, the nucleus evolves, possibly triggered by the introduction of the endosymbiont [71]. (C) The protomitochondrion as the trigger of eukaryogenesis. The α-proteobacterial endosymbiont is acquired by a relatively simple archaeal host. Eukaryotic features develop in parallel with the reduction of the endosymbiont into the LECA-state mitochondrion and as a direct consequence of its presence. Note that it is possible that additional partners (not necessarily endosymbiotic) might have been involved in a syntrophic relationship in the process of establishing the eventual mitochondrial endosymbiosis [72, 73]; these are not shown as they are currently hypothetical, and only the mitochondrion survived as a distinct entity in eukaryotes. .

a cell generates for each nuclear gene (thus the "power per gene" label). The claim is that the existence of numerous organelles entirely dedicated to energy production increases the energy availability per gene, enabling expansion of genome size and gene content and of organizational complexity [92]:

> The situation is more pronounced in terms of gene number. An average bacterial genome (such as *E. coli*) contains nearly 5,000 genes, compared with some 20,000 in an average protist, such as *Euglena* ranging up to 40,000 in *Paramecium* [62,63,70,71]. At a metabolic rate of 0.49 pW per cell, a bacterium with 5000 genes has only 0.1 fW per gene. Smaller bacteria, with around 2,500 genes, have a power of 0.2 fW per gene. Larger bacteria, with around 10,000 genes are surely close to a lower functional limit, with a power of just 0.05 fW per gene. In contrast, at a power of 2,286 pW per cell and 20,000 genes, an average protist has about 115 fW per gene, over 1000-fold more energy per gene than an average bacterium, and more than 2000-fold more than

a large bacterium.

A prominent part of Lane and Martin's argument is that the mitochondrial genome is highly reduced and compact. The ancestral α-proteobacterial genome most likely contained a few thousand genes packed into several megabases of genomic space, as its modern relatives do. While some plant mitochondria (e.g. those of the *Silene* genus) have secondarily ballooned to multimegabase sizes [106], most mitochondrial genomes are only a few tens of kilobases long, and they all contain only a few dozen protein-coding genes [107]; in almost all cases, those genes are a subset of the ∼66 genes found in the mitochondrial genomes of the jakobids [108], which are the least reduced in terms of gene content known so far [109]. This extreme reduction is due to a combination of gene loss and endosymbiotic gene transfer (EGT) to the nuclear genome. The question of why mitochondrial genomes persist at all rather than having been transferred to the nucleus altogether has long been pondered. Given that the retained genes are mostly components of the redox systems in the mitochondrial inner membrane, the best explanation proposed seems to be that direct and rapid regulation of the expression of these genes is necessary ("colocation for redox regulation of gene expression", or "CoRR" hypothesis; [110]), necessitating the retention of a mitochondrial genome physically close to the location of electron transport chains (ETCs).

Lane and Martin suggest that it is not possible for a prokaryote cell to develop an analogous system of a dedicated, membrane-separated organelle with its own small compact genome "servicing" its ETCs. Of course, there are readily available counterexamples – the thylakoids of cyanobacteria are a membrane-bound ETC-containing organelle, and other prokaryotes also have developed analogous structures (e.g. the chromatophores of *Rhodobacter sphaeroides* [13]). The key difference between these structures and mitochondria, according to Lane and Martin, is that only mitochondria contain their own genomes. They then proceed to propose that the only path towards increased size for a prokaryote is whole-genome polyploidy as high-copy number plasmids "servicing" the ETC have never been observed and are claimed to be an impossibility. However, they argue that this does not result in high energy "availability per gene" [92], and increased "power per gene" is proposed as being necessary for increases in organizational complexity:

> Bacteria and archaea respire over their plasma membrane and thus are subject to surface-area-to-volume constraints. Increasing their linear dimension 25-fold increases their surface area 625-fold, and volume 15,000-fold. ATP synthesis could therefore increase 625-fold, but such an increase would require a 625-fold increase in the number of respiratory proteins, ATP synthase enzymes, and all other molecular machinery needed to transcribe and translate the genes. Transcription from a single bacterial genome could hardly be increased 625-fold. We must then scale up the number of genomes accordingly. To a superficial first approximation, scaling up ATP synthesis 625-fold would require haploid genome number to increase by 625-fold. Energy per gene would remain unchanged. If we take internal volume into consideration, the same principles apply. Protein synthesis could not increase 15,000-fold from a single genome, but if the number of haploid genomes were increased 15,000-fold, energy per gene would fall by 25-fold. So scaling up a bacterium to mean eukaryotic volume would cut energy per gene by $5000 \times 25 = 125{,}000$-fold [92]. All else being equal, a eukaryotic-sized bacterium should have about five orders of magnitude less energy per gene than the eukaryote [91, 92].

The Lane and Martin bioenergetic argument has become quite popular in recent years (the original publication has been cited ≥600 times, and we have seen the overall idea referred to as authoritative in numerous papers). However, evolutionary hypotheses should ideally meet higher epistemic standards than the mere statement of an argument in order to be widely accepted as true. A plausible description of the evolutionary *process* and the evolutionary forces in operation needs to be provided.

The "power per gene" hypothesis does not meet these criteria as it is difficult to connect to any known mechanisms for how evolution actually occurs. But another minimum requirement of a set of evolutionary explanations for how different aspects of cell biology evolved is internal consistency, and, as will be seen from what follows, the "power per gene" hypothesis fails that test too, as it is not consistent with several hypotheses concerning other aspects of eukaryogenesis that were either advanced by Lane and Martin themselves or are fairly well established at this point more generally.

In order to demonstrate these inconsistencies, we will treat the "power per gene" and several other hypothesis as much more reasonable than they actually are. These cases of assuming the validity of a proposal that we disagree with for the sake of the argument will be explicitly pointed out.

In 2015, as part of a larger endeavor aiming at understanding cellular bioenergetics and its impact on evolution, we carried out an analysis of the biochemical costs of replicating, transcribing, and translating genes [94]. We made several observations that relate to the "power per gene" hypothesis as stated by Lane and Martin. First, no discontinuity between prokaryotes and single-celled eukaryotes is observed in the metabolic scaling of energy requirements and cell size, meaning that eukaryotes are not endowed with greater bioenergetic capacity than equivalently sized aerobic prokaryotes would be. Second, the evolution of expanded genomes in eukaryotes can be readily explained simply as a consequence of their larger cell size and lower effective population size (N_e), without having to invoke any bioenergetic impacts by the mitochondrion [112].

To briefly summarize, the replication, transcription, and

translation of a gene impose metabolic costs on the cell, which are a function of gene and protein length and expression levels. This cost results in a negative baseline fitness effect $s_c = s_{DNA} + s_{RNA} + s_{PRO}$ of magnitude on the order of the ratio between the absolute cost of maintenance of the gene and the total energy budget of the cell. Then the total fitness effect of the gene is $s_n = s_a - s_c$, where s_a is the direct selective advantage conferred by the functionality provided by the gene. In order for the gene to be actively selected for or against, s_n has to exceed the drift threshold ($1/N_e$ for haploid and $1/2N_e$ for diploid organisms). After calculating the biochemical costs of replication, transcription and translation, accounting for the total energy budget and the total abundance of transcripts and proteins in different types of cells, it turns out that the fractional metabolic cost of genes decreases with increasing cell size, and, combined with the well-known fact that physically larger organisms tend to have lower effective population sizes, it appears that DNA is under selection at the metabolic level in prokaryotes, but is largely free to accumulate in large-sized eukaryotic cells. We thus concluded that a bioenergetic boost triggered by the presence of a mitochondrion in ancient eukaryotes is not necessary to explain their increased complexity.

In a later study [101], using similar approaches, we examined in more detail the energetic and biosynthetic capacity of prokaryotes and eukaryotes as well as the cost of building cellular membranes. We found that, contrary to expectations, eukaryotes do not have a higher ATP production capability than prokaryotes when normalized for cell size. The ATP synthesis capacity of a hypothetical cell in which all ETC complexes reside in the cell membrane is not substantially different from that of a cell of comparable size in which ETCs are located inside mitochondria. We also did not find a discontinuity in the scaling with cell size of the number of ribosomes per cell. Finally, we found that the addition of internal membranes imposes significant metabolic costs on cells due to the high ATP cost of synthesizing lipids.

These two studies have been commented on in a negative light by Martin, Lane and colleagues on numerous occasions [98–100, 102, 103, 105], featuring some rather incorrect statements. For example, in [98] we read that:

> For example, population geneticists contend that mitochondria had nothing whatsoever to do with eukaryote origin (Lynch and Marinov 2015).

Another excerpt states [103]:

> In 2016, many if not most authors publishing on the topic of eukaryogenesis still see no crucial role either for symbiosis or for mitochondria at eukaryote origin (Gray, 2014; Keeling et al., 2015; Booth and Doolittle, 2015; Forterre, 2011; Baum, 2015; Dacks et al., 2016; Lynch and Marinov, 2015, 2016; Koonin, 2015; Martijn and Ettema, 2013). For some, "luck" is preferable to endosymbiosis as an evolutionary mechanism (Keeling et al., 2015; Booth and Doolittle, 2015).

Finally, some quotes from [100]:

> The energetic and symbiogenic argument that mitochondria really made a difference at eukaryote origin (54) prompted Booth and Doolittle (15) to argue from a philosophical standpoint that mitochondria had no impact on evolution, to which an exchange (99, 100) was published. It also prompted Lynch and Marinov to argue from a population genetic standpoint that mitochondria had no impact on eukaryote evolution and that "*variation in the power of random genetic drift has played a central role in the historical diversification of genome and possibly cellular architecture across the tree of life*" (16)
>
> [...]
>
> Lynch usually argues that nothing other than nonadaptive processes or N_e impacted any aspect of the evolutionary process, in any organism group.
>
> [...]
>
> Lynch and Marinov's (16, 17, 102) proposition is that mitochondria are irrelevant to eukaryote origin and that nothing more than point mutations and the power of population bottlenecks are needed to transform a prokaryote into a eukaryote.

As a general comment, the argument that "nothing other than nonadaptive processes or N_e impacted any aspect of the evolutionary process" has never been made; the point has always been that all evolutionary forces, one of which is genetic drift, have to be taken into account in quantitative manner when analyzing evolutionary processes, that neutral evolution is the null hypothesis to be rejected, and that such analyses reveal that selection is invoked as the default explanation in too many cases where it need not be [111–113].

Specifically with respect to mitochondria and eukaryogenesis, the statement (implied from Lane and Martin's writings) that "the mitochondrion had nothing to do with it" is also widely inaccurate. It is the much more specific claim that the bioenergetic impact of the mitochondrion was crucial for the origin and complexity of eukaryotes that is in doubt. This conclusion was initially based on the two studies referred to above, but it is also supported by several other lines of reasoning, to be expanded on below.

We note that for an evolutionary explanation to be a solid and reliable one, it needs to provide an account not just of the effects that the trait in question has today but also of the *process* of evolution and of the impact the trait had during that process. A trait might be having a certain effect at a given point in time but it need not be the case that this effect played a causal selective role its establishment up to that point.

The mitochondrion is indeed the power station of modern eukaryotic cells. It also does indeed have the *apparent*

effect of providing a large "power per (nuclear) gene" at the cost of a relatively small investment in genetic real estate. However, as we will show, even if we accept the premise that "power per gene" is an evolutionarily meaningful concept (which we do not view as correct), that such an apparent effect can be observed in modern eukaryotes does not in any way imply that increased "power per gene" played any causal role in the evolution of what we recognize as core eukaryotic features or that the (proto)mitochondrion had such effects when eukaryotes were originally evolving in the very distant geologic past.

The progression and impact of endosymbiosis

The first important consideration to take into account is that endosymbiosis is a complex process, and that how exactly the establishment of the mitochondrial endosymbiont influenced the origin of eukaryotes can only be understood within the context of that process. After (and possibly even during) initial establishment of endosymbiosis, the following key interrelated aspects of that process need to unfold for an endosymbiont to become an "organelle":

1. Biochemical integration between host and endosymbiont, involving mechanisms for transport of metabolites across membranes, has to evolve.
2. Extensive endosymbiotic gene transfer (EGT) has to happen from the endosymbiont genome to the host genome.
3. Mechanisms for targeting and import of proteins into the endosymbiont have to develop.

The two key endosymbiotic events in the history of the planet are the acquisition of the mitochondrion by the protoeukaryote and the acquisition of a cyanobacterial endosymbiont, later turned into a plastid, by the ancestor of the modern Archeplastida lineage. Both of these events happened in the very distant past and we have very scarce information on which to rely on for reconstructing the details of the process. However, we do have plenty of examples of more recent endosymbiotic events, from which we can obtain an idea of how analogous processes can unfold.

The original archeplastid cyanobacterial symbiosis took a cyanobacterial genome containing several thousand genes down to ≤ 250 genes (in red algae) and even fewer ≤ 100 genes in green algae and land plants. While that original event happened more than a billion years ago, unlike the situation with the mitochondrion, whose origin is, as far as we can tell, a singular event, we do have a recent example of a second primary endosymbiotic event between an eukaryote and a cyanobacterium, that of the rhizarian *Paulinella*. The chromatophore of *Paulinella chromatophora* is estimated to have been acquired 60–100 Mya [114]. In that time, its genome has been reduced to a size of 1.02 Mbp containing 867 protein coding genes [115], and up to 450 now nuclear encoded proteins are trafficked back into the chromatophore [116, 117].

Another well-known example of what is often referred to (correctly or not) as "organelle-formation-in-progress" are the spheroid bodies in rhopalodiacean diatoms, also of cyanobacterial origin. Notably, this is a very recent event, dating to the Miocene some 12 Mya [118]. The sequenced spheroid body genome of *Epithemia turgida* is 2.79 Mbp in size and contains 1,720 protein coding genes [119], while that of *Rhopalodia gibberula* is 3.02 Mbp in size and contains 1,671 protein coding genes [120].

A wide variety of bacterial endosymbionts are found in insects, usually exhibiting a great extent of genome reduction. The smallest endosymbiont genome known to date is that of *Nasuia deltocephalinicola*, found in the phloem-feeding insect *Macrosteles quadrilineatus*, which is only 112 kbp in size [121]. It is thought to descend from a colonization event that happened $\sim$250 Mya or earlier [122].

These and other examples show that while the speed of genome reduction and EGT probably varies between groups, it still takes on the order of hundreds of millions of years for an endosymbiont genome to be reduced to less than 100 Kb. Of course, it is in principle possible that this happened more quickly in the case of the protoeukaryote, but a counter argument can be made that modern eukaryotes are all descendants of organisms that have already at least once successfully managed to integrate an endosymbiont. Thus there are numerous components of the protein targeting and compound import/export machineries that could be reused/adapted in the case of further endosymbiotic events, a luxury that the protoeukaryote did not have. In any case, it is highly doubtful that the protomitochondrion had its genome reduced extremely quickly on a geological time scale. The fact that the jakobids, which place a lower bound on the mitogenome complexity of LECA, have a fairly large complement of genes also argues against such a conclusion. We have no idea how long it took to go from FECA (the first eukaryotic common ancestor) to LECA, but it is likely that it was a substantial amount of time given the sheer number of universal complex eukaryotic features that had to evolve between FECA and LECA. Also, *Prometheoarchaeum syntrophicum* is extremely slow growing[70]; if we are to assume that this is a feature shared with the actual eukaryotic ancestor, it lends further support of the view of eukaryogenesis unfolding over a lengthy geologic (if not necessarily generational) period.

This has some significant consequences for the Lane and Martin hypothesis, a key feature of which is that the small size of the mitochondrial genome and its complete dedication to energy production provides a much larger "power per nuclear" gene relative to what prokaryotes have to live with. The early endosymbiont was not reduced, and it most likely took a long time for it to be diminished. Therefore, long after the endosymbiont was acquired, there was no significant "power per nuclear gene" effect to speak of as the endosymbiont did not have a small genome entirely ded-

icated to energy production (and we would like to again stress that such an effect is in our view not causal with respect to eukaryogenesis).

The impact of the (proto)mitochondrion versus the role of mature mitochondria

The nature of the initial endosymbiotic relationship

An even more serious problem is that the original endosymbiont quite likely was not even a mitochondrion to begin with but had a rather different kind of relationship with the host. The diversity of proposals on how and why that relationship was established is large and difficult to summarize in a short amount of space [123]. We will restrict ourselves to pointing out that the accurate estimation of the timing of the establishment of endosymbiosis and good understanding of the geochemical conditions on Earth at that time and thereafter are key to answering this question. The original endosymbiotic theory paper by Margulis [34] placed much emphasis on the role of oxygen, envisioning a symbiotic relationship between an anaerobic host and an aerobic endosymbiont in the context of the Great Oxidation Event (GOE; the appearance of oxygen in Earth's atmosphere as a result of the activity of photosynthetic cyanobacteria ∼2.5 Gya). More recent papers have also developed the idea that the acquisition of the mitochondrion was beneficial because it conferred oxygen tolerance to the host [78, 124, 125]. However, the fossil record [126, 127] and molecular clocks [128] do not support a timing of eukaryogenesis immediately coinciding with the GOE, and it is also significant that for the rest of the Precambrian oxygen concentrations remained low ($\leq$1%) compared to conditions in the Phanerozoic. To what extent there was a causal relationship between the GOE and eukaryogenesis is not clear. In addition, the oxygen tolerance hypothesis has serious issues given that mitochondria themselves are a major source of reactive oxygen species in modern cells [8, 129].

A popular scenario for how the endosymbiosis was established is the hydrogen hypothesis first stated by Martin and Müller in 1998 [130]. According to the hydrogen hypothesis, the archaeal host had a hydrogen-dependent methanogenic metabolism utilizing hydrogen and CO_2, while the endosymbiont was a facultative anaerobe capable of generating hydrogen and CO_2 as products of anaerobic respiration. If we are to assume the validity of the "power per gene" hypothesis, an obvious problem is immediately apparent – under the hydrogen hypothesis the endosymbiont did not initially function as a mitochondrion but had a very different metabolic relationship with the host. While the host extracted some metabolic benefits from the endosymbiont, the endosymbiont was not yet dedicated to oxidative phosphorylation, with its very high energy yields. Thus the "power per gene" hypothesis does not apply for the period between the acquisition of the endosymbiont and its conversion into a true mitochondrion, and, as outlined above, it also does not hold until the combination of EGT and loss of non-essential genes sufficiently reduced the size of the endosymbiont genome.

This is a general feature of many proposals for how the endosymbiosis was originally established metabolically. To be noted, the analysis of available Asgard archaea and the direct study of *Prometheoarchaeum syntrophicum* cast strong doubt on the hydrogen hypothesis [72, 73, 131, 132]. *Prometheoarchaeum* itself is certainly incompatible with it. It is an anaerobe whose metabolism appears to be based on catabolizing amino acids and growing syntrophically together with hydrogen- and formate-utilizing bacteria. In other words, it is very much not the hydrogen-dependent methanogen envisioned in the hydrogen hypothesis. Admittedly, Asgard archaea as a whole are very diverse metabolically and the possibility that some of them utilize hydrogen cannot yet be completely rejected [72, 73], but it is also clearly not the front runner among competing proposals at present.

In any case, whatever the exact initial metabolic relationship was, the process of evolving an actual mitochondrion could well have taken tens and hundreds of millions of years. But during all that time the endosymbiont impacted the biology of the host in numerous ways unrelated to any bioenergetic contribution it might have made to the host or to the yet-to-evolve mitochondrion (e.g. EGT, reworking of the membrane system of the cell, and others). As we will show shortly, many of the proposed scenarios for how some of the specific complex aspects of eukaryotic biology evolved feature exactly such impacts as key driving forces. We note that we do not necessarily agree with all of the proposals listed below, but they nevertheless illustrate how hallmark eukaryotic traits could have evolved without having to invoke the bioenergetic hypothesis as formulated by Lane and Martin (we also again stress that many of these hypotheses have been originally advanced by Martin and collaborators themselves).

The origin of the endomembrane system

While membrane invaginations and even membrane-bound compartments are known in prokaryotes, the complex endomembrane system consisting of an endoplasmic reticulum, a Golgi apparatus, and a variety of endomembrane vesicles (lysosomes, peroxisomes, autophagosomes, and others) is, as far as we know, unique to eukaryotes. How did it evolve? As with most other aspects of eukaryogenesis, there are numerous scenarios, and we cannot be certain about the validity of any of them. Here we will briefly consider the one proposed by Gould, Gard and Martin [133].

It is based on the observation that prokaryotes produce extracellular membrane vesicles called outer membrane vesicles, or OMVs [134–136]. Both bacteria and archaea have been observed to generate OMVs but the former are key to the OMV hypothesis as is the observation that mitochondrial membranes also generate vesicles, called mitochondria-derived vesicles, or MDVs [136–139]. In ad-

dition, a well-known general dichotomy between bacteria and archaea concerns the composition of their membranes. Archaeal lipids contain isoprenoid hydrocarbon side chains and glycerol-1-phosphate (G1P), connected via an ether linkage. In contrast, bacterial lipids are based on an ester linkage between glycerol-3-phosphate (G3P) and fatty acids. Eukaryote membranes are all of the bacterial type. Thus, if the eukaryote ancestor was an archaeon the transition from archaeal to bacterial lipids has to be explained (although it is to be noted that a recent phylogenomic analyses suggests that Lokiarchaeota together with some Euryarchaeota groups lack G1P synthesis capacity and instead have the capacity to synthesize G3P-based "chimeric lipids" [140]). The OMV hypothesis ties these observations together by postulating that an outward flux of OMVs from the bacterial endosymbiont gave rise to the endomembrane system in protoeukaryotes.

We are not convinced of the validity of this hypothesis; autogenous models for the origin of the endomembrane system, according to which the endomembrane system evolved from inward invaginations of the plasma membrane (e.g. [141]) may turn out to be more in line with what actually happened, but the OMV hypothesis is Martin's own latest model [133]. We just note that the process as outlined could have well unfolded and proceeded quite far along the path towards the modern eukaryotic state long before the α-proteobacterial endosymbiont was reduced to a true mitochondrion so that the bioenergetic hypothesis could even apply. However, the endosymbiosis event itself obviously played a major causative role (under the OMV hypothesis).

The evolution of phagocytosis

Phagocytosis has been central to eukaryogenesis hypotheses for much of the history of the subject, as many have assumed that a phagocytosing host is a prerequisite for the establishment of endosymbiosis [142, 143] (the "failed predation" hypothesis). There are several problems with such a view. First, bacterial endosymbionts inside other bacteria [20, 144–146] and even inside mitochondria [147] are known, meaning that phagocytosis by the host is not necessary for endosymbiosis. Also, on a broader speculative level, it might be argued that a rare and improbable event fits better with the singular nature of the prokaryote-to-eukaryote evolutionary transition than a common and easy to accomplish step. Endosymbioses have been established on innumerable occasions in eukaryotes, precisely because of their phagocytotic capabilities, and if the ancestral archaeon/protoeukaryote was similarly well equipped to establish endosymboses, multiple cases of bacterial endosymbionts within archaea would be expected to have evolved. However, this is not what is observed, at least based on what we currently know about archaea, and, of course, we also know of only one case of evolution to a complex eukaryote-like state.

Second, and more important, phagocytosis, is most likely incompatible with a cellular membrane that also contains ETCs. The reason for this is that lysosomes in modern eukaryotes maintain a highly acidic pH, which is necessary for the proper functioning of the digestive enzymes contained within. But ATP synthases generate ATP by transporting protons along the electrochemical gradient generated by electron transport, and if ATP synthases were to end up in the lysosomal membrane as a result of a phagocytosis event, the acidic pH in the lysosomal lumen would be very quickly dissipated. A similar line of reasoning was advanced by Martin and colleagues in 2017 [148].

Interestingly, this argument for why the archaeal host could not have been phagocytosing can be turned around to explain why eukaryotes did evolve phagocytosis. Once the endosymbiont/protomitochondrion (it need not have fully evolved into the LECA mitochondrial state) had taken over most ATP production in the cell, there was no longer a need to maintain ETCs in the external cell membrane, and the cell was now free to evolve phagocytosis.

This would have been a major influence that the endosymbiont had on the (proto)eukaryote cell but, again, it was at best an indirect effect of its bioenergetic impact. While predation is known in prokaryotes (e.g. *Bdellovibrio* [149]), it involves the predatory cell invading the prey. In contrast, phagocytosis is of a fundamentally different nature, with the predator engulfing the prey. Its appearance likely had significant ramifications for the subsequent evolution of eukaryotes, e.g. by enabling increased cell size and driving the evolution of defense mechanisms, such as rigid and armored cell walls, as a means of protection against predation.

The origin of spliceosomal introns

A radical difference between the organization of eukaryotic and prokaryotic genes is the presence of spliceosomal introns in the former. Some eukaryotes have very few introns, but multiple phylogenetic reconstructions of the evolution of shared intron positions across the whole eukaryotic tree of life unequivocally point to the conclusion that LECA had an intron-dense genome [150–153]. Spliceosomal introns are spliced by the spliceosome, a rather complex biochemical machine, consisting of $\sim$100 proteins plus several small nuclear RNAs (snRNAs), a state that too appears to be ancestral to extant eukaryotes [154].

The origin of spliceosomal introns is one of the pieces of the eukaryogenesis puzzle that we are most certain about. The core of the spliceosome is actually a ribozyme as it is the snRNAs that catalyze the splicing reaction [155]. The overall structure that the snRNAs adopt is highly reminiscent of the conformation of the self-splicing Group II introns [156], which are found in bacteria and the organelles of various eukaryotes [157]. Thus, it is highly likely that the spliceosome evolved from such self-splicing Group II introns. Remarkably, Group II introns appear to be ancestrally absent from archaeal genomes, suporting the view that the endosymbiont giving rise to the mitochondrion

played a central role in the evolution of splicing and the spliceosome [71, 158]. Under that view, the archaeal host had no Group II introns, but once the endosymbiosis was established, Group II introns present in the α-proteobacterial endosymbiont began colonizing the host genome through EGT. The gradual degeneration through mutations of these self-splicing introns left them incapable of splicing in *cis* and led to the evolution of mechanisms to splice them in *trans*. The snRNAs evolved from fragments of Group II introns to take over the role of catalyzing the splicing reaction, and with time the protein components of the spliceosome were added around them.

Given that EGT could have started, and probably indeed did start as soon as the endosymbiont was acquired, this example of a major impact that it had on the host is also almost completely independent from any bioenergetic effect it might have also had once it became a mitochondrion.

The origin of the nucleus

This line of thinking was extended even further Martin and Koonin about a decade ago [71, 158], when they proposed that the invasion by Group II introns is what drove the evolution of the nucleus. The argument goes as follows. In prokaryotes, transcription and translation are directly coupled in time and space because there is no physical separation between the two processes, while in eukaryotes the nuclear membrane compartmentalizes them. However, removal of spliceosomal introns is relatively slow, and therefore abundant spliceosomal introns would impose a major fitness cost in terms of mistranslated proteins in the absence of decoupling between transcription and translation. This is posited to have provided selective pressure for the evolution of the nucleus.

There are significant question marks regarding how an organism so fundamentally flawed might have survived at all for sufficiently long to evolve a nuclear membrane, but this is beside the point. The important thing to note here is that when it comes to the specifics of the models that exist in the literature for how such a major discontinuity between prokaryotes and eukaryotes as the nucleus could have evolved, the bioenergetic effect of the mitochondrion is once again absent from the picture, and the process as proposed could have unfolded entirely within the time period from endosymbiont acquisition to its conversion into a mitochondrion with a reduced genome.

The evolution of sex

EGT has also been postulated to have driven the evolution of meiotic sex [92], another feature that seems to have been present in LECA [159]. The hypothesis is that the invasion of mitochondrial DNA into the host genome through EGT greatly elevated the rate at which genes are inactivated by deleterious mutations (for example, through the insertion of Group II introns with diminished self-splicing activity). The protoeukaryote is proposed to have offset these fitness effects by cell fusion and masking these mutations by intact copies of these genes from other cells (to be noted, many unicellular eukaryotes are haploid for most of their life cycles). This scenario was formulated in the context of a host cell that has been freed of restriction of genome size by a fully formed mitochondrion, but there is really little reason why it would not apply from the onset of endosymbiosis and EGT. In addition, as we noted before [94, 101], the mitochondrion did not in fact provide the cell with a quantum boost in energy availability compared to a prokaryote of similar size.

Another hypothesis for the origin of sex links the evolution of meiosis to the production of reactive oxygen species by the mitochondrion and the resulting DNA damage [125, 160–162]. Many components of the meiotic machinery indeed do trace their origin to more ancient DNA repair mechanisms, but we reserve judgment on the validity of this model too. Once again, we reserve judgement regarding the validity of these proposals. We also have to point out that sex is a prominent feature of the biology of prokaryotes, it is just not of the meiotic kind. Thus scenarios about the origin of meiotic sex need to carefully consider the timing of the key steps in the sequence of events leading to meiosis relative to the origin of the nuclear membrane and linear chromosomes in eukaryotes (which would have presumably precluded exchanges of genetic information through the mechanisms employed by prokaryotes).

In the context of the discussion of the "power per gene" hypothesis, it is important to note that while directly linked to the presence of a mitochondrion, the scenarios outlined above do not in fact invoke its bioenergetic impacts.

The complexity of the archaeal host and its implications

Finally, we come back to the fact that the recent genomic analyses of Asgard archaea [63–66, 163] as well as biochemical experiments using purified expressed proteins from their genomes [87] hint that these cells might possess fairly sophisticated cytoskeletons and membrane trafficking machineries. The cellular organization of the immediate eukaryote ancestor is not yet certain, despite the isolation of *Prometheoarchaeum*, which does not have an endomembrane system but does form membrane protrusions. After all, modern Asgard archaea are billions of years removed from the ancestor of eukaryotes, and we have directly observed only one representative of a diverse and ancient lineage. What is important to note in our context is that the more complex the ancestor turns out to be, the less likely the "power per gene" hypothesis is to be true, as all of that complexity would have developed before the mitochondrial endosymbiosis had even occurred.

Summary and conclusions

The status of the "power per gene" hypothesis

The key observations relevant to the bioenergetic hypothesis are the following:

1. Accounting for cell size, there is no discontinuity between prokaryotes and eukaryotes in terms of energy budget. Eukaryotes do have no more (and possibly less CITE NEW PAPER) available energy than scaled up prokaryotes with the same metabolism.

2. The mitochondrion does not appear to provide aerobic cells with a major advantage in terms of energy availability relative to comparatively-sized aerobic cells without mitochondria.

3. As a result of their greater organismal size and population genetic environment characterized by low N_e, the path towards passive genomic expansion is open to eukaryotes but not to prokaryotes [111].

4. Many of the hallmark features of eukaryotes could have evolved in the absence of a fully fledged mitochondrion, but rather as a result of the impacts of having an endosymbiont and the processes of EGT and exchange of membranes.

5. The endosymbiont likely had not become a true mitochondrion for a lengthy period of time following its establishment. Many of the key events of eukaryogenesis could have advanced quite far long before the proposed by Lane and Martin bioenergetic impacts of the mitochondrion were in effect.

We can further evaluate the "power per gene" hypothesis in the light of three broad eukaryogenesis scenarios outlined in Figure 1. First, it is obviously incompatible with the Archezoan hypothesis, as in that sequence of events all major eukaryotic features are already present prior to the endosymbiosis taking place. Importantly, the mitochondrion plays little role in the development of shared eukaryote complexity under that hypothesis. However, the Archezoan scenario appears to be false as far as we can tell.

Second, complex archaeon scenarios are not incompatible with the "power per gene" hypothesis, but they nevertheless do argue against it to an extent, the more strongly so the more complex the archaeal host and the later the acquisition of the mitochondria turn out to be. In such sequences of events non-trivial cellular complexity would have evolved even without a mitochondrion, and also possibly even under anaerobic conditions.

Third, even the scenario of eukaryogenesis coinciding entirely with the establishment of the mitochondrial endosymbiosis does not support the "power per gene" hypothesis, because, as discussed above, increases in complexity could have been triggered by the presence of the endosymbiont, independent of its bioenergetic effects and likely before they even took shape.

However, the mitochondrion still plays a critical role in the process under that scenario, as it is these impacts (impacts on the host's genome through EGT, possible impacts on the development of an endomembrane system, indirect impacts by enabling the evolution of features such as phagocytosis, and others) that drove much of the process of eukaryogenesis. Thus equating skepticism of the "power per gene" hypothesis with claiming that the mitochondrion had no role in the evolution of eukaryotes is a rather gross misrepresentation.

The passive emergence of eukaryotic complexity by genetic drift and mutational pressures

As has been demonstrated extensively previously [111–113, 164–166], many of the complex features of eukaryote genomic organization, such as the larger size of eukaryote genomes, the presence of introns, the accumulation of transposable elements and intergenic space, the baroque complexity of gene regulation in multicellular organisms, and others, most of them often thought of as adaptive improvements over the simple eukaryote state, can in fact be readily and much more parsimoniously explained by the inability of natural selection to get rid of slightly deleterious features in a population genetic environment characterized by low effective population size. Our bioenergetic analyses [94, 101] provided further support for that view, by identifying differences between the situation in prokaryotes and eukaryotes with respect to the fitness impact that the presence of additional DNA and its expression has on the cell.

The key variable that emerges from these analyses is cell size. Larger cells are more tolerant to accumulating excess DNA (cases of massive whole-genome polyploidization aside) because of the smaller fractional cost of its maintenance, and also because physical size is generally inversely correlated with N_e XX CITE LYNCH + ?, IN PRESS XX. Increased tolerance towards the addition of genes and the expansion of noncoding space can then lead to increases in the complexity of the gene repertoire of the cell and of the regulatory networks governing the expression of these genes. The latter, by virtue of the multiple alternative cell states they can generate, are probably a prerequisite for the emergence of complex multicellularity.

The reasons why the tolerance towards (or, alternatively, inability to purge) extra DNA due to the combination of large cellular size and low N_e is conductive to complexification are severalfold. At the gene level, complexity increases through the de novo origin of new genes from intergenic space, the acquisition of new genes through horizontal gene transfer, or, perhaps most importantly, through subfunctionalization or neofunctionalization of duplicate copies of existing genes. The maintenance of intergenic space imposes metabolic costs on the cell proportional to its length, which, as discussed above, are sufficiently large to be subject to selection against in prokaryotes, plus it represents an increase in the size of the mutational target in the cell, also proportional to its length and also potentially sufficiently

large to be subject to selection in lineages with very large N_e [112]. A similar line of reasoning applies to genes acquired through HGT, which is dependent on the interplay between the magnitude of their positive fitness effect to the cell and the negative fitness effect of the metabolic cost of maintaining and expressing them. Finally, theoretical analyses point to subfunctionalization becoming an increasingly more probable mechanism for the preservation of duplicate genes with decreasing population size [167, 168] (as a side, note, the establishment of endosymbionts and their conversion into organelles through gene loss and EGT can itself be thought of as a grand example of subfunctionalization).

At the level of gene regulation, complexification occurs through the expansions of the set of regulatory elements (REs) controlling the expression of each gene (such as enhancers and insulators). The typical human gene is associated with as many as a dozen distally located such elements [169], and it is often the case that different enhancers are utilized to drive the expression of a gene in different cell types and tissues. This is in marked contrast with prokaryotes and many unicellular eukaryotes where gene regulation is mostly promoter proximal. The proliferation of REs is of key importance to the evolution of new cell types, and in turn, of organismal complexity, as it allows the same set of genes to be used in different combinations in novel contexts. However, this is not necessarily an adaptive process – many of the intermediate steps that lead from simple to more complex gene regulatory networks involve non-adaptive mechanisms, as discussed in detail previously [166]. In short and in a most simplified form, addition of regulatory embellishments increases the size of the target for degenerative mutations, thus it has a negative fitness effect proportional to the mutation rate μ, and it also imposes a metabolic cost through its maintenance. Larger cells and population genetic environments characterized by lower N_e are therefore more conductive to increases in regulatory complexity. From then on, novel paths are open for gene evolution by ordinary descent with modification processes.

More generally, greatly increased complexity is inherently maladaptive as it results in more fragile systems, and is often "locked in place" though nonadaptive means rather than providing an obvious advantage [170, 171].

We can therefore best understand the evolution of eukaryotes and complex life on planet Earth on the broadest level not as a purely adaptive process (as usually presented) but as one also driven by an evolutionary ratchet towards increased complexity emerging passively from the cumulative incremental neutral fixation of genomic alterations in lineages with long-term population genetic environments permissive to it. In contrast, prokaryotes remain relatively simple and streamlined due to their small size and very high N_e, which efficiently eliminate the genomic alterations that might otherwise have led to significantly increased complexity from their populations. The process goes as follows: increases in cell size, decreases in N_e and changes in genomic organization in early eukaryotes make it easier to acquire and tolerate additional DNA. The emergence of features such as phagocytosis and a complex cytoskeletal organization might enable and further drive cell size increases. The increased complexity of gene composition and gene regulation eventually leads to the appearance of simple multicellularity, which in turn lowers N_e even further as multicellular organisms are typically physically larger than their unicellular relatives. The lowered N_e facilitates even further complexification of gene regulatory networks, enabling the evolution of yet larger and more complex organisms, which in turn have even lower N_e. This state need not be an inevitable product – the vast majority of eukaryote lineages are still unicellular [172] – but it is also true that such an evolutionary path is probably closed to populations of small-sized aerobic prokaryotes.

The mitochondrial endosymbiosis was a key event that set this process in motion. Yet it most likely did so through its numerous other impacts on its host rather than through an increase in the available "power per gene" as proposed by Lane and Martin.

Acknowledgments

XXX FUNDING SOURCES, ETC. XXX

References

1. Abreu F, Martins JL, Silveira TS, Keim CN, de Barros HG, Filho FJ, Lins U. 2007. 'Candidatus *Magnetoglobus multicellularis*', a multicellular, magnetotactic prokaryote from a hypersaline environment. *Int J Syst Evol Microbiol* **57**(Pt 6):1318–1322.
2. Arias Del Angel JA, Escalante AE, Martínez-Castilla LP, Benítez M. 2017. An Evo-Devo Perspective on Multicellular Development of Myxobacteria. *J Exp Zool B Mol Dev Evol* **328**(1–2):165–178.
3. Mayerhofer LE, Macario AJ, Conway de Macario E. 1992. Lamina, a novel multicellular form of *Methanosarcina mazei* S-6. *J Bacteriol* **174**(1):309–314.
4. Herrero A, Stavans J, Flores E. 2016. The multicellular nature of filamentous heterocyst-forming cyanobacteria. *FEMS Microbiol Rev* **40**(6):831–854.
5. Grosberg RK, Strathmann RR. 2007. The evolution of multicellularity: A minor major transition?. *Annu Rev Ecol Evol Syst* **38**:621–654.
6. Knoll AH. 2011. The multiple origins of complex multicellularity. *Annu Rev Earth Planet Sci* **39**:217–239.
7. Karnkowska A, Vacek V, Zubáčová Z, Treitli SC, Petrželková 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, Hampl V. 2016. A Eukaryote without a Mitochondrial Organelle. *Curr Biol* **26**(10):1274–1284.

8. Lane N. 2014. Bioenergetic constraints on the evolution of complex life. *Cold Spring Harb Perspect Biol* **6**(5):a015982.
9. Lindsay MR, Webb RI, Strous M, Jetten MS, Butler MK, Forde RJ, Fuerst JA. 2001. Cell compartmentalisation in planctomycetes: novel types of structural organisation for the bacterial cell. *Arch Microbiol* **175**(6):413–429.
10. Fuerst JA. 2013. The PVC superphylum: exceptions to the bacterial definition? *Antonie Van Leeuwenhoek* **104**(4):451–466.
11. Rast A, Heinz S, Nickelsen J. 2015. Biogenesis of thylakoid membranes. *Biochim Biophys Acta* **1847**(9):821–830.
12. McInerney JO, Martin WF, Koonin EV, Allen JF, Galperin MY, Lane N, Archibald JM, Embley TM. 2011. Planctomycetes and eukaryotes: a case of analogy not homology. *Bioessays* **33**(11):810–817.
13. Woronowicz K, Harrold JW, Kay JM, Niederman RA. 2013. Structural and functional proteomics of intracytoplasmic membrane assembly in *Rhodobacter sphaeroides*. *J Mol Microbiol Biotechnol* **23**(1–2):48–62.
14. Clements KD, Bullivant S. 1991. An unusual symbiont from the gut of surgeonfishes may be the largest known prokaryote. *J Bacteriol* **173**(17):5359–5362.
15. Angert ER, Clements KD, Pace NR. 1993. The largest bacterium. *Nature* **362**(6417):239–241.
16. Schulz HN, Brinkhoff T, Ferdelman TG, Mariné MH, Teske A, Jorgensen BB. 1999. Dense populations of a giant sulfur bacterium in Namibian shelf sediments. *Science* **284**(5413):493–495.
17. Mendell JE, Clements KD, Choat JH, Angert ER. 2008. Extreme polyploidy in a large bacterium. *Proc Natl Acad Sci U S A* **105**(18):6730–6734.
18. diCenzo GC, Finan TM. 2017. The Divided Bacterial Genome: Structure, Function, and Evolution. *Microbiol Mol Biol Rev* **81**(3) pii: e00019-17.
19. Volff JN, Altenbuchner J. 2000. A new beginning with new ends: linearisation of circular chromosomes during bacterial evolution. *FEMS Microbiol Lett* **186**(2):143–150.
20. von Dohlen CD, Kohler S, Alsop ST, McManus WR. 2001. Mealybug β-proteobacterial endosymbionts contain γ-proteobacterial symbionts. *Nature* **412**:433–436
21. Rivera MC, Lake JA. 2004. The ring of life provides evidence for a genome fusion origin of eukaryotes. *Nature* **431**(7005):152–155.
22. Cox CJ, Foster PG, Hirt RP, Harris SR, Embley TM 2008. The archaebacterial origin of eukaryotes. *Proc Natl Acad Sci U S A* **105**(51):20356–20361.
23. Yutin N, Makarova KS, Mekhedov SL, Wolf YI, Koonin EV. 2008. The deep archaeal roots of eukaryotes. *Mol Biol Evol* **25**(8):1619–1630.
24. Koonin EV. 2010. The origin and early evolution of eukaryotes in the light of phylogenomics. *Genome Biol* **11**(5):209.
25. Kelly S, Wickstead B, Gull K. 2011. Archaeal phylogenomics provides evidence in support of a methanogenic origin of the Archaea and a thaumarchaeal origin for the eukaryotes. *Proc Biol Sci* **278**(1708):1009–1018.
26. Williams TA, Foster PG, Cox CJ, Embley TM. 2013. An archaeal origin of eukaryotes supports only two primary domains of life. *Nature* **504**(7479):231–236.
27. McInerney JO, O'Connell MJ, Pisani D. 2014. The hybrid nature of the Eukaryota and a consilient view of life on Earth. *Nat Rev Microbiol* **12**(6):449–455.
28. Raymann K, Brochier-Armanet C, Gribaldo S. 2015. The two-domain tree of life is linked to a new root for the Archaea. *Proc Natl Acad Sci U S A* **112**(21):6670–6675.
29. Koonin EV. 2015. Origin of eukaryotes from within archaea, archaeal eukaryome and bursts of gene gain: eukaryogenesis just made easier? *Philos Trans R Soc Lond B Biol Sci* **370**(1678):20140333.
30. Eme L, Spang A, Lombard J, Stairs CW, Ettema TJG. 2017. Archaea and the origin of eukaryotes. *Nat Rev Microbiol* **15**(12):711–723.
31. 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. *J Cell Sci* **129**(20):3695–3703.
32. Mereschkowski C. 1905. Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. *Biol Centralbl* **25**:593–604.
33. Mereschkowsky K. 1910. Theorie der zwei Plasmaarten als Grundlage der Symbiogenesis, einer neuen Lehre von der Ent-stehung der Organismen. *Biol Centralbl* **30**:353–367.
34. Sagan L. 1967. On the origin of mitosing cells. *J Theor Biol* **14**(3):255–274.
35. Schwartz RM, Dayhoff MO. 1978. Origins of prokaryotes, eukaryotes, mitochondria, and chloroplasts. *Science* **199**(4327):395–403.
36. Yang D, Oyaizu Y, Oyaizu H, Olsen GJ, Woese CR. 1985. Mitochondrial origins. *Proc Natl Acad Sci U S A* **82**(13):4443–4447.
37. Andersson SG, Zomorodipour A, Andersson JO, Sicheritz–Pontén T, Alsmark UC, Podowski RM, Näslund AK, Eriksson AS, Winkler HH, Kurland CG. 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature* **396**(6707):133–140.
38. Gray MW, Burger G, Lang BF. 1999. Mitochondrial evolution. *Science* **283**(5407):1476–1481.
39. Esser C, Ahmadinejad N, Wiegand C, Rotte C, Sebastiani F, Gelius–Dietrich G, Henze K, Kretschmann E, Richly E, Leister D, Bryant D, Steel MA, Lockhart PJ, Penny D, Martin W. 2004. A genome phylogeny for mitochondria among α-proteobacteria and

a predominantly eubacterial ancestry of yeast nuclear genes. *Mol Biol Evol* **21**(9):1643–1660.

40. Fitzpatrick DA, Creevey CJ, McInerney JO. 2006. Genome phylogenies indicate a meaningful α-proteobacterial phylogeny and support a grouping of the mitochondria with the Rickettsiales. *Mol Biol Evol* **23**(1):74–85.
41. Thrash JC, Boyd A, Huggett MJ, Grote J, Carini P, Yoder RJ, Robbertse B, Spatafora JW, Rappé MS, Giovannoni SJ. 2011. Phylogenomic evidence for a common ancestor of mitochondria and the SAR11 clade. *Sci Rep* **1**:13.
42. Brindefalk B, Ettema TJ, Viklund J, Thollesson M, Andersson SG. 2011. A phylometagenomic exploration of oceanic alphaproteobacteria reveals mitochondrial relatives unrelated to the SAR11 clade. *PLoS One* **6**(9):e24457.
43. Ferla MP, Thrash JC, Giovannoni SJ, Patrick WM. 2013. New rRNA gene-based phylogenies of the *Alphaproteobacteria* provide perspective on major groups, mitochondrial ancestry and phylogenetic instability. *PLoS One* **8**(12):e83383.
44. Martijn J, Vosseberg J, Guy L, Offre P, Ettema TJG. 2018. Deep mitochondrial origin outside the sampled alphaproteobacteria. *Nature* **557**(7703):101–105.
45. Woese CR, Fox GE. 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci U S A* **74**(11):5088–5090.
46. Woese CR, Kandler O, Wheelis ML. 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci U S A* **87**(12):4576–4579.
47. Huet J, Schnabel R, Sentenac A, Zillig W. 1983. Archaebacteria and eukaryotes possess DNA-dependent RNA polymerases of a common type. *EMBO J* **2**(8):1291–1294.
48. Zillig W, Klenk HP, Palm P, Pühler G, Gropp F, Garrett RA, Leffers H. 1989. The phylogenetic relations of DNA-dependent RNA polymerases of archaebacteria, eukaryotes, and eubacteria. *Can J Microbiol* **35**(1):73–80.
49. Searcy DG, Stein DB. 1980. Nucleoprotein subunit structure in an unusual prokaryotic organism: *Thermoplasma acidophilum*. *Biochim Biophys Acta* **609**(1):180–195.
50. Searcy DG, Delange RJ. 1980. *Thermoplasma acidophilum* histone-like protein. Partial amino acid sequence suggestive of homology to eukaryotic histones. *Biochim Biophys Acta* **609**(1):197–200.
51. Sandman K, Krzycki JA, Dobrinski B, Lurz R, Reeve JN. 1990. HMf, a DNA-binding protein isolated from the hyperthermophilic archaeon *Methanothermus fervidus*, is most closely related to histones. *Proc Natl Acad Sci U S A* **87**(15):5788–5791.
52. Henneman B, van Emmerik C, van Ingen H, Dame RT. 2018. Structure and function of archaeal histones. *PLoS Genet* **14**(9):e1007582.
53. Lake JA, Henderson E, Oakes M, Clark MW 1984, Eocytes: a new ribosome structure indicates a kingdom with a close relationship to eukaryotes. *Proc Natl Acad Sci U S A* **81**(12):3786–3790.
54. Lake JA. 1988. Origin of the eukaryotic nucleus determined by rate-invariant analysis of rRNA sequences. *Nature* **331**(6152):184–186.
55. Gogarten JP, Kibak H, Dittrich P, Taiz L, Bowman EJ, Bowman BJ, Manolson MF, Poole RJ, Date T, Oshima T, et al. 1989. Evolution of the vacuolar H^+-ATPase: implications for the origin of eukaryotes. *Proc Natl Acad Sci U S A* **86**(17):6661–6665.
56. Rivera MC, Lake JA. 1992. Evidence that eukaryotes and eocyte prokaryotes are immediate relatives. *Science* **257**(5066):74–76.
57. Baldauf SL, Palmer JD, Doolittle WF. 1996. The root of the universal tree and the origin of eukaryotes based on elongation factor phylogeny. *Proc Natl Acad Sci U S A* **93**(15):7749–7754.
58. Samson RY, Bell SD. 2009. Ancient ESCRTs and the evolution of binary fission. Trends Microbiol. 2009;17:507–13.
59. Yutin N, Koonin EV. 2012. Archaeal origin of tubulin. *Biol Direct* **7**:10.
60. Ettema TJ, Lindas AC, Bernander R. 2011. An actin-based cytoskeleton in archaea. *Mol Microbiol* **80**:1052–1061.
61. Koonin EV, Yutin N. 2014. The dispersed archaeal eukaryome and the complex archaeal ancestor of eukaryotes. *Cold Spring Harb Perspect Biol* **6**(4):a016188.
62. Wolf YI, Makarova KS, Yutin N, Koonin EV. 2012. Updated clusters of orthologous genes for Archaea: a complex ancestor of the Archaea and the byways of horizontal gene transfer. *Biol Direct* **7**:46.
63. Spang A, Saw JH, Jørgensen SL, Zaremba-Niedzwiedzka K, Martijn J, Lind AE, van Eijk R, Schleper C, Guy L, Ettema TJG. 2015. Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* **521**(7551):173–179.
64. 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 TJ. 2017. Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* **541**(7637):353–358.
65. Spang A, Caceres EF, Ettema TJG. 2017. Genomic exploration of the diversity, ecology, and evolution of the archaeal domain of life. *Science* **357**(6351). pii: eaaf3883.
66. Spang A, Eme L, Saw JH, Caceres EF, Zaremba-Niedzwiedzka K, Lombard J, Guy L, Ettema TJG. 2018. Asgard archaea are the closest prokaryotic relatives of eukaryotes. *PLoS Genet* **14**(3):e1007080.

67. Da Cunha V, Gaia M, Gadelle D, Nasir A, Forterre P. 2017. Lokiarchaea are close relatives of Euryarchaeota, not bridging the gap between prokaryotes and eukaryotes. *PLoS Genet* **13**(6):e1006810.
68. Da Cunha V, Gaia M, Nasir A, Forterre P. 2018. Asgard archaea do not close the debate about the universal tree of life topology. *PLoS Genet* **14**(3):e1007215.
69. Fournier GP, Poole AM. 2018. A Briefly Argued Case That Asgard Archaea Are Part of the Eukaryote Tree. *Front Microbiol* **9**:1896.
70. 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 C, Yamanaka Y, Yamaguchi T, Kamagata Y, Tamaki H, Takai K. 2019. Isolation of an archaeon at the prokaryote-eukaryote interface. *bioRxiv* 726976
71. Martin W, Koonin EV. 2006. Introns and the origin of nucleus-cytosol compartmentalization. *Nature* **440**:41–45.
72. López–García P, Moreira D. 2019. Eukaryogenesis, a syntrophy affair. *Nat Microbiol* **4**(7):1068–1070.
73. 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. *Nat Microbiol* **4**(7):1138–1148.
74. Cavalier-Smith T. 1987. Eukaryotes with no mitochondria. *Nature* **326**(6111):332–333.
75. Cavalier-Smith T. 1989. Molecular phylogeny. Archaebacteria and Archezoa. *Nature* **339**(6220):l00–101
76. Clark CG, Roger AJ. 1995. Direct evidence for secondary loss of mitochondria in *Entamoeba histolytica*. *Proc Natl Acad Sci U S A* **92**(14):6518–6521.
77. Tovar J, Fischer A, Clark CG. 1999. The mitosome, a novel organelle related to mitochondria in the amitochondrial parasite *Entamoeba histolytica*. *Mol Microbiol* **32**(5):1013–1021.
78. Andersson SG, Kurland CG. 1999. Origins of mitochondria and hydrogenosomes. *Curr Opin Microbiol* **2**(5):535–541.
79. Williams BA, Hirt RP, Lucocq JM, Embley TM. 2002. A mitochondrial remnant in the microsporidian *Trachipleistophora hominis*. *Nature* **418**(6900):865–869.
80. Tovar J, León-Avila G, Sánchez LB, Sutak R, Tachezy J, van der Giezen M, Hernández M, Müller M, Lucocq JM. 2003. Mitochondrial remnant organelles of *Giardia* function in iron-sulphur protein maturation. *Nature* **426**(6963):172–176.
81. Hrdy I, Hirt RP, Dolezal P, Bardonová L, Foster PG, Tachezy J, Embley TM. 2004. *Trichomonas* hydrogenosomes contain the NADH dehydrogenase module of mitochondrial complex I. *Nature* **432**(7017):618–622.
82. Hjort K, Goldberg AV, Tsaousis AD, Hirt RP, Embley TM. 2010. Diversity and reductive evolution of mitochondria among microbial eukaryotes. *Philos Trans R Soc Lond B Biol Sci* **365**(1541):713–727.
83. Martijn J, Ettema TJ. 2013. From archaeon to eukaryote: the evolutionary dark ages of the eukaryotic cell. *Biochem Soc Trans* **41**(1):451–457.
84. Pittis AA, Gabaldón T1,2,3. 2016. Late acquisition of mitochondria by a host with chimaeric prokaryotic ancestry. *Nature* **531**(7592):101–104.
85. Martin WF, Roettger M, Ku C, Garg SG, Nelson-Sathi S, Landan G. 2017. Late Mitochondrial Origin Is an Artifact. *Genome Biol Evol* **9**(2):373–379.
86. Degli Esposti M. 2016. Late Mitochondrial Acquisition, Really? *Genome Biol Evol* **8**(6):2031–2035.
87. Akïl C, Robinson RC. 2018. Genomes of Asgard archaea encode profilins that regulate actin. *Nature* **562**(7727):439–443.
88. Yutin N, Wolf MY, Wolf YI, Koonin EV. 2009. The origins of phagocytosis and eukaryogenesis. *Biol Direct* **4**:9.
89. Burns JA, Pittis AA, Kim E. 2018. Gene-based predictive models of trophic modes suggest Asgard archaea are not phagocytotic. *Nat Ecol Evol* **2**(4):697–704.
90. Poole AM, Gribaldo S. 2014. Eukaryotic origins: How and when was the mitochondrion acquired? *Cold Spring Harb Perspect Biol* **6**(12):a015990.
91. Lane N, Martin W. 2010. The energetics of genome complexity. *Nature* **467**(7318):929–934.
92. Lane N. 2011. Energetics and genetics across the prokaryote-eukaryote divide. *Biol Direct* **6**:35.
93. Lane N, Martin WF. 2015. Eukaryotes really are special, and mitochondria are why. *Proc Natl Acad Sci U S A* **112**(35):E4823.
94. Lynch M, Marinov GK. 2015. The bioenergetic costs of a gene. *Proc Natl Acad Sci U S A* **112**(51):15690–15695.
95. Koonin EV. 2015. Energetics and population genetics at the root of eukaryotic cellular and genomic complexity. *Proc Natl Acad Sci U S A* **112**(52):15777–15778.
96. Booth A, Doolittle WF. 2015. Eukaryogenesis, how special really? *Proc Natl Acad Sci U S A* **112**(33):10278–10285.
97. Booth A, Doolittle WF. 2015. Reply to Lane and Martin: Being and becoming eukaryotes. *Proc Natl Acad Sci U S A* **112**(35):E4824.
98. Garg SG, Martin WF. 2016. Mitochondria, the Cell Cycle, and the Origin of Sex via a Syncytial Eukaryote Common Ancestor. *Genome Biol Evol* **8**(6):1950–1970.
99. Lane N, Martin WF. 2016. Mitochondria, complexity, and evolutionary deficit spending. *Proc Natl Acad Sci U S A* **113**(6):E666.
100. Martin WF. 2017. Symbiogenesis, gradualism, and mitochondrial energy in eukaryote evolution. *Periodicum Biologorum* **119**(**3**):141–158

101. Lynch M, Marinov GK. 2017. Membranes, energetics, and evolution across the prokaryote-eukaryote divide. *Elife* **6**. pii: e20437.
102. Lane N. 2017. Serial endosymbiosis or singular event at the origin of eukaryotes? *J Theor Biol* **434**:58–67.
103. Martin WF. 2017. Physiology, anaerobes, and the origin of mitosing cells 50 years on. *J Theor Biol* **434**:2–10.
104. Lynch M, Marinov G. 2018. Correction: Membranes, Energetics, and Evolution Across the Prokaryote-Eukaryote Divide. *Elife* **7**. pii: e35006.
105. Gerlitz M, Knopp M, Kapust M, Xavier JC, Martin WF. 2018. Elusive data underlying debate at the prokaryote-eukaryote divide. *Biol Direct* **13**:21
106. Wu Z, Cuthbert JM, Taylor DR, Sloan DB. 2015. The massive mitochondrial genome of the angiosperm *Silene noctiflora* is evolving by gain or loss of entire chromosomes. *Proc Natl Acad Sci U S A* **112**(33):10185–10191.
107. Smith DR, Keeling PJ. 2015. Mitochondrial and plastid genome architecture: Reoccurring themes, but significant differences at the extremes. *Proc Natl Acad Sci U S A* **112**(33):10177–10184.
108. Kannan S, Rogozin IB, Koonin EV. 2014. MitoCOGs: clusters of orthologous genes from mitochondria and implications for the evolution of eukaryotes. *BMC Evol Biol* **14**:237.
109. Burger G, Gray MW, Forget L, Lang BF. 2013. Strikingly bacteria-like and gene-rich mitochondrial genomes throughout jakobid protists. *Genome Biol Evol* **5**(2):418–438.
110. Allen JF. 2015. Why chloroplasts and mitochondria retain their own genomes and genetic systems: colocation for redox regulation of gene expression. *Proc Natl Acad Sci U S A* **112**:10231–10238.
111. Lynch M, Conery JS 2003. The origins of genome complexity. *Science* **302**(5649):1401–1404.
112. Lynch M. 2007. *The Origins of Genome Architecture*. Sinauer; Sunderland, MA.
113. Lynch M. 2007. The frailty of adaptive hypotheses for the origins of organismal complexity. *Proc Natl Acad Sci U S A* **104**(Suppl 1):8597–8604
114. Delaye L, Valadez-Cano C, Pérez-Zamorano B. 2016. How Really Ancient Is *Paulinella Chromatophora*? *PLoS Curr* **8**. pii: ecurrents.tol.e68a099364bb1a1e129a17b4e06b0c6b.
115. Nowack EC, Melkonian M, Glöckner G. 2008. Chromatophore genome sequence of *Paulinella* sheds light on acquisition of photosynthesis by eukaryotes. *Curr Biol* **18**(6):410–418.
116. Nowack EC, Price DC, Bhattacharya D, Singer A, Melkonian M, Grossman AR. 2016. Gene transfers from diverse bacteria compensate for reductive genome evolution in the chromatophore of *Paulinella chromatophora*. *Proc Natl Acad Sci U S A* **113**(43):12214–12219.
117. Singer A, Poschmann G, MḦlich C, Valadez-Cano C, Hänsch S, Hüren V, Rensing SA, Stühler K, Nowack ECM. 2017. Massive Protein Import into the Early-Evolutionary-Stage Photosynthetic Organelle of the Amoeba *Paulinella chromatophora*. *Curr Biol* **27**(18):2763–2773.e5.
118. Nakayama T, Ikegami Y, Nakayama T, Ishida K, Inagaki Y, Inouye I. 2011. Spheroid bodies in rhopalodiacean diatoms were derived from a single endosymbiotic cyanobacterium. *J Plant Res* **124**(1):93–97.
119. Nakayama T, Kamikawa R, Tanifuji G, Kashiyama Y, Ohkouchi N, Archibald JM, Inagaki Y. 2014. Complete genome of a nonphotosynthetic cyanobacterium in a diatom reveals recent adaptations to an intracellular lifestyle. *Proc Natl Acad Sci U S A* **111**(31):11407–11412.
120. Nakayama T, Inagaki Y. 2017. Genomic divergence within non-photosynthetic cyanobacterial endosymbionts in rhopalodiacean diatoms. *Sci Rep* **7**(1):13075.
121. Bennett GM, Moran NA. 2013. Small, smaller, smallest: the origins and evolution of ancient dual symbioses in a Phloem-feeding insect. *Genome Biol Evol* **5**(9):1675–1688.
122. Moran NA, Bennett GM. 2014. The tiniest tiny genomes. *Annu Rev Microbiol* **68**:195–215.
123. Martin WF, Garg S, Zimorski V. 2015. Endosymbiotic theories for eukaryote origin. *Philos Trans R Soc Lond B Biol Sci* **370**(1678):20140330.
124. de Duve C. 2007. The origin of eukaryotes: a reappraisal. *Nat Rev Genet* **8**(5):395–403.
125. Gross J, Bhattacharya D. 2010. Uniting sex and eukaryote origins in an emerging oxygenic world. *Biol Direct* **5**:53.
126. Eme L, Sharpe SC, Brown MW, Roger AJ. 2014. On the age of eukaryotes: evaluating evidence from fossils and molecular clocks. *Cold Spring Harb Perspect Biol* **6**(8) pii: a016139.
127. Knoll AH. 2014. Paleobiological perspectives on early eukaryotic evolution. *Cold Spring Harb Perspect Biol* **6**(1). pii: a016121.
128. Parfrey LW, Lahr DJ, Knoll AH, Katz LA. 2011. Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proc Natl Acad Sci U S A* **108**(33):13624–13629.
129. Van Der Giezen M, Lenton TM. 2012. The rise of oxygen and complex life. *J Eukaryot Microbiol* **59**(2):111–113.
130. Martin W, Müller M. 1998. The hydrogen hypothesis for the first eukaryote. *Nature* **392**(6671):37–41.
131. Bulzu PA, Andrei AS, Salcher MM, Mehrshad M, Inoue K, Kandori H, Beja O, Ghai R, Banciu HL. 2019. Casting light on Asgardarchaeota metabolism in a sunlit microoxic niche. *Nat Microbiol* **4**(7):1129–1137.
132. Seitz KW, Dombrowski N, Eme L, Spang A, Lombard J, Sieber JR, Teske AP, Ettema TJG, Baker BJ. 2019.

Asgard archaea capable of anaerobic hydrocarbon cycling. *Nat Commun* **10**(1):1822.

133. Gould SB, Garg SG, Martin WF. 2016. Bacterial Vesicle Secretion and the Evolutionary Origin of the Eukaryotic Endomembrane System. *Trends Microbiol* **24**(7):525–534.
134. Hoekstra D, van der Laan JW, de Leij L, Witholt B. 1976. Release of outer membrane fragments from normally growing *Escherichia coli*. *Biochim Biophys Acta* **455**(3):889–899.
135. Kulp A, Kuehn MJ. 2010. Biological functions and biogenesis of secreted bacterial outer membrane vesicles. *Annu Rev Microbiol* **64**:163–184.
136. Schwechheimer C, Kuehn MJ. 2015. Outer-membrane vesicles from Gram-negative bacteria: biogenesis and functions. *Nat Rev Microbiol* **13**(10):605–619.
137. Cook KL, Soto-Pantoja DR, Abu-Asab M, Clarke PA, Roberts DD, Clarke R. 2014. Mitochondria directly donate their membrane to form autophagosomes during a novel mechanism of parkin-associated mitophagy. *Cell Biosci* **4**(1):16.
138. Soubannier V, McLelland GL, Zunino R, Braschi E, Rippstein P, Fon EA, McBride HM. 2012. A vesicular transport pathway shuttles cargo from mitochondria to lysosomes. *Curr Biol* **22**(2):135–141.
139. Braschi E, Goyon V, Zunino R, Mohanty A, Xu L, McBride HM. 2010. Vps35 mediates vesicle transport between the mitochondria and peroxisomes. *Curr Biol* **20**(14):1310–1315.
140. Villanueva L, Schouten S, Damsté JS. 2017. Phylogenomic analysis of lipid biosynthetic genes of Archaea shed light on the 'lipid divide'. *Environ Microbiol* **19**(1):54–69.
141. Jékely G. 2007. Origin of eukaryotic endomembranes: a critical evaluation of different model scenarios. *Adv Exp Med Biol* **607**:38–51.
142. Cavalier-Smith T. 2002. The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. *Int J Syst Evol Microbiol* **52**:297–354.
143. Poole AM, Neumann N. 2011. Reconciling an archaeal origin of eukaryotes with engulfment: a biologically plausible update of the Eocyte hypothesis. *Res Microbiol* **162**(1):71–76.
144. Wujek DE. 1979. Intracellular bacteria in the blue-green-alga *Pleurocapsa minor*. *Trans Am Microsc Soc* **98**:143–145.
145. Thao ML, Gullan PJ, Baumann P. 2002. Secondary (γ-Proteobacteria) endosymbionts infect the primary (β-Proteobacteria) endosymbionts of mealybugs multiple times and coevolve with their hosts. *Appl Environ Microbiol* **68**(7):3190–3197.
146. Husnik F, Nikoh N, Koga R, Ross L, Duncan RP, Fujie M, Tanaka M, Satoh N, Bachtrog D, Wilson AC, von Dohlen CD, Fukatsu T, McCutcheon JP. 2013. Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. *Cell* **153**(7):1567–1578.
147. Sassera D, Beninati T, Bandi C, Bouman EA, Sacchi L, Fabbi M, Lo N. 2006. '*Candidatus* Midichloria mitochondrii', an endosymbiont of the tick *Ixodes ricinus* with a unique intramitochondrial lifestyle. *Int J Syst Evol Microbiol* **56**(Pt 11):2535–2540.
148. Martin WF, Tielens AGM, Mentel M, Garg SG, Gould SB. 2017. The Physiology of Phagocytosis in the Context of Mitochondrial Origin. *Microbiol Mol Biol Rev* **81**(3). pii: e00008–17.
149. Sockett RE. 2009. Predatory lifestyle of *Bdellovibrio bacteriovorus*. *Annu Rev Microbiol* **63**:523–539.
150. Fedorov A, Feisal Merican A, Gilbert W. 2002. Large-scale comparison of intron positions among aminal, plant, and fungal genes. *Proceedings of the National Academy of Sciences*. **99:**16128-16133.
151. Rogozin IB, Wolf YI, Sorokin AV, Mirkin BG, Koonin EV. 2003. Remarkable Interkingdom conservation of intron positions and massive, lineage-specific intron loss and gain in eukaryotic evolution. *Current Biology*. **13**:1512-1517.
152. Roy SW. 2006. Intron-rich ancestors. *Trends in Genetics* **22**:468–471.
153. Csuros M, Rogozin IB, Koonin EV. 2011. A Detailed History of Intron-rich Eukaryotic Ancestors Inferred from a Global Survey of 100 Complete Genomes. *PLoS Computational Biology* **7**(9):e1002150
154. Collins L, Penny D. 2005. Complex spliceosomal organization ancestral to extant eukaryotes. *Mol Biol Evol* **22**:1053–1066
155. Yan C, Hang J, Wan R, Huang M, Wong CC, Shi Y. 2015. Structure of a yeast spliceosome at 3.6-angstrom resolution. *Science* **349**(6253):1182–1191.
156. Toor N, Keating KS, Taylor SD, Pyle AM. 2008. Crystal structure of a self-spliced group II intron. *Science* **320**(5872):77–82.
157. Lambowitz AM, Zimmerly S. 2011. Group II introns: mobile ribozymes that invade DNA. *Cold Spring Harb Perspect Biol* **3**(8):a003616.
158. Koonin EV. 2006. The Origin of Introns and Their Role in Eukaryogenesis: A Compromise Solution to the Introns-Early Versus Introns-Late Debate? *Biology Direct* **1**:22.
159. Ramesh MA, Malik SB, Logsdon JM Jr. 2005. A phylogenomic inventory of meiotic genes; evidence for sex in *Giardia* and an early eukaryotic origin of meiosis. *Curr Biol* **15**(2):185–191.
160. Hörandl E, Hadacek F. 2013. The oxidative damage initiation hypothesis for meiosis. *Plant Reprod* **26**(4):351–367.
161. Speijer D, Lukeš J, Eliáš M. 2015. Sex is a ubiquitous, ancient, and inherent attribute of eukaryotic life. *Proc Natl Acad Sci U S A* **112**(29):8827–8834.
162. Hŕandl E, Speijer D. 2018. How oxygen gave rise to eukaryotic sex. *Proc Biol Sci* **285**(1872). pii: 20172706.

163. Klinger CM, Spang A, Dacks JB, Ettema TJ. 2016. Tracing the Archaeal Origins of Eukaryotic Membrane-Trafficking System Building Blocks. *Mol Biol Evol* **33**(6):1528–1541.
164. Lynch M. 2006. Streamlining and Simplification of Microbial Genome Architecture. *Annu Rev Microbiol* **60**:327-349.
165. Lynch M. 2006. The Origins of Eukaryotic Gene Structure. *Mol Biol Evol* **23**(2):450-468
166. Lynch M. 2007. The evolution of genetic networks by non-adaptive processes. *Nat Rev Genet* **8**:803–813.
167. Lynch M, Force A. 2000. The probability of duplicate gene preservation by subfunctionalization. *Genetics* **154**(1):459–473.
168. Lynch M, O'Hely M, Walsh B, Force A. 2000. The probability of preservation of a newly arisen gene duplicate. *Genetics* **159**(4):1789–1804.
169. ENCODE Project Consortium. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**(7414):57–74.
170. Stoltzfus A. 1999. On the possibility of constructive neutral evolution. *J Mol Evol* **49**:169–181
171. Stoltzfus A. 2012. Constructive neutral evolution: exploring evolutionary theory's curious disconnect. *Biology Direct* **7**:35
172. Adl SM, Simpson AG, Lane CE, Lukeš J, Bass D, Bowser SS, Brown MW, Burki F, Dunthorn M, Hampl V, Heiss A, Hoppenrath M, Lara E, Le Gall L, Lynn DH, McManus H, Mitchell EA, Mozley-Stanridge SE, Parfrey LW, Pawlowski J, Rueckert S, Shadwick RS, Schoch CL, Smirnov A, Spiegel FW. 2012. The revised classification of eukaryotes. *J Eukaryot Microbiol* **59**(5):429–493.