Mitochondria and Major Transitions at Eukaryote Origin

Inaugural dissertation

for the attainment of the title of doctor
in the Faculty of Mathematics and Natural Sciences
at the Heinrich-Heine University Düsseldorf

presented by

Sriram Garg

from Thrissur, India

2017

From the Institute of Molecular Evolution
at the Heinrich Heine University Düsseldorf

Published by permission of the
Faculty of Mathematics and Natural Sciences at
Heinrich Heine University Düsseldorf

Supervisor: Prof. Dr. William F. Martin
Co-supervisor: Prof. Dr. Peter Jahns
Date of the oral examination: 10.03.2017

Statement of declaration

I hereby declare that this dissertation is the result of my own work. No other person's work has been used without due acknowledgment. This dissertation has not been submitted in the same or similar form to other institutions. I have not previously failed a doctoral examination procedure.

Düsseldorf, 10.03.2017 Sriram G. Garg

For Amma and Appa...

The secret of complex life lies in the chimeric nature of the eukaryotic cell — a hopeful monster, born in an improbable merger 2000 million years ago, an event still frozen in our innermost constitution and dominating our lives today.

—Nick Lane

Publications included in this thesis

I Martin WF, **Garg S**, Zimorski V (2015) Endosymbiotic theories for eukaryote origin. *Phil. Trans. R. Soc. B.* **370**:20140330.

II Gould SB, **Garg SG**, Martin WF (2016) Bacterial vesicle secretion and the evolutionary origin of the eukaryotic endomembrane system. *Trends Microbiol.* **24**:525–534.

III **Garg S**, Stölting J, Zimorski V, Rada P, Tachezy J, Martin WF, Gould SB (2015) Conservation of transit peptide-independent protein import into the mitochondrial and hydrogenosomal matrix. *Genome Biol. Evol.* **7**:2716–2726.

IV Rada P, Makki AR, Zimorski V, **Garg S**, Hampl V, Hrdý I, Gould SB, Tachezy J (2015) N-terminal presequence-independent import of phosphofructokinase into hydrogenosomes of Trichomonas vaginalis. *Eukaryot. Cell* **14**:1264–1275.

V **Garg SG**, Gould SB (2016) The role of charge in protein targeting evolution. *Trends Cell Biol.* **12**:894–905

VI **Garg SG**, Martin WF (2016) Mitochondria, the cell cycle, and the origin of sex via a syncytial eukaryote common ancestor. *Genome Biol. Evol.* **8**:1950–1970.

Other Publications

I Ku C, Nelson-Sathi S, Roettger M, **Garg S**, Hazkani-Covo E, Martin WF (2015) Endosymbiotic gene transfer from prokaryotic pangenomes: Inherited chimerism in eukaryotes. *Proc. Natl. Acad. Sci. U.S.A.* **112**:10139–10146.

II Martin WF, Roettger M, Ku C, **Garg SG**, Nelson-Sathi S, Landan G (2016) Late mitochondrial origin is pure artefact. *bioRxiv:055368*.

Table of contents

Zusammenfassung

Der Stammbaum des Lebens unterteilt sich strikt in Prokaryoten und Eukaryoten. Diese Grenze macht sich auf vielen Ebenen bemerkbar – von der biochemischen Diversität bis hin zur morphologischen Komplexität. Während Prokaryoten Einzeller geblieben sind und sich durch vielseitige Stoffwechselwege auszeichnen, ist komplexes, mehrzelliges Leben auf diesem Planeten ausschließlich den Eukaryoten vorbehalten. So viele Merkmale Eukaryoten auszeichnen – letztlich gibt es kein bedeutungsvolleres als Mitochondrien. Es ist inzwischen allgemein anerkannt, dass es sich bei Mitochondrien um einst frei lebende Proteobakterien handelt, die sich mit Archaeen zusammenschlossen und eine symbiotische Beziehung eingingen, deren Produkt Eukaryoten sind. Da es keine Zwischenstufen zwischen Prokaryoten und Eukaryoten gibt, stellte die Eukaryogenese eine extreme Herausforderung für Biologen dar. Genau aus diesem Grund ist die Endosymbiontentheorie auch über jene Theorien erhaben, die sich mit der Existenz von hypothetischen Zwischenprodukten beschäftigen. Solange keine Beweise für solche Intermediate existieren, ergibt es mehr Sinn, die Grenzen der Endosymbiontentheorie hinsichtlich der Eukaryogenese zu testen. Eukaryoten verfügen über viele einzigartige, biologische Prozesse und Wege, die in Prokaryoten fehlen. Evolutionär betrachtet lassen sich so wichtige Übergänge festlegen, von denen drei in dieser Arbeit untersucht werden. Der Ursprung des eukaryotischen Endomembransystems durch die äußere Membran-Vesikelsekretion der Endosymbionten liefert das Ausgangsmaterial für die Evolution der morphologischen Komplexität in Eukaryoten. Wirts- und Symbiontenphysiologie mussten sich aneinander anpassen. Das erforderte einen Gentransfer, wozu sich zunächst ein neuer Protein-Import-Mechanismus für den Symbionten entwickeln musste. Die Entwicklung des mitochondrialen Protein-Targeting war entscheidend für die Stärkung der Symbiose und das Übertragen der Kontrolle an den Wirt. Spuren dieser alten Protein-Targeting und Proteinimport Mechanismen können auch heute noch bei Organismen beobachtet werden, die wieder in einen einfacheren, anzestralen Zustand zurückkehren, wie *Trichomonas vaginalis*. Schließlich wird auch der Ursprungs des Zellzyklus aus einer mitochondrialen Perspektive untersucht. Die Bildung des Zellkerns und damit die des Cytosols entkoppelte die Genom- von der Zellteilung. Ein mehrkerniger Zwischenzustand ermöglicht nicht nur eine von der Zellteilung unabhängige Genomteilung, sondern schafft auch den Übergang von einem autarken Endosymbionten zu einem, dessen Teilung durch den Wirt gesteuert wird. Das ist auch ein wichtiger Grund für die Monophylie der Eukaryoten, die in phylogenetischen Bäumen zu beobachten ist. Die Endosymbiose von Bakterien in Archaeen war ein bemerkenswertes Ereignis, das Zellen vor neue Herausforderungen stellte und die Epoche komplexer Zellen einleiten.

Abstract

The tree of life is deeply divided into the prokaryotes and eukaryotes. This division is visible on many levels, ranging from biochemical diversity to morphological complexity. Eukaryotes are the source of all complex life on this planet, while the prokaryotes have remained unicellular, simple in morphology but metabolically versatile. Although there are many features that unite eukaryotes, none is more significant than the mitochondria. Mitochondria are now accepted to stem from once free-living proteobacteria that associated with an archaeal host and ultimately entered an endosymbiotic relationship. The product of this relationship are the eukaryotes. Eukaryogenesis remains an extremely challenging problem for biologists, since no intermediates between prokaryotes and eukaryotes are known. This is precisely why endosymbiotic theory triumphs over other theories that propose the existence of hypothetical intermediates. In the absence of evidence for such intermediates, it is prudent to test the limits of the endosymbiotic theory with respect to its explanatory power regarding eukaryogenesis. Eukaryotes have many unique biological processes and pathways that are absent in prokaryotes. Evolutionarily speaking, these constitute major transitions, three of which are explored in this thesis: (i) The origin of the eukaryotic endomembrane system through outer membrane vesicle secretion by the endosymbiont provided the initial raw material for the evolution of the intracellular complexity of eukaryotes. (ii) Concurrent with the natural transformation of the host's and endosymbiont's physiology to accommodate one another, gene transfer from symbiont to the host genome required the evolution of novel protein import mechanisms into the endosymbiont. The evolution of mitochondrial protein targeting was crucial in reinforcing endosymbiosis and transferring ever more control over the endosymbiont, to the host. Traces of the ancestral mechanisms of protein targeting and import can still be observed today in cases where organisms (such as *Trichomonas vaginalis*) are reverting back to a simpler, ancestral state of protein targeting. (iii) Lastly, the problem of the origin of the cell cycle is explored from a mitochondrial perspective. The formation of the nucleus allowed to uncouple the process of genome division from cell division. Not only does a multi-nucleated intermediate, a syncytial cell, allow for genome division in the absence of cell division, it also accommodates the transition between a self-sufficient endosymbiont to one whose division is controlled by the host. Importantly, it also accounts for the monophyly of eukaryotes as observed in phylogenetic trees. Endosymbiosis of a bacterium within an archaeon was a truly remarkable event and one that precipitated novel and challenging obstacles in biology and that paved the way to life at the level of complex cells.

Outline of the thesis

In this thesis, eukaryogenesis is explored from the standpoint of endosymbiotic theory which briefly posits that eukaryotes emerged as a result of the symbiotic (syntrophic) association of an archaeal host with a bacterial symbiont. The thesis is structured to highlight and explore three aspects of eukaryotic biology and how they might have evolved from prokaryotic ancestors namely the eukaryotic endomembrane system, protein targeting to the mitochondria and the origin of the cell cycle and sex.

Origin of subcellular complexity

Conventional models to explain the origin of the endomembrane system traditionally start from invagination of the plasma membrane to form a rudimentary endoplasmic reticulum or ER. That however results in the ER lumen being homologous to the environment. Studies on ER have shown that the lumen is functionally homologous to the bacterial periplasm rather than the extracellular environment. This thesis explores a new idea for the origin of the endomembrane system, starting from outer membrane vesicle (OMV) secretion by a bacterial endosymbiont within an archaeal host.

Inventing protein import

Following the acquisition of the mitochondrion, genes from the endosymbiont were transferred to the host genome. This prompted the invention and evolution of machineries that would import proteins encoded by these genes. Current paradigms on protein import involve an N-terminal targeting sequence on proteins which is recognized by a receptor platforms on the mitochondria and subsequently translocated across membranes by translocases. The order of origin of these three components remain a mystery. The discovery of organisms with reduced mitochondria with concomitantly reduced import machineries such as *Trichomonas vaginalis* can be used to identify ancestral import pathways. With the origin of plastids in the plant lineage, new constraints were imposed upon the cell for imparting specificity to protein import.This thesis explores the impact of plastid acquisition on the evolution on mitochondrial and plastid import in organisms harboring both organelles.

Origin of the cell cycle and sex

Among all the major cellular innovations that encompass eukaryogenesis none is more fundamental than the origin of the eukaryotic cell cycle. The cell cycle governs how eukaryotic cells replicate and divide. While a vast number of theories have attempted to explain the origin, prevalence, and importance of sex in eukaryotes almost all of them exclusively rely on the exsistence of the eukaryotic cell cycle. In an attempt to bridge the gap between the origin of the cell cycle and reciprocal recombination in eukaryotes, this thesis provides an alternative set of explanations from the perspective of endosymbiosis.

The prokaryote–eukaryote divide

Although coined by Robert Hooke in his now famous book *Micrographia* in 1665 (Hooke, 1665), it was Antonie van Leeuwenhoek, who first observed living cells invisible to the naked eye and coined the term "*Animalcules*" (Lane, 2015). Thus began the era of microbiology — fathered by van Leuwenhoek — and the observation and characterization of microscopic cells. In the year 1892 Dimitry Iwanowski, a Russian Botanist discovered the first virus (Iwanowski, 1892), while studying the tobacco mosaic disease. Subsequently confirmed by Martinus Beijerinck in 1898 (Beijerinck, 1942; Lechevalier, 1972), the idea of particulate matter causing infections was radical at that time but, eventually the evidence was irrefutable. With viruses added to the mix, biology had become increasingly diverse by the time Edouard Chatton (Chatton, 1925; Sapp, 2005), first proposed the terms "procaryotic" and "eucaryotic". This distinction remained largely insignificant to the scientific community until Stanier and van Neil revisited this problem in their 1962 paper "*The concept of a bacterium*", in which they provided a detailed definition of "procaryotic" and "eucaryotic" cells (Stanier and van Niel, 1962). It was not until the late 1960s that classification of organisms into "prokaryotes" and "eukaryotes" became common place in biology. Even though, at its onset, the term microbiology encompassed the study of all organisms invisible to the naked eye, the advent of newer microscopes with higher magnification and resolutions allowed researchers to ascertain that among "microorganisms" there existed a lot of diversity in terms of shape. Biochemical analysis of these organisms revealed an even larger diversity in metabolism and physiology of microbes

In 1977, Carl Woese and George Fox discovered that bacteria fall into two distinct classes (Woese and Fox, 1977). Thus biologists had a new group of prokaryotes — the "archaebacteria" or "archaea" with its own unique biology. It quickly became apparent that the archaea had more in common with eukaryotes than bacteria (Klenk and Doolittle, 1994; Brown *et al.*, 2007; Gribaldo *et al.*, 2010; Cavicchioli, 2011; Poole and Gribaldo, 2014). This incited renewed interest to explain the relationships of eukaryotes to prokaryotes. Faster sequencing techniques, opened up access to a vast breadth of microbial diversity for biologists to study (Fraser *et al.*, 2000). Along with that came a lot of revelations about the relationships between the three major kingdoms of life. In phylogenetic trees, the archaeal components of eukaryotes appear to emerge from within the archaea, rather than as sisters to them, supporting a two domain view of life (Cox *et al.*, 2008; Williams *et al.*, 2012; 2013; McInerney *et al.*, 2014; Raymann *et al.*, 2015).

It is now evident that prokaryotes encode the building blocks for a vast majority of eukaryotic traits for including linear chromosomes (Bentley *et al.*, 2002), histone-like proteins (Reeve *et al.*, 1997; Slesarev *et al.*, 1998; Peeters *et al.*, 2015), kinases and phosphatases (Pereira *et al.*, 2011; Kennelly, 2014), recombination (Shinohara *et al.*, 1992; Camerini-Otero and Hsieh, 1995; Cohan and Aracena, 2012), internal membrane structures (Santarella-Mellwig *et al.*, 2013; Whiddon and Konopka, 2015), cytoskeletal proteins (Carballido-Lopez, 2006; Dyer, 2009; Duggin *et al.*, 2015), endosymbionts (Husnik *et al.*, 2013; Husnik and McCutcheon, 2016; von Dohlen *et al.*, 2001), and vesicular membrane trafficking (Mast *et al.*, 2014; Schlacht *et al.*, 2014; Klinger *et al.*, 2016). However, in spite of having all or nearly all the necessary

building blocks for natural section to chisel its way to "eukaryote perfection", prokaryotes still lack the morphological complexity commonly attributed to eukaryotes. Trying to explain the origin of eukaryotes has remained a major conundrum for evolutionary biologists. Where does this complexity stem from? What drove the major transitions that transformed a prokaryote to an eukaryote? Mitochondria might hold the answers to both these questions.

Endosymbiosis

Ever since microbiologists started investigating eukaryotic cells under the microscope, it was evident that the mitochondrion and the plastid were radically different from all other eukaryotic compartments in both form and function. Richard Altmann is occasionally credited with endosymbiotic theory however, what he observed were granules in stained slides of cells, which he called bioblasts (Altmann, 1890; Martin *et al.*, 2015). According to Altmann, bioblasts were functional units of the cell, but he did not comment on the origin of these bioblasts (Martin *et al.*, 2015). It was in 1905 that a Russian biologist named Konstatin Mereschkowsky provided an explanation for the origin of chloroplasts as a symbiotic event between a bacteria and a host (Mereschkowsky, 1905; Martin and Kowallik, 1999). Portier extended this idea by including the mitochondrion; suggesting that they descended from bacteria (Portier, 1918; Archibald, 2014). Wallin further developed the endosymbiotic theory for the mitochondria but like Portier, he too falsely believed and claimed that the mitochondria could be cultured in a cell-free system (Wallin, 1925). These claims together with the radical nature of the proposition invited relentless criticism among biologists who at that time were staunch believers in Darwin's gradual model of evolution through natural selection. It was not until 1967 that Lynn Marguilis (then Lynn Sagan) revitalized the endosymbiotic theory and included a symbiotic origin of the flagella from Spirochaetes in her seminal paper *"On the origin of mitosing cells"* (Sagan, 1967).

Acceptance of the endosymbiotic theory gave rise to new questions. What was the exact nature of the plastid and the mitochondrial ancestor? And who was the host? The physiological nature of the "host" that acquired the endosymbiont was being speculated on by Christian de Duve in 1969 (De Duve, 1969) and Roger Stanier in 1970 (Stanier, 1970). Both were among the first to suggest the cell that gave home to the mitochondrion was a phagotroph, which at the same time explained how the bacterial symbionts was acquired. As the endosymbiotic theory garnered more evidence through the sequencing of bacterial and organelle genomes, the plastid ancestor was attributed to cyanobacteria (Gray and Doolittle, 1982; Gray, 1983) and the mitochondrial ancestor to α-proteobacteria (**?**Yang *et al.*, 1985). But what about the host? In an illustration by Ford Dolittle in 1980 (Doolittle, 1980), the host was depicted as a phagotrophic cell (like a eukaryote, but without defined structures) and with that the search for evidence of such an intermediate began.

Many theories were proposed to explain the origin and evolution of this "amitochondriate" organism — the "protoeukaryote". By the early 1980s it was already known that organisms such as *Giardia* and *Trichomonas* lack classical mitochondria and that they branched deep in rRNA phylogenetic trees (Sogin, 1989). Cavalier-Smith classified these two species and other organisms which seem to lack mitochondria into a new group he termed *"Archezoa"* and argued that *"...the most primitive eukaryote was a phagotrophic*

archezoan...” (Cavalier-Smith, 1987). This was not radical and was rather readily accepted since it followed all of the Darwinian principles of gradual evolution by natural selection. Not unlike the *Archaeopteryx* — a transitional organism between reptiles and avians — organisms like *Giardia* and *Microsporidia* conformed with the concept of a phagotrophic host and the evolutionary intermediate between prokaryotes and eukaryotes.

However, investigations of organelle proteins revealed that some members of the “Archezoa” either harbored a mitochondrion or once possessed one shared ancestry with one (Roger *et al.*, 1996; Horner *et al.*, 1996; Bui *et al.*, 1996; Germot *et al.*, 1996). This finding was merely the first in a series of papers that came forth against the “Archezoa hypothesis”. It was finally put to rest by genomic evidence that all “Archezoa” were chimeric with respect to their genomes, similar to their eukaryotic counterparts making them divergent eukaryotes rather than being ancestral eukaryotes (Embley and Hirt, 1998; Pfanner and Geissler, 2001; Embley *et al.*, 2003; Embley and Martin, 2006). This paved the way for another theory for the nature of the host — an archaeaon. The archaeal ancestry of eukaryotes was put forth by James Lake in the “Eocyte hypothesis” based on analyses of ribosome structure (Lake *et al.*, 1984). The archaeal origin of eukaryotes is now supported by many independent studies (Cox *et al.*, 2008; Williams *et al.*, 2012; 2013; Raymann *et al.*, 2015; Spang *et al.*, 2015), but the (morphological) nature of the archaeal host cell remains a matter of dispute.

While understanding of the relationships between eukaryotes and archaea progressed, the biological nature of the host that engulfed the endosymbiont had to be adapted to keep pace with the newer findings. This has led many in the field to propose the concept of an archaeal host with phagocytic capabilities (Yutin *et al.*, 2009; Makarova *et al.*, 2010; Guy and Ettema, 2011), more specifically formulated in the “Phagocytosing Archaeon Theory” or PhAT (Martijn and Ettema, 2013). PhAT postulates the presence of a phagocytosying archaeon (with or without a nucleus) that engulfed a α-proteobacterium to become the common ancestor of all eukaryotes. Evidence seemingly in support of this theory comes from sequencing of newer archaeal lineages which seem to encode many proteins that share homology with eukaryotic proteins involved in membrane trafficking (Spang *et al.*, 2015; Saw *et al.*, 2015; Klinger *et al.*, 2016; Dacks *et al.*, 2016). Yet there are many caveats in this interpretation; As Dey and colleagues phrase it best *“...Inferring the form and behavior of an organism from genomic information alone is difficult, especially when the gene families of relevance are ancient and their relationships uncertain...”* (Dey *et al.*, 2016). The absence of any experimental evidence for a phagocytosing prokaryote is might be due to insufficient sampling but, it could also be the result of a natural barrier that hinders prokaryotes from performing phagocytosis.

Phagocytosis is a specialized process which starts with receptor based recognition of a particle (usually $\geq 0.5\,\mu m$) followed by changes in lipid composition, and cortical actin based membrane protrusion to surround the particle, and subsequent invagination (Haas, 2007; Doherty and McMahon, 2009). The so formed phagosome then proceeds to fuse with various vesicles of the eukaryotic membrane trafficking machinery culminating in a fusion with the lysosome, which is crucial to digest the engulfed particle and derive energy from it (Flannagan *et al.*, 2012). Phagocytosis is thus meaningless without the other components of the eukaryotic endomembrane system for which there is no prokaryotic homolog yet.

More importantly most endosymbiotic theories deal with the origin of eukaryotes and the origin of the mitochondria as independent events. However, the ubiquitous presence of mitochondria in all extant eukaryotic lineages and the absence of eukaryotes that never possessed a mitochondrion (Embley and Martin, 2006), inadvertently ties the origin of the mitochondria to eukaryote origins.

The hydrogen hypothesis put forth by Martin and Müller in (1998) , posits an origin for eukaryotes that involves the endosymbiosis of an α-proteobacterium and an archaeal host, following anaerobic metabolic syntrophy. While providing an explanation for the driving force behind the endosymbiosis the hydrogen hypothesis" accounts for (i) anaerobic forms of mitochondria (ii) the bacterial glycolytic pathway in the eukaryotic cytosol and (iii) the absence of intermediate eukaryotes that resemble Archezoans. The hydrogen hypothesis makes mitochondrial acquisition crucial for the evolution of eukaryotes (Martin and Müller, 1998). Moreover, as Lane and Martin argue, without the mitochondria, the origin of eukaryotes poses insurmountable energetic barriers (Lane and Martin, 2010).

Morphological complexity in eukaryotes in part stems from the myriad of cytoskeletal proteins that provide dynamic structure to the cell (Pollard, 2003). One µm of a microtubule filament consists of approximately 1,300 α/β-tubulin dimers, which are encoded by only two genes. A typical cell would have hundreds of microtubule filaments dynamically polymerizing and depolymerizing, the synthesis of which requires as much ATP s a typical prokaryote requires for an entire cell division (Garg and Martin, 2016). Morphological complexity thus arises not from the presence of genes alone, but rather from the energetic ability to express large amounts of the protein (Lane and Martin, 2010; Garg and Martin, 2016). Mitochondria allowed the host cell to express large amounts of protein and thus allowed the cell to explore alternative solutions to the problems that endosymbiosis brought with it. Criticisms of this idea are ever present (Lynch and Marinov, 2015) but fail to explain why the eukaryote common ancestor possessed a mitochondria. The endosymbiotic theory for the origin of eukaryotes is a robust hypothesis with far reaching implications for major transitions in eukaryotic evolution. More than ever, mitochondria appear to be the key to the eukaryotic state (Gould *et al.*, 2016; Garg and Martin, 2016; Garg and Gould, 2016).

The advent of the eukaryotic cell — complex cells with a nucleus — was truly one of the most important evolutionary transitions, it heralded the appearance of complex life. Development of the nucleus, the one feature that separates eukaryotes from prokaryotes (*pro* – before and *karyon* – nucleus), was long considered the defining invention that gave rise to eukaryotes. However, numerous findings now suggest that mitochondria are as ubiquitous as the nucleus and provide the energetic basis for the invention of eukaryotic specific traits (Lane and Martin, 2010; Lane, 2014; Lane and Martin, 2016).The literature on eukaryogenesis is vast and an overview of the current state of the field is provided in publication I of this thesis.

Major transitions at eukaryote origin

The origin of eukaryotes from prokaryotes has always been a difficult problem. Microfossil evidence indicates that around 1.7 billion years ago (Parfrey *et al.*, 2011), eukaryotes appeared on our planet. How their appearance came to be is an issue of continuing debate (Martin *et al.*, 2015). A premise of the present thesis is that evidence for the origin of the eukaryotic cell is preserved in both its genes and structures.

Origin of subcellular complexity

One of the most striking differences, immediately apparent from microscopic observations, between prokaryotes and eukaryotes is their morphology. With respect to their cellular morphology prokaryotes are simple, with a bilayer membrane, in some cases two; surrounding a cytoplasm which houses the genome, the information processing machinery (proteins that perform transcription,translation and their regulation), and all of their metabolic enzymes. This common theme unites bacteria and archaea, although some variation on this theme exists among prokaryotes. Bacteria are divided into two major classes, the Gram-negative and the Gram-positive bacteria. The outer membrane of Gram-negative bacteria surrounds the inner membrane, forming a periplasm between the two. The periplasm serves many important functions and houses enzymes for reactions that are unfavorable to the cytoplasm, such as oxidative folding and calcium storage (Missiakas and Raina, 1997; Jones *et al.*, 2002; Merdanovic *et al.*, 2011). Gram-positive bacteria on the other hand are rich in peptidoglycan; a glycoprotein mesh that imparts stability from the exterior of the cell (Brown *et al.*, 2015). Similar to the Gram-positive bacteria, a majority of the archaea have a so-called S-layer to impart stability to the cell with some resembling Gram-negative bacteria (Klingl, 2014). This simplicity in structure is accompanied with average sizes ranging between 5 µm and 0.2 µm(Portillo *et al.*, 2013). Photosynthetic bacteria and some other species have internal membrane structures, however these internal membranes are not in a state of constant flux, instead they increase the surface area of the cell membrane and usually attuned for a specific function (Lane and Martin, 2010; Niederman, 2016; Pinevich, 1997). In prokaryotes, there is no true homolog for the dynamic eukaryotic endomembrane system.

The eukaryotic endomembrane system; composed of the nucleus, the endoplasmic reticulum (ER), Golgi, lysosomes, peroxisomes, endosomes, autophagosomes, multi-vesicular bodies and other small compartments interconnected by vesicular flux; is a visually distinguishing hallmark of eukaryotes (Field *et al.*, 2011; Gould *et al.*, 2016). The eukaryotic endomembrane system is not only complex in space but also in time. The dynamic nature of the endomembrane system coupled with vesicular flux between compartments sets it apart from membrane systems formed in prokaryotes.

Discussion of the eukaryotic endomembrane system typically begin with the ER. It is the largest subcellular compartment and the source of all other compartments. During cell division the ER initiates formation of the nucleus and remains connected to it such that the ER lumen is continuous with the nuclear lumen (Olmos *et al.*, 2015). Crossing the ER membrane is the first step for secretory proteins (Lippincott-Schwartz *et al.*, 2000), which occurs co-translationally via the SEC61 complex. COPI and COPII coated vesicles that bud from the ER go to other compartments, the plasma membrane supplies them with lipids that are synthesized in the ER (Blom *et al.*, 2011). The ER stores calcium (Ca^{+2}) and regulates the intracellular levels of the ion (Koch, 1990) which is an integral aspect of many cellular signaling pathways. The ER also initializes the process of glycosylation which is perfected in the Golgi before the secretion of proteins from the latter (Aebi, 2013). Disulfide bond formation is an essential component of protein folding, stability and multimerization in the extracellular environment (Thangudu *et al.*, 2008). Similar to the bacterial periplasm oxidative protein folding in eukaryotes occurs in the ER (Herrmann and Riemer, 2014). These processes form the core set of functions performed by the ER.

The Golgi which communicates with the ER by means of COPI and COPII coated vesicles acts like a sorting station for proteins destined for other internal compartments and secretion (Watson and Stephens, 2005; Brandizzi and Barlowe, 2013). It is also the major site for O-linked and N-linked glycosylation, which is required for stability of proteins that are secreted or membrane proteins that are exposed to the environment (Wilson *et al.*, 2011; Stanley, 2011; Reynders *et al.*, 2011). Functionally the Golgi is subdivided into the cis, medial, and trans Golgi, wherein the proteins destined for other compartments (and secretion) bud off (Glick and Nakano, 2009). Exocytosis is the process of fusion and/or secretion of vesicles from within the cell to the plasma membrane, resulting in release of the contents of the vesicles into the environment (Wu *et al.*, 2014). Since continuous exocytosis will result in increased surface area of the cell, exocytosis is balanced with endocytosis in order to maintain cell size and shape (Gauthier *et al.*, 2009; Doherty and McMahon, 2009). Endocytosis of small vesicles containing membrane proteins and small cargo from the environment is processed in a series of transient compartments of the endosomal machinery from where proteins can either go back to the plasma membrane or be sent for degradation (Doherty and McMahon, 2009). Protein degradation in the eukaryotic cytosol typically occurs in the lysosome (Luzio *et al.*, 2007). The lysosome is acidified by means of a vacuolar–ATPase (V-type ATPase) which is homologous to the archaeal plasma membrane ATPase (Mulkidjanian *et al.*, 2007). It hydrolyses ATP to pump protons into the lysosome and acidify the lumen of lysosomes (De Hertogh *et al.*, 2004; Mindell, 2012; Marshansky *et al.*, 2014). This activates acid proteases that proceed with degradation of lysosomal contents. An acidified lysosome allows the cell to compartmentalize its proteases thereby preventing unwanted proteolysis of its cytoplasmic contents. The acidified lysosome thus forms a very essential component of the endomembrane system. Phagocytosis which is distinct from endocytosis, is used by some organisms or cell (in case of multicellular eukaryotes) to engulf large food particles or organisms in an actin dependent process (Freeman and Grinstein, 2014). The phagosome thus formed fuses with an acidified lysosome wherein the contents of the phagosome are degraded and exported as energy-rich small molecules for metabolism in the cytosol. In addition to these structures, autophagosomes and multivesicular bodies are sometimes generated as a quality control mechanism for the cytosol as and when the need arises (Fader and Colombo, 2009). All these vesicles stem from the ER and eventually find their way to other compartments. With continuous connections to the nucleus and the mitochondria (Lang *et al.*, 2015; Voeltz *et al.*, 2002) the eukaryotic endomembrane system forms the central hub of a typical

eukaryotic cell.

The interconnected and interdependent nature of the endomembrane system hinders the independent evolution of one sub-compartment without others. It is therefore imperative that when one invokes a prokaryote with the ability to perform phagocytosis (as the host that acquired the α-proteobacterial endosymbiont), they must invoke a prokaryote with an endomembrane system. That, however, takes us back to a eukaryotic cell without a mitochondrion. A core assumption of all models that envision the evolution of a complex host before the entry of the endosymbiont, is that the ER evolved in a prokaryote and was performing all the said functions. A caveat in doing so is that until now no reasoning is given to the selection pressure that would drive a prokaryote to evolve a functional ER given the fact that prokaryotes perform all of its functions. Being an energetically expensive compartment, if the ER does not provide an immense evolutionary advantage to a prokaryote, the question arises then what it would have been selected for during evolution? While solving evolutionary problems, more often than not, we are compelled to ask why would a cell evolve a particular process/trait/mechanism, rather than how. Finding the selection pressures that led to the development of these processes or traits can help discriminate between the plausible and the probable scenarios in evolution.

Gould *et al.*, 2016 explores an hitherto underappreciated aspect of prokaryotic biology overlooked when addressing eukaryogenesis and the origin of the eukaryotic endomembrane system and is the topic of publication II

Inventing protein import

That mitochondria are the result of the endosymbiosis of a bacterium has been long since agreed upon (Gray and Doolittle, 1982). Biochemists had realized early on, however, that the genomes of these organelles encoded only some but not all proteins of the organelle (Borst, 1972; Whitfeld and Bottomley, 1983) and hence depended on cytoplasmic protein synthesis (Schatz and Mason, 1974; Bottomley and Bohnert, 1982; Ellis, 1981). By the 1980s, it was known, that precursor protein translocation across microsomes or the ER occurred co-translationally via the SEC61 complex, through the recognition of a signal sequence by a signal recognition particle (SRP) (Gilmore *et al.*, 1982; Walter *et al.*, 1984; Meyer *et al.*, 1982; Kreil, 1981; Matlin, 2013). This was however not the case for organellar protein translocation, which was shown to be post-translational. It had become known that organellar proteins are synthesized as precursors, too, but unlike ER protein translocation, they were translocated into their respective organelles post-translationally (Chua and Schmidt, 1978; Maccecchini *et al.*, 1979). Targeting to the organelles was mediated by N-terminal targeting sequences (NTSs) that associated with receptor platforms on the outer membrane of the organelle (Karlin-Neumann and Tobin, 1986; von Heijne, 1986). A complete picture of the receptor platform is still not available as newer proteins get implicated in organelle targeting and translocation as experimental tools get refined (Neupert, 2015; Di Maio *et al.*, 2016; van Dooren *et al.*, 2016; Murcha *et al.*, 2006).

The plastid and mitochondrial translocation machineries share similarities (Neupert, 2015; Becker *et al.*, 2012; Dudek *et al.*, 2013; Soll and Schleiff, 2004; Chacinska *et al.*, 2009; Lister *et al.*, 2005; Doležal *et al.*, 2006). In both plastids and mitochondria protein translocation is facilitated by a complex of proteins called the translocase of the outer membrane of chloroplasts or the translocase of the outer membrane of mitochondria (TOC and TOM, respectively) (Schleiff and Becker, 2010). The TOC and TOM complexes are composed of a central translocase, the Toc75 and the Tom40 respectively, which are associated with other accessory proteins that aid in signal recognition and binding. Toc34, Toc64 and Toc159 in the plastids and Tom20,Tom70 and Tom22 in the mitochondria are some of the more conserved proteins of the TOC and TOM complexes with smaller proteins that help in formation and maintenance of the complex itself (Neupert, 2015; Bölter and Soll, 2016). The core translocases (Toc75 and Tom40) of the TOC and TOM complexes are thought to be derived from bacterial transporter proteins (Arnold *et al.*, 2007; Remmert *et al.*, 2010; Zeth and Thein, 2010) which is more clear for plastid Toc75 (Hsu and Inoue, 2009; Schleiff *et al.*, 2011; Voulhoux and Tommassen, 2004; Richardson *et al.*, 2014), than for mitochondrial Tom40 (Doležal *et al.*, 2006; Lithgow and Schneider, 2010). Given the bacterial ancestry of both organelles it is likely that protein targeting to the organelles share some similarities to bacterial protein secretion (Doležal *et al.*, 2006; Tong *et al.*, 2011; Perry *et al.*, 2006; Lithgow and Schneider, 2010).

The more recent acquisition of the plastid, relative to the mitochondria, might be a contributing factor for the ease of finding evolutionary relationships of the plastid import machinery. In case of proteins destined for the plastid stroma or the mitochondrial matrix, import proceeds via the translocase of the inner membrane of chloroplasts (TIC) or the translocase of the inner membrane of the mitochondria (TIM) complexes (Neupert, 2015; Bölter and Soll, 2016). As in the case for TOC and TOM, the TIC and TIM have several accessory proteins aiding a central translocase. In addition to soluble proteins, separate systems such as the OEP80 and SAM complex, are essential for integration of membrane proteins in the plastids and mitochondria respectively.

Similar to the translocases the N-terminal targeting sequences (NTSs) that targets proteins to the plastids (pNTS) or the mitochondria (mNTS) have some unifying features (Schleiff and Becker, 2010; Steppuhn and Herrmann, 1989). They both form amphiphilic helices, are mainly in the N-terminus and cleaved off by independent and organelle specific signal processing proteases for maturation. The mitochondrial targeting sequences, however, are highly enriched in arginine and lysine while their plastid counterparts are enriched in serine, at least in the case of green algae and embryophytes (Steppuhn and Herrmann, 1989; Franzén *et al.*, 1989; von Heijne and Abrahmsèn, 2001). Some of these serine residues are known to be phosphorylated and subsequently bound by a guidance complex, a 14-3-3 and Hsp70 protein, before binding to receptor proteins of the TOC complex (May and Soll, 2000). While in both cases the import of proteins into the organelle is energy-dependent (Shi and Theg, 2013; Soll and Schleiff, 2004; Geissler *et al.*, 2001), in case of aerobic mitochondria the import of proteins; especially into the matrix; is additionally dependent on the transmembrane potential generated by the electrochemical gradient ($\Delta\Psi$), which is a result of the electron transport chain that is coupled to oxidative phosphorylation (Martin *et al.*, 1991; Geissler *et al.*, 2000).

Hydrogenosomes like those of *Trichomonas vaginalis* were initially thought to be autonomous compartments, like the the peroxisome, but specialized for hydrogen production (Müller, 1973). However,

genetic and biochemical evidence pointed towards hydrogenosomes having a bacterial ancestry. The protein import machinery of hydrogenosomes, and later confirmed for mitosomes, were homologous the import machineries of mitochondria (Burri *et al.*, 2006; Van der Giezen, 2009; Lithgow and Schneider, 2010). Hydrogenosomes and mitosomes have homologs of the core outer membrane translocase Tom40, a few accessory proteins such as Tom70 and Tom20 and the translocase of the inner membrane (Doležal *et al.*, 2006). However, the complete machinery for import is highly reduced, both in terms of number and complexity of individual subunits (Doležal *et al.*, 2006; Burri *et al.*, 2006). Hydrogenosomes lack the electron transport chain for oxidative phosphorylation and hence an electrochemical gradient. ATP production in hydrogenosomes occurs via substrate level phosphorylation (Müller *et al.*, 2012). Mitosomes on the other hand do not produce ATP at all rather they import itvfor iron-sulfur cluster biogenesis (Goldberg *et al.*, 2008).

Given that hydrogenosomes and mitosomes are evolutionarily related to the mitochondria, the reduced import machinery has two evolutionary implications. (i) The reduction of the organelles themselves allowed for convergent evolution of simpler mechanisms of protein import or (ii) the protein import machinery has reverted to an evolutionarily conserved ancestral state. The diversity of the receptor platform (like the Tom 22) associated with the core conserved set of proteins seem to suggest that the latter is a more plausible scenario (Perry *et al.*, 2006; Clements *et al.*, 2009). The evolutionary reduction of import machineries sheds some light on the selection pressures involved in the evolution and maintenance of the machinery.

In a series of papers central to this thesis (publications III, IV and V), it has been proposed that a loss of the electrochemical gradient ($\Delta\Psi$) in organelles such as the hydrogenosomes and mitosomes is coupled to the loss of positively charged N-termini in proteins that are imported into the organelle (Garg *et al.*, 2015). A positively charged N-terminus is essential for mitochondrial protein import (Martin *et al.*, 1991) while negatively charged N-terminus can be detrimental (McBride *et al.*, 1995). This has major implications regarding relocation of metabolic pathways and evolution of eukaryotic metabolism (Rada *et al.*, 2015; Martin, 2010). The selection pressures involved in the origin and maintenance of positively charged NTSs for protein import into mitochondria has implications for the evolution of protein import into primary and secondary plastids, too (Garg and Gould, 2016)

Origin of the cell cycle and sex

The differences that separate eukaryotes from prokaryotes involve not only structures such as mitochondria or the endomembrane system, there are also processes that differ acrosss the prokaryote-eukaryote divide. Among those processes , none is more significant than sex. Eukaryotes have it and prokaryotes do not. The problem concerning the origin of eukaryotic sex has been addressed in various forms over the years from the standpoint of population genetics (Muller, 1964; Felsenstein, 1974; Kondrashov, 1988), a developmental standpoint (Kondrashov, 1993; 1997; Kondrashov and Crow, 1991), from the anisogamic cost of sex (Williams, 1975; Maynard Smith, 1978; Lewis Jr, 1987) to the benefits of recombination (Cavalier-Smith, 2002; Bernstein and Bernstein, 2010; Hörandl and Hadacek, 2013) and numerous other aspects of sexual reproduction (Bell, 1982; Margulis and Sagan, 1986; Bell, 1988; Bernstein and Bernstein, 1991; Perrot *et al.*, 1991; Hutson and Law, 1993; Lynch *et al.*, 1993; Allen, 1996; Barton and Charlesworth, 1998; Solari, 2002; Otto, 2009; Feigel *et al.*, 2009; Hörandl, 2009; Cavalier-Smith, 2010; Gross and Bhattacharya, 2010; Goodenough and Heitman, 2014; Havird *et al.*, 2015; Heitman, 2015; Speijer, 2016). However sex has has rarely been tied to the origin of eukaryotes.

The unique cell cycle of eukaryotes and its importance in sexual reproduction tie them together with eukaryote origins. A cell cycle in its broadest sense entails the division of the genetic material followed by the division of the cell. In multicellular organisms division of the entire individual is called the life-cycle. The immense number of variations in achieving the cell- or life-cycle is tied into the variation we see in life (Raikov, 1994). The nature of differences between the DNA recombination in prokaryotes and eukaryotes relates to the localization of DNA within the cell.

The organization, duplication, division, and storage of DNA inside the nucleus differs significantly between prokaryotes and eukaryotes. In contrast to almost all prokaryotes, eukaryotes have linear chromosomes and more often than not their genome is fragmented in multiple chromosomes which replicate with multiple origins of replication (Masai *et al.*, 2010; Toro and Shapiro, 2010). This complicates segregation, because it has to be ensured that all chromosomes are segregated equally. This is compounded by the fact that eukaryotes have a nucleus, which separates the genome from the cytoplasm. In prokaryotes the genetic material is organized in a circular molecule of DNA in the majority of cases, which lies naked in the cytosol and transcription, translation and replication occur concurrently with cell division in the cytosol (Cooper, 2006; Reyes-Lamothe *et al.*, 2012; Egan and Vollmer, 2013). Distinct phases of prokaryotic chromosome and cell division akin to those in the mitotic cell division of eukaryotes do not exist in prokaryotes (Cooper, 2006). It has been argued that bacteria like the *Caulobacter* or archaea of the TACK superphylum have a cell cycle (Biondi *et al.*, 2006; Lindås and Bernander, 2013), but the eukaryotic cell cycle is unique and distinct with phases of chromosome and cell-division and regulatory checkpoints that separate the two (Norbury and Nurse, 1992; Murray, 2004). To state it clearly, eukaryote-like chromosome condensation during cell division is lacking in prokaryotes.

Briefly, a typical eukaryotic cell cycle is divided into two major stages the interphase and the mitotic or M-Phase. The interphase is further separated into synthesis (S) phase during which the genome is replicated and the gap (G) phases. At the end of a successful cell cycle, a cell has faithfully replicated its genome and separates to form two new daughter cells with exactly the same genome as the mother cell (Norbury and Nurse, 1992; Alberts, 2008; Lodish, 2008). Many variations in this theme exists with the presence (closed mitosis) or absence (open mitosis) of the nuclear envelope like in closed and open mitosis being one of them (Sazer *et al.*, 2014; Boettcher and Barral, 2014) and others like cell senescence or specialization the cells enter into a stage in which they stop dividing or cannot divide anymore. The transition from one cell cycle phase to another is a highly controlled process involving cyclin dependent kinases (CDKs) which are temporally regulated by their respective cyclins (Morgan, 1997). These act as checkpoints that ensure faithful chromosome replication and segregation followed by cell division (Hartwell and Weinert, 1989). The mechanistic similarities and differences are far too many to list here (reviewd in Bendich and Drlica, 2000), and they illustrate a very crucial point: cell division was one of the most important and difficult transitions in eukaryote origins and more so than others since, cell division is imperative for survival for without cell division, there are no progeny upon which natural selection can act.

In the most modern rendition of Darwin's theory of evolution by natural selection, mutations in all forms (spontaneous or adaptive, point mutations, chromosome additions, deletions, inversions, partial genome duplications) generates variation upon which selection acts to ensure the survival of individuals most suited for the environment as measured by their fitness (Nei, 2013). Since evolution does not have foresight, the

mutations can be both beneficial and detrimental to the individual, which contributes to the slow pace of evolution. Frequently however, the strength of the mutation may not be severe enough to affect fitness of the individual immediately and thus may survive in the population. An accumulation of such mutations can drive the population towards extinction. This is the principle behind Muller's ratchet (Muller, 1964; Crow, 2005). The solution as postulated by Muller is recombination, which in eukaryotes is realized through sex. Muller's ratchet is generally considered the main reason for the evolution and persistance of eukaryotic sex but, Muller's ratchet is not unique to eukaryotes. Prokaryotes can also be affected by Muller's ratchet, and they escape it through mechanisms of lateral gene transfer, which include transformation, transduction and conjugation (Takeuchi *et al.*, 2014). Eukaryotic sex however is the product of syngamy or fusion of gametes which is a consequence of meiotic division of cells (Solari, 2002).

Origin of sex thus entails the lack of prokaryotic like recombination mechanisms in eukaryotes and the lack of meiotic processes in prokaryotes. Meiosis is deeply embedded in mitosis and results in a reduction of ploidy, which is subsequently restored by syngamy. However, the advantage of meiosis comes from recombination that happens during the phase of chromosome segregation. It is known now from biochemical and bioinformatic studies that the enzymes mediating this recombination are homologous to prokaryotic enzymes (Camerini-Otero and Hsieh, 1995; Ramesh *et al.*, 2005; Hörandl and Hadacek, 2013; Speijer *et al.*, 2015). This has lead to many prominent theories that address the evolution of meiosis from mitosis, which unequivocally assume the presence of a mitosing eukaryote lacking sex (Cleveland, 1947; Heywood and Magee, 1976; Hurst and Nurse, 1991; Maguire, 1992; Kondrashov, 1994; Villeneuve and Hillers, 2001; Solari, 2002; Wilkins and Holliday, 2008; Bernstein and Bernstein, 2010; Hörandl and Hadacek, 2013; Markov and Kaznacheev, 2016; Loidl, 2016). The last publication (publication VI) of this thesis addresses the origin of the eukaryotic cell cycle and sex.

Publication I

Endosymbiotic theories for eukaryote origin

William F. Martin[1], **Sriram Garg**[1] and Verena Zimorski[1]

1 - Institute for Molecular Evolution, Heinrich-Heine Universität Düsseldorf, Universitätsstraße 1, Düsseldorf 40225, Germany

The presented manuscript was peer-reviewed and published as part of a theme issue on *"Eukaryotic origins:progress and challenges"* compiled and edited by T.Martin Embley and Tom A. Williams in **Philosophical Transactions of the Royal Society of London: Biological Sciences** with an impact factor of **7.224** on 6th May 2015. It was also featured in the Cover

Contribution as **second author**
33.33%
Illustrated and designed Figures 1,2,3 and 4 and contributed towards writing parts of the text including mining for literature.

PHILOSOPHICAL TRANSACTIONS B

rstb.royalsocietypublishing.org

Review

Cite this article: Martin WF, Garg S, Zimorski V. 2015 Endosymbiotic theories for eukaryote origin. *Phil. Trans. R. Soc. B* **370**: 20140330. http://dx.doi.org/10.1098/rstb.2014.0330

Accepted: 6 May 2015

One contribution of 17 to a theme issue 'Eukaryotic origins: progress and challenges'.

Subject Areas:
evolution, cellular biology, plant science

Keywords:
endosymbiosis, eukaryotes, nucleus, mitochondria, plastids, anaerobes

Author for correspondence:
William F. Martin
e-mail: bill@hhu.de

THE ROYAL SOCIETY PUBLISHING

Endosymbiotic theories for eukaryote origin

William F. Martin, Sriram Garg and Verena Zimorski

Institute for Molecular Evolution, Universität Düsseldorf, Universitätsstraße 1, Düsseldorf 40225, Germany

For over 100 years, endosymbiotic theories have figured in thoughts about the differences between prokaryotic and eukaryotic cells. More than 20 different versions of endosymbiotic theory have been presented in the literature to explain the origin of eukaryotes and their mitochondria. Very few of those models account for eukaryotic anaerobes. The role of energy and the energetic constraints that prokaryotic cell organization placed on evolutionary innovation in cell history has recently come to bear on endosymbiotic theory. Only cells that possessed mitochondria had the bioenergetic means to attain eukaryotic cell complexity, which is why there are no true intermediates in the prokaryote-to-eukaryote transition. Current versions of endosymbiotic theory have it that the host was an archaeon (an archaebacterium), not a eukaryote. Hence the evolutionary history and biology of archaea increasingly comes to bear on eukaryotic origins, more than ever before. Here, we have compiled a survey of endosymbiotic theories for the origin of eukaryotes and mitochondria, and for the origin of the eukaryotic nucleus, summarizing the essentials of each and contrasting some of their predictions to the observations. A new aspect of endosymbiosis in eukaryote evolution comes into focus from these considerations: the host for the origin of plastids was a facultative anaerobe.

1. Introduction

Early evolution is an important part of life's history, and the origin of eukaryotes is certainly one of early evolution's most important topics, as the collection of papers in this special issue attests. There are various perspectives from which eukaryote origins can be viewed, including palaeontological evidence [1], energetics [2], the origin of eukaryote-specific traits [3,4] or the relationships of the different eukaryotic groups to one another [5]. This paper will look at eukaryote origins from the standpoint of endosymbiotic theory, and how different versions of endosymbiotic theory tend to square off with the data that we have for eukaryotic anaerobes and with regard to data from gene phylogenies. Endosymbiotic theory has a long and eventful history, virtuously summarized in Archibald's book [6], and speaking of history, here is a good place to dispel a myth—about Altmann.

One can occasionally read (though we will politely provide no examples) that Altmann [7] is to be credited with the idea of symbiotic theory for the origin of mitochondria, but that is incorrect. Those of us who can read German and who have a copy of Altmann's 1890 book can attest: in the 1890 book, Altmann was not interested in mitochondria, and he did not propose their symbiotic origin. He mentioned neither mitochondria (nor their older name, chondriosomes) nor endosymbiosis in his book on 'bioblasts'. To Altmann, everything in eukaryotic cells consisted of bioblasts, including the cytosol, the nucleus and the chromosomes. His bioblasts corresponded to a chemical organization state of matter that was larger than the molecule but smaller than the cell 'the smallest morphological unit of organized material' ('*die kleinste morphologische Einheit der organisirten Materie*') [8, p. 258]. They would maybe correspond in size roughly to what we today call macromolecular complexes, which however cannot be seen in the light microscopes of Altmann's day. He also distinguished autoblasts,

cytoblasts, karyoblasts and somatoblasts, which are mentioned far less often than bioblasts. A scholarly treatise of Altmann in the context of symbiotic theory, and why he cannot be credited with having suggested endosymbiotic theory, can be found in Höxtermann & Mollenhauer [8].

The concept of symbiosis (Latin, 'living together'), that two different organisms can stably coexist and even give rise to a new type of organism, traces to Simon Schwendener [9], a Swiss botanist who discovered that lichens consist of a fungus and a photosynthesizer. The German botanist Heinrich Anton de Bary (1878) coined the term '*Symbiose*' to designate this type of coexistence [10]. Schimper [11] is sometimes credited with the discovery of endosymbiotic theory, but his treatise of the topic is wholly contained in a footnote that translates to this: 'If it can be conclusively confirmed that plastids do not arise de novo in egg cells, the relationship between plastids and the organisms within which they are contained would be somewhat reminiscent of a symbiosis. Green plants may in fact owe their origin to the unification of a colorless organism with one uniformly tinged with chlorophyll' [11, pp. 112–113]. That was all he wrote on the possibility of symbiotic plastid origin. The sentence immediately following that one in Schimper's famous footnote, however, is also significant, as we will see in a later passage about Portier and the symbiotic origin of mitochondria; it translates to this: 'According to Reinke (Allg. Botanik, p. 62) the chlorophyll bodies [Chlorophyllkörner, another name for plastids in Schimper's day] might even have the ability to live independently; he observed this phenomenon, as communicated to me, and published with kind permission, in a rotting pumpkin, the chloroplastids of which, surrounded by Pleosporahyphae, continued to vegetate in dead cells and multiplied by division' [11, p. 113]. Clearly, Reinke was observing the proliferation of contaminating bacteria, not of free-living organelles.

Schimper [11,12] did, however, champion the case that plastids proliferate through division. That was important for the Russian biologist Constantin Mereschkowsky, who probably delivered the first thoroughly argued case that some cells arose through the intracellular union of two different kinds of cells (endosymbiosis), in his 1905 paper [13] that has been translated into English [14]. Mereschkowsky [13] said three things: (i) plastids are unquestionably reduced cyanobacteria that early in evolution entered into a symbiosis with a heterotrophic host, (ii) the host that acquired plastids was itself the product of an earlier symbiosis between a larger, heterotrophic, amoeboid host cell and a smaller 'micrococcal' endosymbiont that gave rise to the nucleus, and (iii) the autotrophy of plants is an inheritance, *in toto*, from cyanobacteria [13].

Mereschkowsky's scheme was more fully elaborated but basically unchanged in his 1910 series [15]: there were two kinds of fungi, those that evolved a nucleus without endosymbiosis and those that once possessed plastids but became secondarily non-photosynthetic, today we call them the oomycetes, and there is still no consensus on the issue of whether they ever had plastids or not. The branches in Mereschkowsky's tree occasionally unite via endosymbiosis to produce fundamentally and radically new kinds of organisms (plants, for example) [15,16]. A more modern version of symbiosis in cell evolution would have to include the symbiotic origin of mitochondria, archaea and the concept of secondary endosymbiosis. Endosymbiotic theories have it that cells unite, one inside the other, during evolution to give rise to novel lineages at the highest taxonomic levels, via combination. That is not the kind of evolution that Darwin had in mind; his view of evolution was one of gradualism.

Many biologists still have a problem with the notion of endosymbiosis and hence prefer to envisage the origin of eukaryotes as the product of gene duplication, point mutation and micromutational processes [17]. A 2007 paper by the late Christian de Duve [18] is now often taken as the flagpole for micromutational theories of eukaryote origin, but de Duve, like the late Lynn Margulis [19], always categorically rejected the evidence that mitochondria and hydrogenosomes—anaerobic forms of mitochondria [20,21]—share a common ancestor. No anaerobic form of mitochondria ever fits into classical endosymbiotic theory. This is because classical (Margulis's version of) endosymbiotic theory [19] was based on the premise that the benefit of the endosymbiotic origins of mitochondria was founded in oxygen utilization, while de Duve's versions went one step further and suggested that even the endosymbiotic origin of peroxisomes was founded in oxygen utilization [18]. Anaerobic mitochondria were never mentioned and hydrogenosomes, if they were mentioned, were explained away as not being mitochondria [18,19]. The overemphasis of oxygen in endosymbiotic theory and how the focus on oxygen led to much confusion concerning the phylogenetic distribution and evolutionary significance of anaerobic forms of mitochondria has been dealt with elsewhere [22–24].

There is one alternative to classical endosymbiotic theory that took anaerobic mitochondria and hydrogenosomes into account, the hydrogen hypothesis [25]; it predicted (i) all eukaryotes to possess mitochondria or to have secondarily lost them, (ii) that the host for mitochondrial origins was an archaeon, the eukaryotic state having arisen in the wake of mitochondrial origins, and (iii) that aerobic and anaerobic forms should interleave on the eukaryotic tree. Though radical at the time, prediction (i) was borne out [26–29], and so was prediction (ii) [30–32], as well as (iii) [21,33]. Furthermore, only recently, it has been recognized that the invention of eukaryotic specific traits required more metabolic energy per gene than prokaryotes have at their disposal, and that mitochondria afforded eukaryotic cells an orders of magnitude increase in the amount of energy per gene, which (finally) explains *why* the origin of eukaryotes corresponds to the origin of mitochondria [2,34]. But there is more to eukaryote origins than just three predictions and energy. There is the origin of the nucleus to deal with [35], and the role that gene phylogenies have come to play in the issues. In addition, there is the full suite of characters that distinguish eukaryotes from prokaryotes to consider (meiosis, mitosis, cell cycle, membrane traffic, endoplasmic reticulum (ER), Golgi, flagella and all the other eukaryote-specific attributes, including a full-blown cytoskeleton—not just a spattering of prokaryotic homologues for cytoskeletal proteins [31]), but here our focus is on endosymbiotic theories, not the autogenous origin of ancestrally shared eukaryotic characters, whose origins for energetic reasons come in the wake of mitochondrial origin [34].

2. Gene trees, not as simple as it sounds

To get a fuller picture of eukaryote origins, we have to incorporate lateral gene transfer (LGT) among prokaryotes, endosymbiosis and gene transfer from organelles to the nucleus into the picture. That is not as simple as it might

seem, because it has become apparent that individual genes have individual and differing histories. Thus, in order to get the big picture, we would have to integrate all individual gene trees into one summary diagram in such a way as to take the evolutionary affinities of the plastid (a cyanobacterium), the mitochondrion (a proteobacterium) and the host (an archaeon) into account. Nobody has done that yet, although there are some attempts in that direction [36]. In 2015, our typical picture of eukaryotic origins entails either a phylogenetic tree based on one gene or, more commonly now, a concatenated analysis of a small sample of genes (say 30 or so from each genome), which generates a tree, the hope being that the tree so obtained will be representative for the genome as a whole and thus will have some predictive character for what we might observe in phylogenies beyond the 30 or so genes used to make the tree. The 30 or so genes commonly used for such concatenated phylogenies are mostly ribosomal proteins or other proteins involved in information processing, genes that Jim Lake called informational genes in 1998 [37].

But because of the role of endosymbiosis in eukaryote cell evolution, eukaryotes tend to have two evolutionarily distinct sets of ribosomes (archaeal ribosomes in the cytosol and bacterial ribosomes in the mitochondrion), or sometimes three (an additional bacterial set in the plastid [38]) and in rare cases four sets of active ribosomes (yet one more set in algae that possess nucleomorphs) [39]. The 'core set of genes' approach, in all of its manifestations so far, only queried cytosolic ribosomes for eukaryotes, and thus only looked at the archaeal component of eukaryotic cell history. Some of us have been worried that by looking only at genes that reflect the archaeal component of eukaryotic cells we might be missing a lot, because it was apparent early on that many genes in eukaryotes do not stem from archaea, but from bacteria instead and, most reasonably under endosymbiotic theory, from organelles [40,41].

An early study looking at the phylogeny of the core gene set, which largely but not entirely corresponds to the ribosomal protein superoperon of prokaryotes, came to the conclusion that the information contained within the alignment is problematic because of the low amount of sequence conservation involved across many of the sites [42]. Concerns were also voiced that the 30 genes of the set, if analysed individually, might not have the same history and that concatenation might thus be a problem [43], but that did not stop bioinformaticians [44] from rediscovering the same set of 30 or so genes and making a tree that looked remarkably similar to the rRNA tree in most salient aspects, in particular as regards the position of the eukaryotes. By that time it was reasonably well-known that the genes of archaeal origin in eukaryotes are not representative of the genomes as a whole; they constitute a minority of the genome and are vastly outnumbered by genes of bacterial origin [45]. Despite that, the attention in the issue of eukaryote origins has, with few exceptions [46–48], remained focused on the archaeal component, and it will probably stay that way until improved methods to summarize the information contained in thousands of trees come to the fore.

Always critical of the branches in trees that phylogenetic methods produce [49], Embley and colleagues looked at the conserved core set with more discerning phylogenetic methods [30,50,51] and found that the archaeal component of eukaryotes branches within the archaea. These new trees tend to group the eukaryotes with the crenarchaeotes, specifically with the TACK superphylum of archaea [31], while at the same time tending to locate the root of the archaea among the euryarchaeotes, sometimes among the methanogens [52].

Now is a good time to have a look at endosymbiotic theories and related ideas for the origin of eukaryotes, their nucleus and their mitochondria. In doing so, we pick up on our own earlier reviews of the topic [22,53], the figures of which have become popular [31]. In the next section, we summarize what various models say, starting with models for the origin of the nucleus, and then move on to models for the origins of chloroplasts and mitochondria.

3. The nucleus

The nucleus is a defining feature of eukaryotes [54]. Theories for the evolution of the nucleus are usually based (i) on invaginations of the plasma membrane in a prokaryote or (ii) on endosymbiosis of an archaeon in a eubacterial host or (iii) on an autogenous origin of a new membrane system including the nuclear envelope in a host of archaeal origin after acquisition of mitochondria. The endosymbiotic theory for the origin of the nucleus started with Mereschkowsky [13]. He postulated that the nucleus evolved from a prokaryote (mycoplasma), which was engulfed by an amoeboid cell homologous to the eukaryotic cytosol (figure 1*a*; [15]).

Cavalier-Smith argued that nuclear and ER membranes originated through invaginations of the plasma membrane of a prokaryotic cell (figure 1*b*; [55–58]). He suggested that the prokaryote initially lost its cell wall and thereby gained the ability to phagocytose food particles. Ribosomes, primarily attached to the plasma membrane, became internalized, but still attached to the membrane, resulting first in the rough ER and out of it the nuclear envelope. Gould & Dring [59] presented a different model in 1979 where they described that endospore formation of Gram-positive bacteria resulted in the origin of the nucleus. The protoplast of a single cell divides during endospore formation in such a manner that the cell engulfs a portion of its own cytoplasm, which than becomes surrounded by a double membrane resulting in the cell's nucleus (figure 1*c*; [59]). In the 1990s, several models for the origin of the nucleus via endosymbiosis (sometimes called endokaryotic theories) were published, but only few refer to Mereschkowsky's original suggestion. They have in common that they envisage a eubacterial host that engulfed an archaebacterial endosymbiont that underwent a transformation into the nucleus (figure 1*d*; [60,61]). Fuerst & Webb [62] observed that the DNA in the freshwater budding eubacterium *Gemmata obscuriglobus* (a member of the *Planctomyces-Pirella* group) appears to be surrounded by a folded membrane, the organization of which was thought to resemble the nucleus (figure 1*e*; [62]). Later papers were less cautious and called this structure a nucleus outright [63]; subsequent work on *Gemmata* showed that the inner membrane is simply an invagination of the plasma membrane [64], as had been previously pointed out [53]. Searcy & Hixon [65] interpreted thermophilic acidophilic sulfur-metabolizing archaebacteria lacking a rigid cell wall but having a well-developed cytoskeleton as a primary stage for the evolution of eukaryotic cells (figure 1*f*; [65]).

Lake & Rivera [66] suggested an endosymbiosis in which a bacterium engulfed an archaeon (crenarchaeon) for the origin of eukaryotes (figure 1*g*). A vesicular model for the origin of the nucleus in a cell that had a mitochondrial endosymbiont

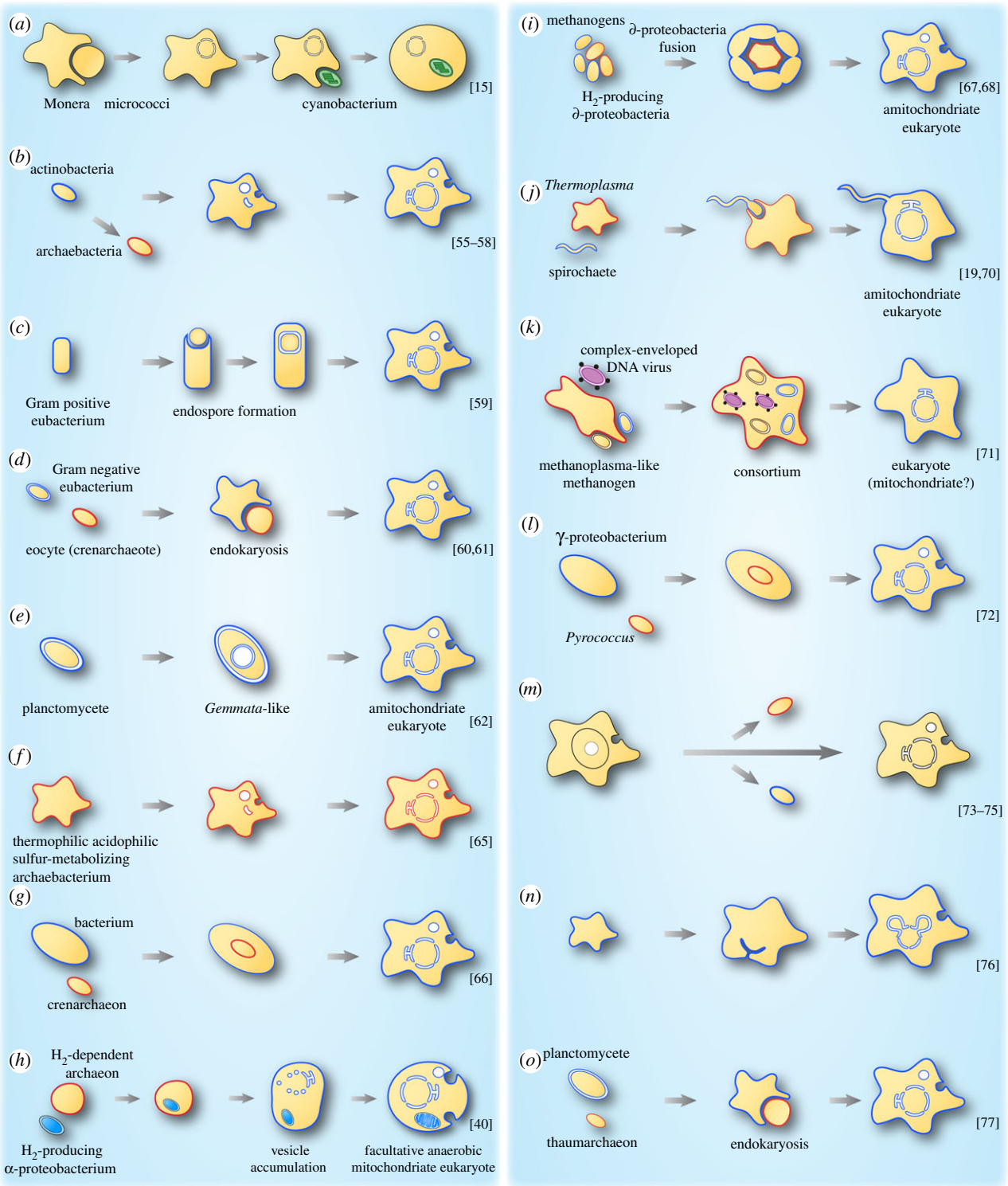


Figure 1. Models describing the origin of the nucleus in eukaryotes. (*a–o*) Schematic of various models accounting for the origin of the nucleus. Archaeal cells/membranes are represented with red, while blue indicates eubacterial cells/membranes. Black membranes are used when the phylogenetic identity of the cell is not clear or not specified. See also [22,53].

was proposed (figure 1*h*; [40]). It posits a role for gene transfer and the origin of bacterial lipids in the origin of the eukaryotic endomembrane system, and in a subsequent formulation [35] it posits a causal relationship between the origin of spliceosomes and the origin of nucleus–cytosol compartmentation (this aspect is discussed in more detail in a later section). Moreira & López-García [67,68] modified the endokaryotic model, invoking the principle of anaerobic syntrophy (H_2-dependence) for the origin of the nucleus. They postulated a fusion of plasma membranes in an agglomeration of δ-proteobacteria entrapping a methanogenic archaebacterium, which evolved to the nucleus (figure 1*i*; [67,68]). The kind of fusion of plasma membranes among free-living cells that Moreira & Lopez-Garcia [67,68] envisage has not been observed for bacteria, but it is known to occur among archaea [69]. Lynn Margulis presented another symbiogenic theory for the origin of the nucleus. She suggested a symbiosis between a spirochaete and an archaebacterium without a cell wall (most likely *Thermoplasma*-like in her view), leading to both the eukaryotic flagellum and the nucleus (figure 1*j*; [19,70]). A viral origin for the nucleus involving poxviruses was suggested in 2001 by Bell in the context of syntrophic consortia involving methanogens (figure 1*k*; [71]). Horiike postulated a model in which the nucleus emerged from an

archaeal endosymbiont (*Pyrococcus*-like), which was engulfed by a γ-proteobacterium (figure 1*l*; [72]). An origin of eukaryotes (hence implicitly or explicitly their nucleus) prior to prokaryotes has also been repeatedly suggested (figure 1*m*; [73–75]). Penny argues that prokaryotes, which he and Forterre [73] sometimes call 'akaryotes' [75], arose from this eukaryote ancestor via Forterre's thermoreduction hypothesis—a transition to the prokaryotic state from a eukaryotic ancestor in response to higher temperatures.

More recently, the community of scientists interested in cytoskeletal evolution have—in unaltered form—rekindled Cavalier-Smith's hypothesis of an autogenous (non-symbiotic) origin of a phagocytosing amitochondriate eukaryote (an archezoon) via point mutational changes leading to a host that does not need a mitochondrion at all to enjoy its phagocytotic lifestyle, but acquires one nonetheless (figure 1*n*; [76]).

Forterre [77] departed from thermoreduction and introduced a new variant of the endokaryotic hypothesis, one that got planctomycetes (a member of the PVC group: Planctomycetes, Verrucomicrobia, Chlamydiae) involved in eukaryote origin as the bacterial host for the engulfment of a thaumarchaeon as the nucleus, followed by invasions of retroviruses and nucleo-cytoplasmic large DNA viruses (NCLDV). In this theory, the PTV (for PVC–thaumarchaeon–virus) fusion hypothesis, the PVC bacterium provides universal components of eukaryotic membranes required also for the formation of the nucleus and the thaumarchaeon provides informational and operational proteins and precursors of the modern eukaryotic cytoskeleton and vesicle trafficking system (figure 1*o*; [77]).

A problem with all models that envisage a role for planctomycetes in eukaryote origin is that there is no molecular phylogenetic evidence that would link any lineage of planctomycetes with eukaryotes [78]. The problems with theories that derive the nucleus from an endosymbiont are numerous and have been listed in detail elsewhere [40]; in essence, they fail to explain why the nuclear compartment is so fundamentally different from any free-living cell from the standpoints of (i) biosynthetic or ATP-generating physiology (altogether lacking in the nuclear compartment), (ii) membrane topology (no free-living cell is bounded similarly), (iii) permeability (no prokaryotic cytosol is contiguous with the environment via pores), and (iv) division (dissolution of a superficial homologue to the plasma membrane once per cell division in eukaryotes with open mitosis). Endosymbiotic theories for plastid and mitochondrial origin do not have those problems. A problem with the thermoreduction hypothesis is that it does not address the issue of where eukaryotes come from in the first place, it just takes their origin as a given. The recognition that the common ancestor of eukaryotes possessed a mitochondrion [30,32,79] is a severe problem for thermoreduction hypotheses, because the eukaryote has to first give rise to a prokaryote (the mitochondrial ancestor) that is required for its own origin, a sequence of events that, at face value, requires time to run backwards. Thermoreduction hypotheses are generally silent regarding the origin of mitochondria. Very few models for the origin of the nucleus, possibly only one, derive the nucleus in an archaeal host that possessed a mitochondrion. That model posits the nuclear membrane to arise from vesicles of membranes consisting of bacterial lipids [40] and invokes the need to separate splicing from translation as the selective pressure that led to the fixation of the compartmentation into nucleoplasm and cytoplasm [35].

The recent focus both on the evolution of cytoskeletal components [76] and on an autogenous (non-symbiotic) origin of a phagocytosing amitochondriate eukaryote point to a problem that should be mentioned. That theory, once called the archezoa hypothesis [55,56], now sometimes called the phagocytosing archaeon theory [31], envisages that point gradual changes lead to a prokaryotic host that can perform fully fledged eukaryotic phagocytosis (a quite complex process). These theories have it that phagocytosis is the key character that enabled the endosymbiotic origin of mitochondria. A problem common to those theories is that the phagocytotic, primitively amitochondriate eukaryote does not need a mitochondrion at all, and if there were some construable selective advantage then eukaryotes should have arisen from prokaryotes in multiple lineages independently. That has always been one of the weakest aspects of autogenous theories, in addition to the bioenergetic aspects [34].

4. The origin of mitochondria (and chloroplasts)

Endosymbiotic theory for the origin of chloroplasts and mitochondria started again with Mereschkowsky [13] and his idea about a symbiosis between 'chromatophores' (plastids) and a heterotrophic amoeboid cell. He contradicted the orthodox view that chromatophores are autogenous organs of the plant cells; he saw them as symbionts, extrinsic bodies or organisms, which entered into the host's plasma establishing a symbiotic relationship. The host for the origin of plastids itself originated, in his view, from an earlier symbiosis between a heterotrophic, amoeboid cell and a 'micrococcal' endosymbiont that gave rise to the nucleus (figure 2*a*; [13]). Comparison of physiological and anatomic attributes of plastids and cyanobacteria known at that time led him to the certain conclusion that the endosymbionts were 'cyanophyceae' (cyanobacteria) that entered into symbioses with amoeboid or flagellated cells on several independent occasions, leading to a plant kingdom having several independent origins. That is, he viewed the different coloured plastids of algae (red, green, brown, golden) as inheritances from different endosymbionts, each having those different pigmentations. Although he was wrong on that specific interpretation—today there is broad agreement that the plastids of all plants and algae have a single origin [80–82]—he was right with the endosymbiotic, cyanobacterial origin of plastids.

Mereschkowsky failed, however, to recognize the endosymbiotic origin of mitochondria, although the physiological properties of cells that he explained with the endosymbiotic origin of the nucleus are, from today's perspective, properties of mitochondria [15]. As very readably explained by Archibald [6], Portier developed (in French) the idea that there was a close relationship between bacteria and mitochondria and that mitochondria were involved in numerous processes in the cell. But like Schimper in his footnote regarding plastids, which we translated above, Portier proposed that mitochondria could be cultured outside their host cells, and this precipitated unforgiving criticism from his contemporaries [6]. Clearly, both Reinke (as cited in Schimper's footnote that we translated above) and Portier were observing the proliferation of contaminating bacteria, not of free-living organelles. Wallin [83] developed the endosymbiotic theory further for mitochondria, in English. He recognized that these organelles are descendants of endosymbiotic bacteria, but it remained

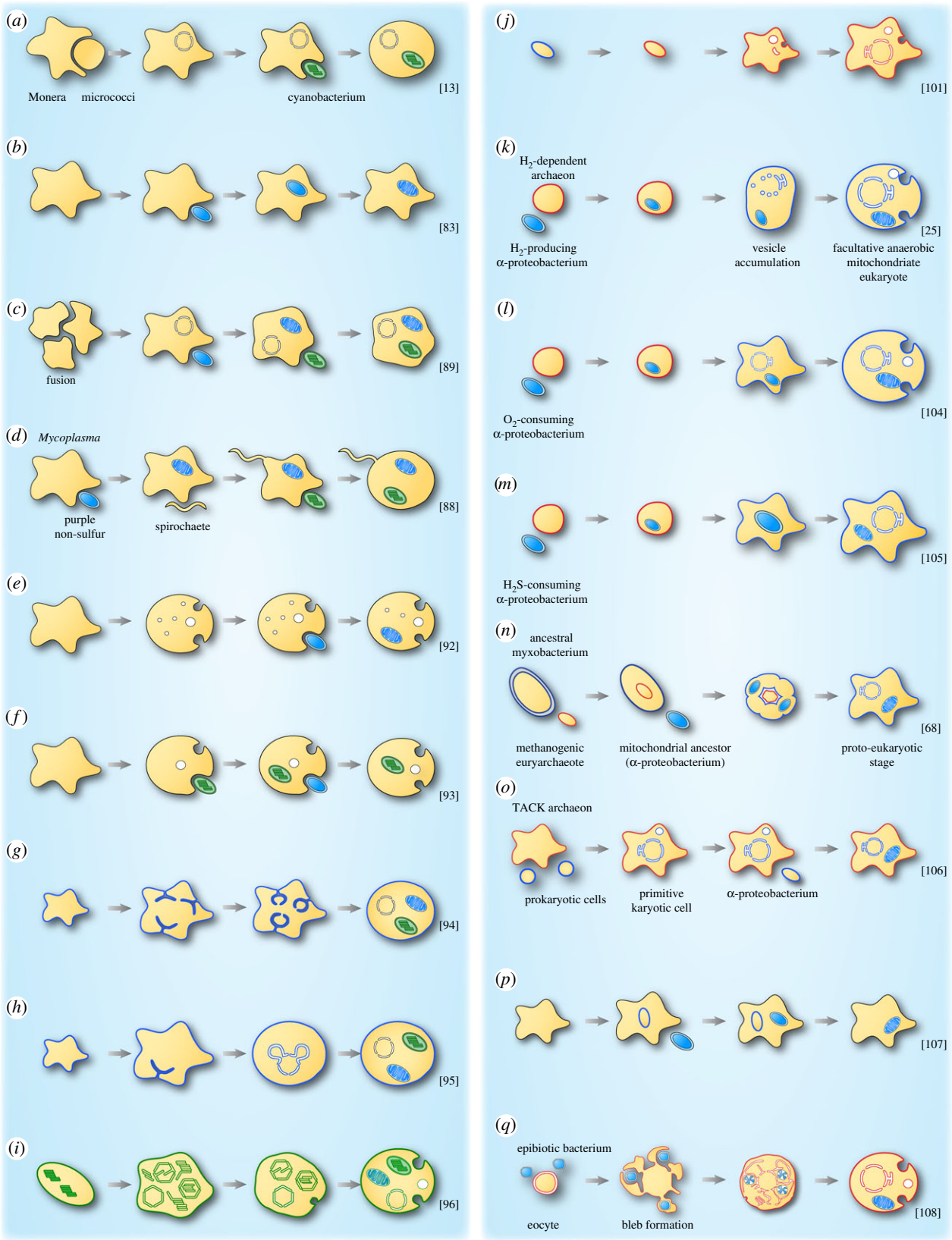


Figure 2. Models describing the origin of mitochondria and/or chloroplasts in eukaryotes. (*a*–*q*) Schematic of various models accounting for the origin of mitochondria and/or chloroplasts. Archaeal cells/membranes are represented with red, while blue indicates eubacterial cells/membranes. Black membranes are used when the identity of the cell is not clear and green is used for cyanobacterial derived cells/membranes. See also [22].

very unclear what his idea about the host was (figure 2*b*; [83]). Like Portier, he thought the cultivation of mitochondria outside their host to be possible. But he had the concept of gene transfer from organelles to the nucleus in mind: 'It appears logical, however, that under certain circumstances, [. . .] bacterial organisms may develop an absolute symbiosis with a higher organism and in some way or another impress a new character on the factors of heredity. The simplest and most readily conceivable mechanism by which the alteration takes place would be the addition of new genes to the chromosomes from the bacterial symbiont' [84, p. 144].

In print, cell biologists rejected endosymbiotic theory during the 1920s and through into the 1970s. A few prominent trouncings were (i) from Wilson [85] who wrote (pp. 738–739)

'Mereschkowsky ('10), in an entertaining fantasy, has developed the hypothesis' . . . 'in further flights of the imagination Mereschkowsky suggests', (ii) from Buchner [86] (pp. 79–80), who discussed endosymbiotic theory in a chapter entitled '*Irrwege der Symbioseforschung*' (translation: Symbiosis research gone astray) and (iii) from Lederberg [87], who surmised (p. 424): 'We should not be too explicit in mistaking possibilities for certainties. Perhaps the disrepute attached to some of the ideas represented in this review follows from uncritical over-statements of them, such as the Famintzin–Merechowsky theory of the phylogeny of chloroplasts from cyanophytes (28, 126) or the identity of mitochondria with free-living bacteria (198)'.

Endosymbiotic theory was repopularized in 1967 by Lynn Sagan (later Margulis) [88] and also mentioned in a very curious paper by Goksøyr [89]. As far as we can tell, those were the initial suggestions in endosymbiotic theory that both chloroplasts and mitochondria are descended from endosymbionts, but from separate endosymbionts. Goksøyr suggested an evolutionary development of mitochondria and later, in an independent symbiosis, chloroplasts from prokaryotic forms through a coenocytic relationship in which anaerobic prokaryotes (most likely of a single species) were brought into contact without intervening cell walls (figure 2*c*; [89]). The DNA of these cells accumulated in the centre of the agglomerate, a nuclear membrane arose from an endoplasmic reticulum, establishing an anaerobic eukaryotic cell. Aerobic eukaryotes trace back to an endocellular symbiotic relationship of anaerobic eukaryotes with aerobic prokaryotes, which emerged with the enrichment of oxygen in the atmosphere. The later loss of autonomy by the aerobic prokaryote to become a mitochondrion came along with gene transfer to the host's nucleus. An uptake of a primitive cyanobacterium, involving gene transfers to the nucleus again, led to photosynthetic eukaryotes. Goksøyr assumed that coenocytic systems occurred several times from different prokaryotic forms, making the origin of eukaryotes a non-monophyletic one [89]. Goksøyr's paper contains only one reference, to a 1964 paper by Stanier, and no mention of the older symbiotic literature.

Lynn Sagan rekindled the idea of a prokaryotic ancestry of mitochondria and chloroplasts and extended the idea to include a spirochaete origin of flagella [88]. On the second page of her 1967 paper, which was reported to have been rejected by 15 different journals [90], she states 'Although these ideas are not new. . .' while referring to Mereschkowsky's 1910 paper [15], although Mereschkowsky does not appear in the bibliography of her 1970 book [91]. She suggested the origin of eukaryotes from prokaryotes to be related to the increasing production of free oxygen by photosynthetic prokaryotes and the increasing proportion of oxygen in the atmosphere. Her host was a heterotrophic anaerobic prokaryote (perhaps similar to *Mycoplasma*), in whose cytoplasm an aerobic prokaryotic microbe (the proto-mitochondrion) was ingested, resulting in the evolution of an aerobic amoeboid organism, which later acquired a spirochaete, resulting in the eukaryotic flagellum (figure 2*d*; [88]; her later versions modified that order of events). She depicted the evolution of plastids as several ingestions of different photosynthetic prokaryotes (protoplastids—evolved from oxygen-consuming prokaryotes, homologous to cyanobacteria) by heterotrophic protozoans (figure 2*d*; [88]).

Countering Margulis, de Duve [92] outlined that the primitive phagocyte, which symbiotically adopted different types of microorganisms, was a primitive aerobe that remained dependent on hydrogen peroxide-mediated respiration during its early evolution, establishing through the loss of the cell wall and the evolution of membrane invagination processes (endocytosis) a primitive phagocyte with peroxisomes as the main (aerobic) respiratory organelle. This amitochondriate, peroxisome-bearing organism became later the host of an aerobic bacterium with oxidative phosphorylation, the ancestor of mitochondria (figure 2*e*; [92]). Stanier suggested an anaerobic, heterotrophic host in the evolution of chloroplasts [93] and placed the origin of chloroplasts before the origin of mitochondria, arguing that since mitochondria use oxygen, and since eukaryote origin took place in anaerobic times, there must have been first a sufficient and continuous source of oxygen before mitochondria were able to develop (figure 2*f*; [93]).

In the early 1970s, there was considerable resistance to the concept of symbiosis in cell evolution. Raff & Mahler [94] presented an alternative, non-symbiotic model for the origin of mitochondria, proposing that the proto-eukaryote was an advanced, heterotrophic, aerobic cell of large size, which enlarged the respiratory membrane surface achieved by invaginations of the inner cell membrane, which then formed membrane-bound vesicles blebbing off the respiratory membrane, generating closed respiratory organelles acquiring an outer membrane later on (compartmentalization, figure 2*g*; [94]). Bogorad [95] described a cluster clone hypothesis for the origin of eukaryotic cells from an uncompartmentalized single cell. He suggested that the cell's genome split into gene clusters (representing a new genome), followed by a membrane development around each gene cluster to create one or more gene-containing structures from which nuclei, mitochondria and chloroplasts evolved (figure 2*h*; [95]). Cavalier-Smith [96] explained the origin of chloroplasts and mitochondria by fusion and restructuring of thylakoids in a cyanobacterium. Plastids resulted through restructuring of photosynthetic thylakoids and mitochondria through restructuring of respiratory thylakoids, respectively (figure 2*i*; [96]). Though molecular evolutionary studies put non-symbiotic models for the origin of plastids and mitochondria more or less out of business [97], skepticism regarding endosymbiotic theory tends to run deep. Anderson *et al.* [98] in their publication on human mitochondrial DNA concluded that the data 'make it difficult to draw conclusions about mitochondrial evolution. Some form of endosymbiosis, involving the colonization of a primitive eukaryotic cell by a respiring bacteria-like organism, is an attractive hypothesis to explain the origin of mitochondria. However, the endosymbiont may have been no more closely related to current prokaryotes than to eukaryotes' [98, p. 464].

During the 1970s and 1980s, some other models for the origin of eukaryotes were developed, which are not presented in figure 2. John & Whatley [99] presented a very explicit symbiotic model for the origin of mitochondria with an anaerobic, fermenting, mitochondrion-lacking 'proto-eukaryote' as the host for a free-living aerobic respiring bacterium (similar to *Paracoccus denitrificans*), giving rise to the mitochondria where again the host's origin is not addressed. Woese [100] recognized that the archaebacteria might be related to the host lineage in endosymbiotic theory, but his model for the origin of mitochondria suggested a mitochondrial origin early in Earth's history, when the atmosphere was anaerobic, that mitochondria might descend from an initially photosynthetic organelle, that gained the ability of oxygenic respiration after becoming an endosymbiont [100].

In 1980, both van Valen & Maiorana (figure 2*j*; [101]) and Doolittle [102] put archaebacteria into the context of endosymbiosis, suggesting that they are the sister groups of the host that acquired the mitochondrion. Margulis [103] adjusted her version of endosymbiotic theory to accommodate the discoveries of archaea accordingly, but she kept the symbiotic (spirochaete) origin of flagella.

The hydrogen hypothesis posits anaerobic syntrophy as the ecological context linking the symbiotic association of an anaerobic, strictly hydrogen-dependent and autotrophic archaebacterium as the host with a facultatively anaerobic, heterotrophic eubacterium as endosymbiont (figure 2*k*; [25]). It entails an ancestral mitochondrion that could use either its electron transport chain or use mixed acid (H_2-producing) fermentations, thus it directly accounts for the common ancestry of mitochondria and hydrogenosomes as well as for intermediate forms between the two, the anaerobic mitochondria [21]. The model of Vellai and Vida [104] operates with a prokaryotic host for the origin of mitochondria (figure 2*l*), as does the sulfur cycling theory of Searcy (figure 2*m*; [105]), but neither accounts for hydrogenosomes or anaerobic mitochondria.

López-García & Moreira [68] proposed an evolutionary scenario for the origin of mitochondria that also includes an endosymbiotic origin of the nucleus. Their model is also a syntrophic symbiosis mediated by interspecies hydrogen transfer between a strict anaerobic, methanogenic archaeon, that became the nucleus, and a fermenting, heterotrophic, hydrogen-producing ancestral myxobacterium (δ-proteobacterium) [68] that served as its host; the mitochondrial ancestor (an α-proteobacterium) was then surrounded by the syntrophic couple, which led to an obligatory (endo)symbiotic stage with metabolic compartmentation as selective force to avoid interference of opposite anabolic and catabolic pathways. After the mitochondrion was stabilized, a loss of methanogenesis occurred generating the proto-eukaryote stage, in which the archaeal endosymbiont became the nucleus (figure 2*n*; [68]).

The phagocytosing archaeon theory was proposed by Martijn & Ettema [106], which posits an archaeon (most probably belonging to the TACK superphylum) and an α-proteobacterium (the proto-mitochondrion). The archaeon first phagocytotically took up various forms of other prokaryotic cells and digested them, resulting in gene transfers, whereby we note that phagocytosis is not required for gene transfer among prokaryotes. To protect its genetic material from such 'contamination' a membrane was formed by invagination (the nuclear envelope), resulting in a primitive karyotic cell type. At that stage, an α-proteobacterium was engulfed, establishing an endosymbiotic interaction with the host, leading to a protomitochondrial cell type (figure 2*o*; [106]). This model that has quite a bit in common with that of Cavalier-Smith [57] in that the origin of eukaryotic cell complexity (phagocytosis and nucleus) preceeds the origin of mitochondria, which for energetic reasons is unlikely [34]. Gray [107] recently proposed the pre-mitochondrion hypothesis, which does not account for the origin of eukaryotes but assumes that the host was already more or less eukaryotic in organization, and furthermore assumes that the host was aerobic prior to the origin of mitochondria, emphasizing, like de Duve & Margulis [18,19], oxygen in endosymbiotic theory. The origin of mitochondria was preceded by an ATP-consuming 'compartment', the pre-mitochondrion, presumably surrounded by one membrane (he is not explicit on this point), that became converted into the mitochondrion via retargeting of its proteins into a *Rickettsia*-like α-proteobacterial endosymbiont (figure 2*p*; [107]). The pre-mitochondrion hypothesis is silent on the origin of the archaeal components of eukaryotes, on the presence or the absence of a nucleus in the host, and on anaerobic forms of mitochondria.

The perhaps latest model for the origin of the eukaryotic cell and mitochondria is the inside-out theory by David & Buzz Baum [108]. They argued that an increasing intimate mutualistic association between an archaeal host (eocyte) and an epibiotic α-proteobacterium (the proto-mitochondrion), which initially lived on the host cell surface, drove the origin of eukaryotes. The host cell started to form protrusions and bleb enlargements to achieve a greater area of contact between the symbiotic partners, resulting in the outer nuclear membrane, plasma membrane and cytoplasm, whereas the spaces between the blebs generated the ER. The symbionts were initially trapped in the ER, but penetrated the ER's membrane to localize to the cytosol during evolution (figure 2*q*; [108]).

This section has shown that much thought has been invested on the topic of how the mitochondrial endosymbiont could have entered its host. Many theories place a premium on phagocytosis and predation upon bacteria as the essential step for allowing the symbiont to enter its host. Predation is actually very widespread among bacteria [109], but it never involves phagocytosis, instead it involves *Bdellovibrio*-like penetration mechanisms, an ability that has evolved in many independent lineages of bacteria, including *Micavibrio*, and that has been suggested to have possibly played a role in mitochondrial origin [110,111]. But predation, whether involving phagocytosis or bacterial predation, leaves mitochondria looking like leftovers of indigestion. Endosymbiosis and organelle origins are not about digestion. Microbial symbiosis, the process that gave rise to bioenergetic organelles, is about chemistry.

5. Anaerobes and mitochondrial origin in a prokaryotic host

Endosymbiotic theory is traditionally founded in comparative physiology (core carbon and energy metabolism). That is true for Mereschkowsky [13,15], for Margulis's 1970 formulation [91], for John and Whatley's version [99], and for van Valen and Maiorana's version [101]. The only formulation of endosymbiotic theory that directly accounts for anaerobic mitochondria and the (largely phylogeny-independent) distribution of anaerobes across all major eukaryotic groups and their use of the same small set of enzymes underlying their anaerobic ATP synthetic pathways [21] is the hydrogen hypothesis, which is also founded in comparative physiology.

The theories in the foregoing have different strengths and weaknesses; they are also designed to explain different aspects of eukaryotic cells too numerous to outline here. It is not our aim to defend them all or criticize them all. Instead we wish to focus on one of them, the one that accounts for the anaerobes. Theories are supposed to make testable predictions; in that respect the hydrogen hypothesis [25] has done fairly well. It posits that the host for the origin of mitochondria (hereafter, the host) was an archaeon, not a eukaryote, a view that is now current [30,31]. It predicted that no eukaryotes are primitively amitochondriate. That view is now

conventional wisdom on the issue [28,30,32,33], though it was far from common wisdom when proposed. Other theories ultimately generated the same prediction with regard to mitochondrial ubiquity but were not explicit on organisms like *Entamoeba*, *Giardia* and microsporidia, which harbour neither respiring mitochondria nor fermenting hydrogenosomes and were later found to harbour relict organelles that came to be known as mitosomes [26,27,112–114]. The hydrogen hypothesis did not directly predict the existence of mitosomes, but it did explicitly predict that organisms like *Entamoeba* and *Giardia* are derived, via reduction, from organisms that possessed the same endosymbiont as gave rise to mitochondria and hydrogenosomes. It also clearly predicted the chimaeric nature of eukaryotic genomes [32], which well into the late 1990s were supposed to represent a pure archaeal lineage [115].

The nature of host–symbiont interactions at the onset of mitochondrial symbiosis in the hydrogen hypothesis was posited to be anaerobic syntrophy, the host being a H_2-dependent archaeon, the symbiont being a facultative anaerobe that was able to respire in the presence of O_2, or to perform H_2-producing fermentations under anaerobic conditions. This is sketched in figure 3*a* for the example of methanogenesis, the metabolic model upon which the hypothesis was based, but, clearly, there are many H_2-dependent archaea, and it was clearly stated that any strictly H_2-dependent host would fit the bill [25]. This is the strength of the hydrogen hypothesis, because its host actually needs its mitochondrial symbiont. This is not true for any other version of endosymbiotic theory. Variants have been proposed that invoke anaerobic syntrophy to derive the nucleus via endosymbiosis [67,68,118], but they posit no metabolic demand or requirement for the involvement of mitochondria at eukaryote origin. In all versions of the endosymbiont hypothesis that entail a heterotrophic host, the host does not need its (mitochondrial) endosymbiont.

Anaerobic syntrophy (H_2-transfer) is thus the metabolic context of host–symbiont association, leading to hosts that tend to interact tightly with and adhere to their symbionts (figure 3*b*), similar to the symbiotic associations between methanogens in hydrogenosomes in the cytosol of anaerobic ciliates [119]. This can, in principle, lead to a situation like that sketched in figure 3, with a prokaryotic (bacterial) symbiont residing within a prokaryotic (archaeal) host. This was a fairly radical proposal of the theory, because it did not invoke phagocytosis as the mechanism of endosymbiont entry, an aspect that drew fierce criticism from Cavalier-Smith [57]. In the meantime, examples of prokaryotes that have come to reside as stable endosymbionts within other prokaryotes have been well studied [120,121]. In those examples, the host prokaryotes are definitely not phagocytotic, so phagocytosis is clearly not a prerequisite for the establishment of intracellular symbiosis. Without question, phagocytosis greatly increases the frequency with which endosymbionts become established within eukaryotic cells [122], but—notably—none of those countless cases of phagocytosis-dependent bacterial symbiosis have ever led to anything resembling a second origin of mitochondria. Conversely, a bacterial–archaeal symbiotic association that clearly resembles a second origin of eukaryotes—from the standpoint of physiology, metabolism and the direction of gene transfer—has been described; it gave rise to the haloarchaea [123,124].

The H_2-dependent nature of the host leads to a curious situation in phase depicted in figure 3*c*. In order to generate H_2 for the host, the symbiont requires reduced organic compounds (fermentable organic substrates), but the host is a strict autotroph and cannot supply them in excess of its own needs because H_2-dependent autotrophs live from gases and do not import reduced organic compounds. This phase of the symbiosis is thus unstable because the symbiont will eventually consume the host's cytosol. In order for the symbiosis to persist, either the host needs to invent importers for organics, or the symbiont's preexisting genes for importers are transferred to the host's chromosomes and can be expressed there, and the bacterial importers need to be functional in the archaeal membrane, which is true in haloarchaea [123]. Gene transfer could merely involve occasional lysis of an endosymbiont, just as it occurs in endosymbiotic gene transfer (gene transfer from organelles to the nucleus) in eukaryotes today [117], except that at this stage of the symbiosis, the host is still an archaeon and lacks a nucleus, although the bipartite cell has a bacterial endosymbiont and gene transfer from symbiont to host has commenced (figure 3*d*).

Expression of carbon importers in the host's membrane does not completely solve the problem though, because the hydrogen hypothesis posits that the host was an autotroph, hence its carbon metabolism was specialized to anabolic pathways. A good example of such enzymatic specialization is the bifunctional fructose 1,6 bisphosphate aldolase/bisphosphatase that is characteristic of archaeal autotrophs [125] but is altogether missing in eukaryotes, but many other examples of archaeal-specific enzymes of sugar-phosphate (and unphosphorylated sugar) metabolism are known [126,127]. Thus, either the enzymes of the host's anabolic metabolism need to acquire, one point mutation at a time, the substitutions required to make carbon metabolism run backwards, or, more likely and more rapidly achieved, genes for the symbiont's heterotrophic carbon metabolism are also expressed in the host's chromosomes. As in the case of the importers, this also involves straight endosymbiotic gene transfer, without targeting of the protein product to the donor symbiont, just expression in the archaeal cytosol.

This transfer does a variety of important things. First, it allows carbon to be directed to the symbiont, so that it can produce H_2 via fermentation to satisfy the host. Second, it confers heterotrophy upon the host compartment (the cytosol), but only if transfer of the symbiont's entire glycolytic pathway is successful (the enzymatic steps all the way to pyruvate), because the first net gain of ATP in glycolysis is at the pyruvate kinase step. Third, if that occurs, it directly accounts for the bacterial origin of eukaryotic glycolytic enzymes (except enolase: [128]). No other formulation of endosymbiotic theory accounts for the observation that eukaryotes, though their ribosomes stem from archaea, have a bacterial glycolytic pathway; indeed, for other versions of endosymbiotic theory it is not even an explanandum.

Fourth, and quite unexpectedly, the selective pressure associating the two partners from the beginning and selecting the transfer of importers and glycolysis to the host compartment was the host's dependence upon H_2 to run its carbon and energy metabolism. But the expression of genes for heterotrophic carbon flux in the host compartment supply it with reduced carbon species and ATP and there is no longer any selective pressure to maintain the host's autotrophic lifestyle, which will necessarily have involved membrane bioenergetics because all autotrophs are dependent upon chemiosmotic coupling. As a result, the host can relinquish its autotrophy;

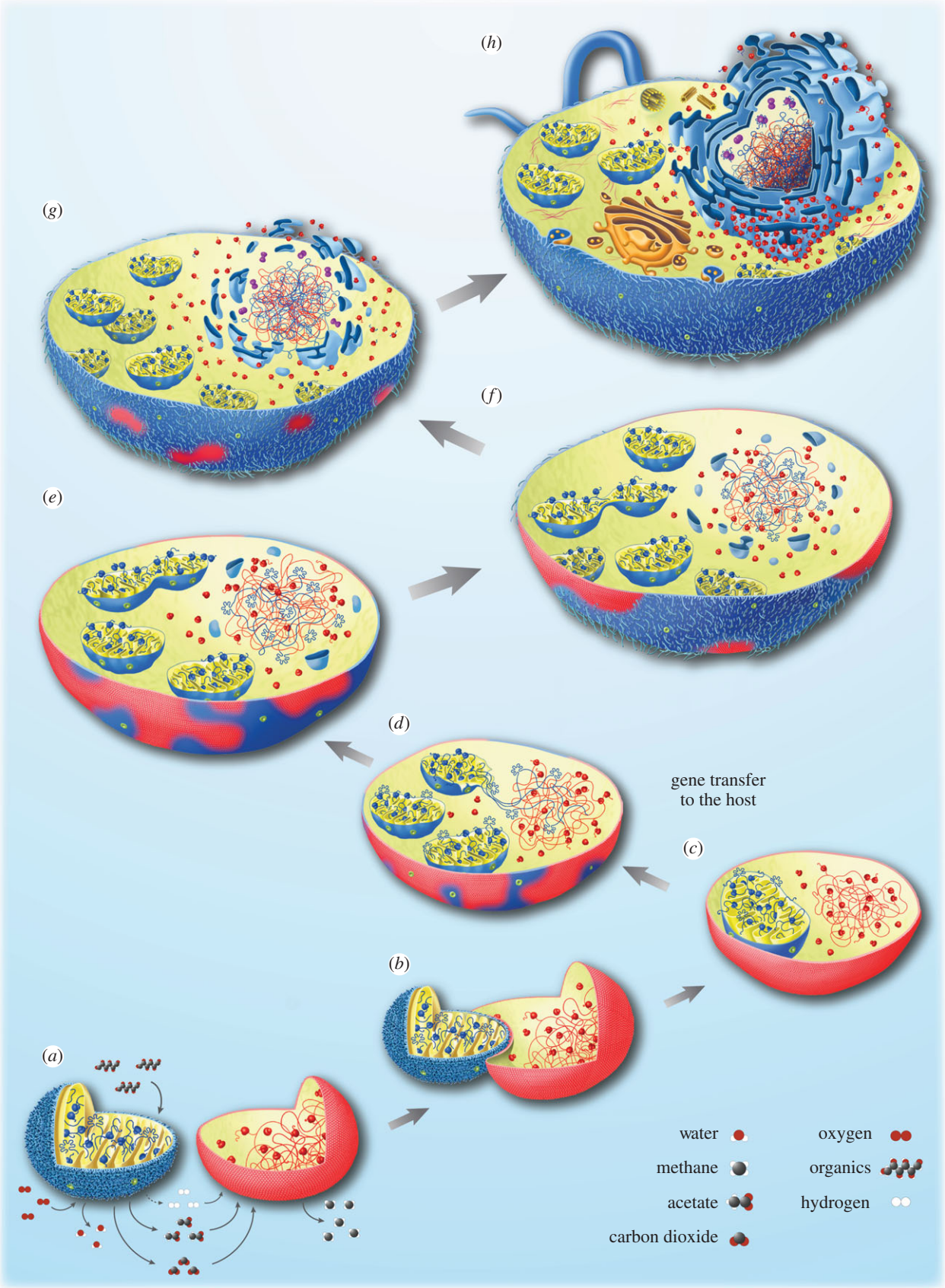


Figure 3. Mitochondrial origin in a prokaryotic host. (*a*–*h*) Illustrations for various stages depicting the transition of a H_2-dependent archaeal host (in red) and a facultatively anaerobic α-proteobacterium (in blue) to an eukaryote. See also [25,34,35] regarding this transition, and [116,117] regarding gene transfer from organelles to the nucleus.

it has become a heterotroph with chimaeric chromosomes harbouring archaeal and bacterial genes, and archaeal ribosomes and glycolysis in the cytosol. In addition, the cytosol harbours a facultatively anaerobic bacterial endosymbiont with a respiratory chain and H_2-producing fermentations (figure 3*d*) that can donate a full genome's worth of bacterial

genes over and over again, replacing many indigenous archaeal pathways with bacterial counterparts, and thus transforming the archaeon from within. Part of this transformation involves the establishment of bacterial lipid synthesis (indicated in blue in figure 3); although the archaeal pathway of lipid synthesis (the mevalonate pathway) has been retained in eukaryotes [129], it is not just used for isoprene ether lipid synthesis, rather it is used for isoprenes in general, such as cholesterol (which requires only trace, that is, non-molar amounts of oxygen [130]), or for the hydrophobic tails of quinone or for dolichol phosphate.

Gene transfer from symbiont to host carries some fateful hitchhikers—self-splicing group II introns. These are indicated in figure 3 as hand-shaped structures in the symbiont's genome. Group II introns are important because their transition into spliceosomal introns is thought to have precipitated the origin of the nucleus [35]. How so? Group II introns occur in prokaryotic genomes [131,132], they are mobile, they can spread to many copies per genomes [133] and they remove themselves via a self-splicing mechanism that involves the intron-encoded maturase [134]. Their splicing mechanism is similar to that in spliceosomal intron removal [135], for which reason they have long been viewed as the precursors of both (i) spliceosomal introns and (ii) their cognate snRNAs in the spliceosome: one 'master' intron in the genome could provide all necessary splicing functions *in trans*; resident group II introns could degenerate so as to become dependent on the *trans* functions and thus to end up as small elements having conserved residues only at the splice sites and the lariat site A.

The crux of the splicing hypothesis for nuclear origins [35] is this: introns entered the eukaryotic lineage via gene transfer from the mitochondrial endosymbiont to an archaeal host (figure 3*d*), where they subsequently spread to many sites in the host's chromosomes (figure 3*e*). Evidence for this is the observation that about half of introns in eukaryotic genes are ancient, being present at positions that are conserved across divergent eukaryotic lineages, indicating their presence in the eukaryote common ancestor [35]. Once they begin to undergo the transition to spliceosomal introns a curious situation arises: splicing is slow, of the order of minutes per intron [136], while translation is fast, of the order of 10 peptide bonds per second. As the transition to spliceosomal introns set in, the host's cytosol was still a prokaryotic compartment in that there was cotranscriptional translation, with active ribosomes synthesizing proteins on nascent transcripts (figure 3*f*). That is not a problem for group II introns, which use their maturase from one ribosome passage to block the mRNA 5′ end until the intron is removed. But with the origin of fully fledged spliceosomes (symbolized as purple dumbbells in figure 3*g*) transitioning to spliceosomal splicing, nascent transcripts are translated before they can be spliced. This means that introns are translated, leading to defective gene expression at hundreds of loci simultaneously, a surely lethal condition for the host unless immediately remedied. There are a finite number of solutions to this problem, in addition to precipitating the origin of nonsense-mediated decay (nmd), a eukaryote-specific machinery that recognizes and inactivates intron-containing mRNAs [137].

One solution would be to simply remove all the introns in the chromosomes. That did not happen, because many intron positions are ancient [138,139]. Another solution would be to invent a spliceosome that is much faster than ribosomes, but that is almost like asking for a miracle, because the modern spliceosome has had more than a billion years to refine its function, but it has not become faster. Another solution would be to physically, hence spatiotemporally, separate the slow process of splicing from the fast process of translation so that the former could go to completion before the latter set in. Separation in cells usually involves membranes, and that is the central tenet of the splicing hypothesis: the initial pressure that led to selection for the nuclear membrane was to exclude active ribosomes from active chromatin (figure 3*h*), allowing the slow process of splicing to go to completion around the chromosomes, and thereby initially allowing distal diffusion, later specific export of processed mRNAs to the cytosol for translation [35]. The nuclear pore complex mediates the translocation of proteins and mRNA between the cytosol and the nucleus. Comparative genomics of nuclear pore complex proteins and proteins that make up the nucleolus shows that many of them share domains with both archaeal and bacterial proteins [140,141].

In that view, the origin of the nucleus marks the origin of a genuinely new cell compartment—not the nucleus itself, but the eukaryotic cytosol—that is free of active chromatin, where protein–protein interactions, rather than protein–DNA interactions, move to the fore in signalling and regulation, and where proteins can spontaneously aggregate and interact in such a way as to generate new structures and functions, including the true cytoskeleton and membrane traffic processes that distinguish eukaryotes from prokaryotes. A curious property of this model for the origin of the nucleus is that it only requires eukaryotes to possess a nuclear membrane when they are expressing genes, which directly points to another very curious (and vastly underappreciated) character that separates eukaryotes from prokaryotes: prokaryotes express their genes continuously during cell division, while eukaryotes shut down the expression all of their genes before chromosome partitioning and cell division. To us, this suggests an evolutionary link between splicing the splicing-dependent origin of the nucleus, the origin of genome-wide gene silencing mechanisms [142], which generally involve chemical modifications of chromatin and histones, and the origin of the eukaryotic cell cycle.

This set of events leads to a bipartite cell (figure 3*h*) (i) that requires a nucleus in order to express genes, (ii) that has retained archaeal ribosomes in the cytosol as a vestige of the host, (iii) that has bacterial energy metabolism both in the cytosol and in the mitochondrion, (iv) that has lost all electron-transfer phosphorylation functions in the plasma membrane, (v) that has nonetheless retained the archaeal ATPase, which however now operates backwards to acidify the vacuole, and (vi) that has typical eukaryotic features. It is true that many theories for eukaryote origin surveyed here address many of the same aspects, but what everyone has overlooked for the now nearly 50 years since Margulis revived endosymbiotic theory [88] is that the myriad inventions that distinguish eukaryotes from prokaryotes do not come for free. The origin of eukaryotic novelties had an energetic price, and that price was paid by mitochondria [34]. The internalization of bioenergetic membranes in eukaryotes frees them from the bioenergetic constraints that keep prokaryotes prokaryotic in organization. Since the late 1990s, there has been a growing realization that all eukaryotes have or had mitochondria, but it had not been clear why that is the case, until the calculations were done [34]. That puts the mitochondrial symbiosis at the very beginning of eukaryogenesis.

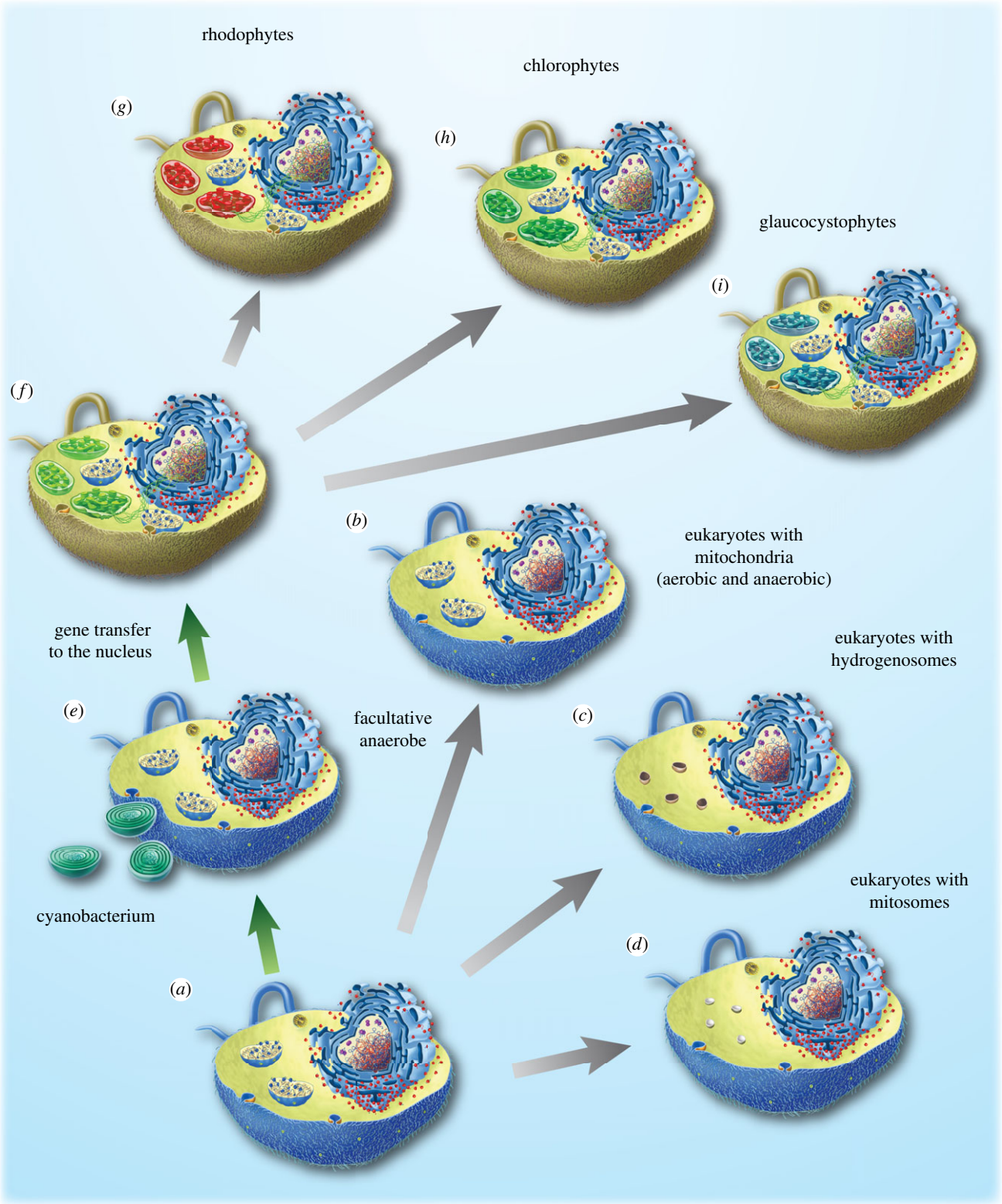


Figure 4. Evolution of anaerobes and the plastid. (*a–d*) Diversification of the mitochondria-containing ancestor to eukaryotes containing specialized forms of the organelle, hydrogenosomes, mitosomes and anaerobic mitochondria. See also [21,143]. (*e,f*) Primary symbiotic origin of a plastid involving a cyanobacterium in a facultative anaerobic host (see text), followed by gene transfer to the nucleus resulting in a plastid-bearing ancestor. See also [144]. (*g–i*) Diversification of the plastid-bearing ancestor to glaucocystophytes, chlorophytes and rhodophytes. See also [25].

6. Rounding out the picture: the plastid

Of course, there was one additional and crucial prokaryotic endosymbiont in eukaryotic history: a cyanobacterium that became the plastid. This is outlined in figure 4. The ancestral eukaryote was, seen from the standpoint of energy metabolism [21], a facultative anaerobe. It underwent specialization to aerobic and anaerobic environments in multiple independent lineages, giving rise to eukaryotes specialized to either aerobic or anaerobic environments [143], as well as giving rise to facultative anaerobes, like *Euglena* [21,145,146] or *Chlamydomonas* [147–149]. The prevalence of enzymes for anaerobic energy metabolism in eukaryotes in general [143], and in particular among algae like *Chlamydomonas* [149], together with the

circumstance that they use the same enzymes that *Trichomonas* and *Giardia* use to survive under anaerobic conditions, not to mention their conservation in *Cyanophora* [150], lead to a novel inference of some interest: the host for the origin of plastids was a facultative anaerobe.

The origin of plastids has been the subject of several recent papers [41,81,82,151]. In terms of endosymbiotic theory, the situation is clear: a eukaryote that already possessed a mitochondrion—a facultative anaerobe, as we just pointed out—obtained a cyanobacterium as an endosymbiont (figure 4*e*); possible metabolic contexts [152] for that symbiosis could have involved carbohydrate produced by the plastid, oxygen produced by the plastid [25], nitrogen supplied by the plastid [153] or a combination thereof. Although the phylogenetic affinity of the cyanobacterium that became the plastid is complicated by the circumstance that prokaryotes avidly undergo LGT, current analyses point to large-genomed, nitrogen-fixing forms [151,154]. Similar to the case for mitochondria, many genes were transferred from the endosymbiont to the host's chromosomes [144], which in the case of plastids were surrounded by a nucleus (figure 4*f*). The origin of protein import machineries of organelles played an important role, both in the case of mitochondria [155] and in the case of plastids [156], because it allowed the genetic integration of host and endosymbiont while allowing the endosymbiont to maintain its biochemical identity. The three lineages of algae harbouring primary plastids—the chlorophytes, the rhodophytes and the glaucocystophytes—diverged early in plastid evolution (figure 4*g–i*). At least two secondary endosymbioses involving green algae occurred [157–159], and at least one, but possibly more, secondary symbioses involving red algal endosymbionts occurred during evolution, whereby protein import probably also played an important role in the establishment of red secondary endosymbioses [82].

Since the inception of endosymbiotic theory by Mereschkowsky [13,15], the founding event that gave rise to primary plastids has been seen as the incorporation of the cyanobacterial endosymbiont. Over the past few years, a variant of endosymbiotic theory has, however, emerged that sees the plastid symbiosis as beginning with a chlamydial infection of a eukaryotic cell, an infection that was cured by the cyanobacterium. The chlamydial story for plastid origin developed slowly but has made its way into prominent journals lately [160]. There are several very severe problems with the chlamydia story, as several authors have recently pointed out [41,82,152,161,162]. Perhaps the most serious problem is that the gene trees upon which the current versions of the chlamydial theory are based do not say what the proponents of the chlamydial theory claim. This is shown in new analyses both by Deschamps [162], who provides an excellent historical overview of the chlamydial theory, and by Domman *et al*. [152]. Both papers show that the suspected chlamydia connection to plastid origin is founded in phylogenetic artefacts—trees that do not withstand critical methodological inspection. Because of phylogenetic factors and because of LGT among prokaryotes, trees can be misleading in the context of inferring endosymbiont origins [41], and it is prudent to look at other kinds of evidence as well. As it concerns the origin of mitochondria, Degli-Esposti [163] surveyed the components of proteobacterial membrane bioenergetics and inferred that the ancestor of mitochondria was methylotrophic.

7. Organelles have retained genomes (why?)

An important component of endosymbiotic theory is the circumstance that organelles have retained genomes. The observation that organelles had DNA at all was one of the key observations that supported endosymbiotic theory in the first place [102]. Indeed, several autogenous (non-endosymbiotic) alternatives to the endosymbiont hypothesis were designed specifically to explain the existence of DNA in organelles [94–96]. With very few important exceptions (that prove the rule, explained below), organelles have retained DNA.

Why have organelles retained DNA? The answer to that question is satisfactorily explained by only one theory: John F. Allen's CoRR hypothesis (co-location for redox regulation) [164,165]. It posits that organelles have retained genomes so that individual organelles can have a say in the expression of components of the respiratory and photosynthetic electron transport chains in order to maintain redox balance in the bioenergetic membrane. The CoRR hypothesis directly accounts for the observation that plastids and mitochondria have converged in gene content to encode almost exclusively genes involved in their respective electron transport chains, and components of the ribosome necessary to express them in the organelle. It has also recently come to the attention of some of us interested in endosymbiosis that plastids and mitochondria (and to some extent nucleomorphs) have furthermore converged in gene content to encode the same set of ribosomal proteins [38]. A compelling explanation for the otherwise puzzling and long overlooked convergence for ribosomal protein content in plastid and mitochondrial genomes is ribosome assembly; the process of ribosome biogenesis requires that some proteins need to be coexpressed in the same compartment as their nascent rRNAs [38]. The convergence observed in gene content in plastid and mitochondrial genomes is striking.

One of the burgeoning strengths of Allen's CoRR hypothesis for the evolutionary persistence of organelle genomes concerns its predictions with regard to hydrogenosomes. Hydrogenosomes have more or less everything that mitochondria have, but they have lost the respiratory chain in their inner membrane. CoRR posits the selective pressure to maintain organelle DNA to be the necessity to maintain redox balance. Some readers might ask: What is redox balance? Redox balance refers to the smooth flow of electrons through the electron transport chain. The concept of redox balance applies both to mitochondria and to chloroplasts, because both have electron transport chains that generate proton gradients to drive their respective ATPase. In both electron transport chains, quinols and quinones are an essential component. These membrane soluble electron carriers can transfer electrons non-enzymatically to O_2, generating the superoxide radical (O_2^-), which is the starting point for reactive oxygen species (ROS) [166]. If the flow of electrons through the bioenergetic membrane (the inner mitochondrial membrane or the thylakoid) is impaired, for example, because downstream components in the chain are present in insufficient amounts, or because upstream components in the chain are too active, then the steady-state quinol concentration increases (quinols are the reduced form of the quinones) and the quinols generate ROS. If an organelle relinquishes its electron transport chain, then there is, according to CoRR, no need to retain the genome, it can become lost, and precisely this

has happened in hydrogenosomes, in no less than four independent lineages: trichomonads, ciliates, fungi and amoeboflagellates [21]. Other theories for organelle genome persistence, for example the theory that organelles encode hydrophobic proteins [167], do not make that prediction.

8. Eukaryotes tug and twist the archaeal tree

There is currently much buzz about the possibility that a group of crenarchaeotes, the TACK superphylum (for Thaumarchaeota, Aigarchaeota, Crenarchaeota and Korarchaeota) might harbour the closest ancestors of the host that acquired the mitochondrion. Several different trees that address the issue have appeared recently ([30,31,50,51]; discussed in [168]). One aspect of those trees that has so far gone unmentioned is that trees that place the eukaryotic informational genes within the crenarchaeotes also root the archaea either with euryarchaeotes basal [50], within the euryarchaeotes [169] or within the methanogens [31,50–52]. Also, archaeal trees that do not include eukaryotes also tend to root the archaea within methanogens or within euryarchaeotes [30,51,52,170]. There are a number of traits that make methanogens excellent candidates for the most ancient among the archaeal lineages [171], methanogenesis is currently the oldest biological process for which there is evidence in the geological isotope record, going back some 3.5 Ga [172], and microbiologists considered methanogenesis to be one of the most primitive forms of prokaryotic metabolism even before archaea were discovered [173]. A methanogenic ancestry of archaea makes sense in many ways.

In line with that, abiotic (geochemical) methane production occurs spontaneously at serpentinizing hydrothermal vents [174–176] (for a discussion of serpentinization, see [177]). Of all naturally occurring geochemical reactions currently known, only the process of serpentinization at hydrothermal vents involves exergonic redox reactions that emulate the core bioenergetic reactions of some modern microbial cells [177–181]. The point is this: if the ancestral state of archaeal carbon and energy metabolism is methanogenesis, then all archaea are ancestrally methanogenic and ancestrally hydrogen dependent. This is relevant for models of eukaryote origins that involve anaerobic synthrophy (a hydrogen-dependent archaea host for the origin of mitochondria), because then hydrogen dependence becomes a very widespread trait affecting the evolution of all archaeal lineages, including those that gave rise to the eukaryotic host lineage.

Indeed, recent findings have it that many archaeal lineages stem from methanogenic ancestors via gene transfers [124]. In particular, the origin of haloarchaea is noteworthy because it entailed exactly the same physiological transformation (from strictly anaerobic H_2-dependent chemolithoautotroph to facultatively anaerobic heterotroph) as the hydrogen hypothesis posits for the origin of eukaryotes [123], and the mechanism underlying that transformation—gene transfer from bacterium to archaeon—is the same as in the hydrogen hypothesis. The main difference between the origin of the respiratory chain of haloarchaea and of mitochondria is that the former operates in an archaeal cytoplasmic membrane whereas the latter operates in the internalized bioenergetic membranes of mitochondria within eukaryotic cells [123]. It is precisely that difference, however, that separates the eukaryotes from the prokaryotes in terms of the metabolic energy available to drive the evolution of novel protein families and thus novel cell biological traits [34].

Thus, as the position of eukaryotes starts to come into focus within the archaeal tree, so does the position of the root among archaea, and multiple evolutionary transitions from an ancestrally H_2-dependent state seems to be a recurring theme within the archaea, with gene transfers from bacteria providing the physiological capabilities to access electron and energy sources other than H_2. Early archaeal evolution and the origin of eukaryotes are ancient events, so ancient that they push phylogenetic methods to their limits, and possibly beyond. The book of early evolution holds many exciting chapters, and the origin of eukaryotes is clearly one of the most crucial, because eukaryotes—and only eukaryotes, the cells that have mitochondria—brought forth genuinely complex life.

Authors' contributions. W.M., S.G. and V.Z. wrote the paper and prepared the figures.

Competing interests. We declare we have no competing interests.

Funding. This work was made possible by a grant to W.F.M. from the ERC.

Acknowledgements. We thank Katharina Brandstädter for help in preparation of the manuscript. We thank Svetlana Kilian for preparing graphic elements used in figures 3 and 4.

References

1. Knoll AH. 2014 Paleobiological perspectives on early eukaryotic evolution. *Cold Spring Harb. Perspect. Biol.* **6**, a016121. (doi:10.1101/cshperspect.a016121)
2. Lane N. 2014 Bioenergetic constraints on the evolution of complex life. *Cold Spring Harb. Perspect. Biol.* **6**, a015982. (doi:10.1101/cshperspect.a015982)
3. Wideman JG, Leung KF, Field MC, Dacks JB. 2014 The cell biology of the endocytic system from an evolutionary perspective. *Cold Spring Harb. Perspect. Biol.* **6**, a016998. (doi:10.1101/cshperspect.a016998)
4. Koonin EV. 2012 *The logic of chance: the nature and origin of biological evolution*. Upper Saddle River, NJ: FT Press.
5. Adl SM *et al.* 2005 The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *J. Eukaryot. Microbiol.* **52**, 399–451. (doi:10.1111/j.1550-7408.2005.00053.x)
6. Archibald JM. 2014 *One plus one equals one: symbiosis and the evolution of complex life*. Oxford, UK: Oxford University Press.
7. Altmann R. 1890 *Die Elementarorganismen und ihre Beziehungen zu den Zellen*. Leipzig, Germany: Verlag von Veit & Comp.
8. Höxtermann E, Mollenhauer D. 2007 Symbiose und Symbiogenese—Entdeckung und Entwicklung eines biologischen Problems. In *Evolution durch Kooperation und Integration* (eds A Geus, E Höxtermann), p. 258. Marburg an der Lahn: Basilisken-Presse.
9. Schwendener S. 1867 Über die wahre Natur der Flechten. *Verhandlungen der Schweizerischen Naturforschenden Gesellschaft in Rheinfelden* **51**, 88–90.
10. De Bary A. 1878 Über Symbiose. *Tageblatt der 51. Versammlung deutscher Naturforscher und Aerzte in Cassel*, pp. 121–126.
11. Schimper AFW. 1883 Über die Entwickelung der Chlorophyllkörner und Farbkörper. *Bot. Z.* **41**, 105–120, 121–136, 137–152, 153–162.
12. Schimper AFW. 1885 Untersuchungen über die Chlorophyllkörner und die ihnen homologen Gebilde. *Jahrb. wiss. Bot.* **16**, 1–247.
13. Mereschkowsky C. 1905 Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. *Biol. Centralbl.* **25**, 593–604.

14. Martin W, Kowallik K. 1999 Annotated English translation of Mereschkowsky's 1905 paper 'Über Natur und Ursprung der Chromatophoren im Pflanzenreiche'. *Eur. J. Phycol.* **34**, 287–295. (doi:10.1080/09670269910001736342)
15. Mereschkowsky C. 1910 Theorie der zwei Plasmaarten als Grundlage der Symbiogenesis, *einer neuen Lehre von der Entstehung der Organismen. Biol. Centralbl*. **30**, 353–442.
16. Geus A, Höxtermann E. 2007 *Evolution durch Kooperation und Integration*. Marburg an der Lahn: Basilisken-Presse.
17. Schlacht A, Herman EK, Klute MJ, Field MC, Dacks JB. 2014 Missing pieces of an ancient puzzle: Evolution of the eukaryotic membrane-trafficking system. *Cold Spring Harb. Perspect. Biol*. **6**, a016048. (doi:10.1101/cshperspect.a016048)
18. De Duve C. 2007 The origin of eukaryotes: a reappraisal. *Nat. Rev. Genet*. **8**, 395–403. (doi:10.1038/nrg2071)
19. Margulis L, Chapman M, Guerrero R, Hall J. 2006 The last eukaryotic common ancestor (LECA): acquisition of cytoskeletal motility from aerotolerant spirochetes in the Proterozoic Eon. *Proc. Natl Acad. Sci. USA* **103**, 13 080–13 085. (doi:10.1073/pnas.0604985103)
20. Lindmark DG, Müller M. 1973 Hydrogenosome, a cytoplasmic organelle of the anaerobic flagellate *Tritrichomonas foetus*, and its role in pyruvate metabolism. *J. Biol. Chem*. **248**, 7724–7728.
21. Müller M *et al.* 2012 Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. *Microbiol. Mol. Biol. Rev*. **76**, 444–495. (doi:10.1128/MMBR.05024-11)
22. Martin W, Hoffmeister M, Rotte C, Henze K. 2001 An overview of endosymbiotic models for the origins of eukaryotes, their ATP-producing organelles (mitochondria and hydrogenosomes) and their heterotrophic lifestyle. *Biol. Chem*. **382**, 1521–1539. (doi:10.1515/BC.2001.187)
23. Martin W, Rotte C, Hoffmeister M, Theissen U, Gelius-Dietrich G, Ahr S, Henze K. 2003 Early cell evolution, eukaryotes, anoxia, sulfide, oxygen, fungi first (?), and a tree of genomes revisited. *IUBMB Life* **55**, 193–204. (doi:10.1080/1521654031000141231)
24. Martin W. 2007 Eukaryote and mitochondrial origins: Two sides of the same coin and too much ado about oxygen. In *Primary producers of the sea* (eds P Falkowski, AH Knoll), pp. 53–73. New York, NY: Academic Press.
25. Martin W, Müller M. 1998 The hydrogen hypothesis for the first eukaryote. *Nature* **392**, 37–41. (doi:10.1038/32096)
26. Tovar J, Fischer A, Clark CG. 1999 The mitosome, a novel organelle related to mitochondria in the amitochondrial parasite *Entamoeba histolytica*. *Mol. Microbiol*. **32**, 1013–1021. (doi:10.1046/j.1365-2958.1999.01414.x)
27. 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**, 172–176. (doi:10.1038/nature01945)
28. van der Giezen M. 2009 Hydrogenosomes and mitosomes: conservation and evolution of functions. *J. Eukaryot. Microbiol*. **56**, 221–231. (doi:10.1111/j.1550-7408.2009.00407.x)
29. Embley TM, van der Giezen M, Horner DS, Dyal PL, Foster P. 2003 Mitochondria and hydrogenosomes are two forms of the same fundamental organelle. *Phil. Trans. R. Soc. Lond. B* **358**, 191–202. (doi:10.1098/rstb.2002.1190)
30. Williams TA, Foster PG, Cox CJ, Embley TM. 2013 An archaeal origin of eukaryotes supports only two primary domains of life. *Nature* **504**, 231–236. (doi:10.1038/nature12779)
31. Guy L, Saw JH, Ettema TJG. 2014 The archaeal legacy of eukaryotes: a phylogenomic perspective. *Cold Spring Harb. Perspect. Biol*. **6**, a016022. (doi:10.1101/cshperspect.a016022)
32. McInerney JO, O'Connell M, Pisani D. 2014 The hybrid nature of the Eukaryota and a consilient view of life on Earth. *Nat. Rev. Microbiol*. **12**, 449–455. (doi:10.1038/nrmicro3271)
33. Simpson AGB, Roger AJ. 2004 The real 'kingdoms' of eukaryotes. *Curr. Biol*. **14**, R693–R696. (doi:10.1016/j.cub.2004.08.038)
34. Lane N, Martin W. 2010 The energetics of genome complexity. *Nature* **467**, 929–934. (doi:10.1038/nature09486)
35. Martin W, Koonin EV. 2006 Introns and the origin of nucleus–cytosol compartmentalization. *Nature* **440**, 41–45. (doi:10.1038/nature04531)
36. Thiergart T, Landan G, Schenk M, Dagan T, Martin WF. 2012 An evolutionary network of genes present in the eukaryote common ancestor polls genomes on eukaryotic and mitochondrial origin. *Genome Biol. Evol.* **4**, 466–485. (doi:10.1093/gbe/evs018)
37. Rivera MC, Jain R, Moore JE, Lake JA. 1998 Genomic evidence for two functionally distinct gene classes. *Proc. Natl Acad. Sci. USA* **95**, 6239–6244. (doi:10.1073/pnas.95.11.6239)
38. Maier U-G, Zauner S, Woehle C, Bolte K, Hempel F, Allen JF, Martin WF. 2013 Massively convergent evolution for ribosomal protein gene content in plastid and mitochondrial genomes. *Genome Biol. Evol*. **5**, 2318–2329. (doi:10.1093/gbe/evt181)
39. Gilson PR, Maier UG, McFadden GI. 1997 Size isn't everything: Lessons in genetic miniaturisation from nucleomorphs. *Curr. Opin. Genet. Dev*. **7**, 800–806. (doi:10.1016/S0959-437X(97)80043-3)
40. Martin W. 1999 A briefly argued case that mitochondria and plastids are descendants of endosymbionts, but that the nuclear compartment is not. *Proc. R. Soc. Lond. B* **266**, 1387–1395. (doi:10.1098/rspb.1999.0792)
41. Ku C, Nelson-Sathi S, Roettger M, Garg S, Hazkani-Covo E, Martin WF. In press. Endosymbiotic gene transfer from prokaryotic pangenomes: inherited chimaerism in eukaryotes. *Proc. Natl Acad. Sci. USA.* (doi:10.1073/pnas.1421385112)
42. Hansmann S, Martin W. 2000 Phylogeny of 33 ribosomal and six other proteins encoded in an ancient gene cluster that is conserved across prokaryotic genomes: influence of excluding poorly alignable sites from analysis. *Int. J. Syst. Evol. Microbiol*. **50**, 1655–1663. (doi:10.1099/00207713-50-4-1655)
43. Charlebois RL, Doolittle WF. 2004 Computing prokaryotic gene ubiquity: rescuing the core from extinction. *Genome Res*. **14**, 2469–2477. (doi:10.1101/gr.3024704)
44. Cicarelli FD, Doerks T, von Mering C, Creevey CJ, Snel B, Bork P. 2006 Toward automatic reconstruction of a highly resolved tree of life. *Science* **311**, 1283–1287. (doi:10.1126/science.1123061)
45. Esser C *et al.* 2004 A genome phylogeny for mitochondria among alpha-proteobacteria and a predominantly eubacterial ancestry of yeast nuclear genes. *Mol. Biol. Evol*. **21**, 1643–1660. (doi:10.1093/molbev/msh160)
46. Pisani D, Cotton JA, McInerney JO. 2007 Supertrees disentangle the chimerical origin of eukaryotic genomes. *Mol. Biol. Evol*. **24**, 1752–1760. (doi:10.1093/molbev/msm095)
47. Cotton JA, McInerney JO. 2010 Eukaryotic genes of archaebacterial origin are more important than the more numerous eubacterial genes, irrespective of function. *Proc. Natl Acad. Sci. USA* **107**, 17 252–17 255. (doi:10.1073/pnas.1000265107)
48. Lane CE, Archibald JM. 2008 The eukaryotic tree of life: endosymbiosis takes its TOL. *Trends Ecol. Evol*. **23**, 268–275. (doi:10.1016/j.tree.2008.02.004)
49. Embley TM, Hirt RP. 1998 Early branching eukaryotes? *Curr. Opin. Genet. Dev*. **8**, 624–629. (doi:10.1016/S0959-437X(98)80029-4)
50. Cox CJ, Foster PG, Hirt RP, Harris SR, Embley TM. 2008 The archaebacterial origin of eukaryotes. *Proc. Natl Acad. Sci. USA* **105**, 20 356–20 361. (doi:10.1073/pnas.0810647105)
51. Williams TA, Embley M. 2014 Archaeal 'dark matter' and the origin of eukaryotes. *Genome Biol. Evol*. **6**, 474–481. (doi:10.1093/gbe/evu031)
52. 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. R. Soc. B* **278**, 1009–1018. (doi:10.1098/rspb.2010.1427)
53. Martin W. 2005 Archaebacteria (Archaea) and the origin of the eukaryotic nucleus. *Curr. Opin. Microbiol*. **8**, 630–637. (doi:10.1016/j.mib.2005.10.004)
54. Pederson T. 2011 The nucleus introduced. *Cold Spring Harb. Perspect. Biol*. **5**, a000521. (doi:10.1101/cshperspect.a000521)
55. Cavalier-Smith T. 1987 The origin of eukaryote and archaebacterial cells. *Ann. NY Acad. Sci*. **503**, 17–54. (doi:10.1111/j.1749-6632.1987.tb40596.x)
56. Cavalier-Smith T. 1988 Origin of the cell-nucleus. *Bioessays* **9**, 72–78. (doi:10.1002/bies.950090209)
57. Cavalier-Smith T. 2002 The phagotrophic origin of eukaryotes and phylogenetic classification of protozoa. *Int. J. Syst. Evol. Microbiol*. **52**, 297–354. (doi:10.1099/ijs.0.02058-0)
58. Cavalier-Smith T. 2004 Only six kingdoms of life. *Proc. R. Soc. B* **271**, 1251–1262. (doi:10.1098/rspb.2004.2705)

59. Gould GW, Dring GJ. 1979 Possible relationship between bacterial endospore formation and the origin of eukaryotic cells. *J. Theor. Biol*. **81**, 47–53. (doi:10.1016/0022-5193(79)90079-1)
60. Zillig W, Klenk H-P, Palm P, Leffers H, Pühler G, Gropp F, Garrett RA. 1989 Did eukaryotes originate by a fusion event? *Endocyt. Cell Res*. **6**, 1–25.
61. Gupta RS, Golding GB. 1996 The origin of the eukaryotic cell. *Trends Biochem. Sci*. **21**, 166–171. (doi:10.1016/S0968-0004(96)20013-1)
62. Fuerst JA, Webb RI. 1991 Membrane-bounded nucleoid in the eubacterium *Gemmata obscuriglobus*. *Proc. Natl Acad. Sci. USA* **88**, 8184–8188. (doi:10.1073/pnas.88.18.8184)
63. Santarella-Mellwig R, Pruggnaller S, Roos N, Mattaj IW, Devos DP. 2013 Three-dimensional reconstruction of bacteria with a complex endomembrane system. *PLoS Biol*. **11**, e1001565. (doi:10.1371/journal.pbio.1001565)
64. Devos DP. 2013 PVC bacteria: Variation of, but not exception to, the Gram-negative cell plan. *Trends Microbiol.* **22**, 14–20. (doi:10.1016/j.tim.2013.10.008)
65. Searcy DG, Hixon WG. 1991 Cytoskeletal origins in sulfur-metabolizing archaebacteria. *Biosystems* **25**, 1–11. (doi:10.1016/0303-2647(91)90008-9)
66. Lake JA, Rivera MC. 1994 Was the nucleus the first endosymbiont? *Proc. Natl Acad. Sci. USA* **91**, 2880–2881. (doi:10.1073/pnas.91.8.2880)
67. Moreira D, López-García P. 1998 Symbiosis between methanogenic archaea and δ-proteobacteria as the origin of eukaryotes: the syntrophic hypothesis. *J. Mol. Evol*. **47**, 517–530. (doi:10.1007/PL00006408)
68. López-García P, Moreira D. 2006 Selective forces for the origin of the eukaryotic nucleus. *Bioessays* **28**, 525–533. (doi:10.1002/bies.20413)
69. Kuwabara T, Minaba M, Ogi N, Kamekura M. 2007 *Thermococcus celericrescens* sp. nov., a fast-growing and cell-fusing hyperthermophilic archaeon from a deep-sea hydrothermal vent. *Int. J. Syst. Evol. Microbiol*. **57**, 437–443. (doi:10.1099/ijs.0.64597-0)
70. Margulis L, Dolan MF, Guerro R. 2000 The chimeric eukaryote: Origin of the nucleus from the karyomastigont in amitochondriate protists. *Proc. Natl Acad. Sci. USA* **97**, 6954–6959. (doi:10.1073/pnas.97.13.6954)
71. Bell PJL. 2001 Viral eukaryogenesis: Was the ancestor of the nucleus a complex DNA virus? *J. Mol. Evol*. **53**, 251–256. (doi:10.1111/j.1749-6632.2009.04994.x)
72. Horiike T, Hamada K, Miyata D, Shinozawa T. 2004 The origin of eukaryotes is suggested as the symbiosis of *Pyrococcus* into γ-proteobacteria by phylogenetic tree based on gene content. *J. Mol. Evol*. **59**, 606–619. (doi:10.1007/s00239-004-2652-5)
73. Forterre P, Gribaldo S. 2010 Bacteria with a eukaryotic touch: a glimpse of ancient evolution? *Proc. Natl Acad. Sci. USA* **107**, 12 739–12 740. (doi:10.1073/pnas.1007720107)
74. Kurland CG, Collins LJ, Penny D. 2006 Genomics and the irreducible nature of eukaryote cells. *Science* **312**, 1011–1014. (doi:10.1126/science.1121674)
75. Penny D, Collins LJ, Daly TK, Cox SJ. 2014 The relative ages of eukaryotes and akaryotes. *J. Mol. Evol*. **79**, 228–239. (doi:10.1007/s00239-014-9643-y)
76. Field MC, Sali A, Rout MP. 2011 On a bender—BARs, ECRTs, COPs, and finally getting your coat. *J. Cell Biol*. **193**, 963–972. (doi:10.1083/jcb.201102042)
77. Forterre P. 2011 A new fusion hypothesis for the origin of Eukaryota: better than previous ones, but probably also wrong. *Res. Microbiol*. **162**, 77–91. (doi:10.1016/j.resmic.2010.10.005)
78. 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**, 810–817. (doi:10.1002/bies.201100045)
79. Embley TM, Martin W. 2006 Eukaryotic evolution, changes and challenges. *Nature* **440**, 623–630. (doi:10.1038/nature04546)
80. McFadden GI, van Dooren GG. 2004 Evolution: red algal genome affirms a common origin of all plastids. *Curr. Biol*. **14**, R514–R516. (doi:10.1016/j.cub.2004.06.041)
81. Keeling PJ. 2013 The number, speed, and impact of plastid endosymbiosis in eukaryotic evolution. *Annu. Rev. Plant. Biol*. **64**, 583–607. (doi:10.1146/annurev-arplant-050312-120144)
82. Zimorski V, Ku C, Martin WF, Gould SB. 2014 Endosymbiotic theory for organelle origins. *Curr. Opin. Microbiol.* **22**, 38–48. (doi:10.1016/j.mib.2014.09.008)
83. Wallin IE. 1927 *Symbionticism and the origin of species*, 171. London, UK: Bailliere, Tindall and Cox.
84. Wallin IE. 1925 On the nature of mitochondria. IX. Demonstration of the bacterial nature of mitochondria. *Am. J. Anat*. **36**, 131–139. (doi:10.1002/aja.1000360106)
85. Wilson EB. 1928 *The cell in development and heredity*, 3rd revised edtion. New York, NY: Macmillan. (Reprinted 1987 by Garland Publishing, New York.)
86. Buchner P. 1953 *Endosymbiose der Tiere mit pflanzlichen Mikroorganismen*. Basel, Switzerland: Birkhäuser, Basel.
87. Lederberg J. 1952 Cell genetics and hereditary symbiosis. *Physiol. Rev*. **32**, 403–430.
88. Sagan L. 1967 On the origin of mitosing cells. *J. Theoret. Biol*. **14**, 225–274. (doi:10.1016/0022-5193(67)90079-3)
89. Goksøyr J. 1967 Evolution of eucaryotic cells. *Nature* **214**, 1161. (doi:10.1038/2141161a0)
90. Knoll AH. 2012 Lynn Margulis, 1938–2011. *Proc. Natl Acad. Sci. USA* **109**, 1022. (doi:10.1073/pnas.1120472109)
91. Margulis L. 1970 *Origin of eukaryotic cells*. New Haven, CT: Yale University Press.
92. de Duve C. 1969 Evolution of the peroxisome. *Ann. NY Acad. Sci*. **168**, 369–381. (doi:10.1111/j.1749-6632.1969.tb43124.x)
93. Stanier Y. 1970 Some aspects of the biology of cells and their possible evolutionary significance. *Symp. Soc. Gen. Microbiol*. **20**, 1–38.
94. Raff RA, Mahler HR. 1972 The non symbiotic origin of mitochondria. *Science* **177**, 575–582. (doi:10.1126/science.177.4049.575)
95. Bogorad L. 1975 Evolution of organelles and eukaryotic genomes. *Science* **188**, 891–898. (doi:10.1126/science.1138359)
96. Cavalier-Smith T. 1975 The origin of nuclei and of eukaryotic cells. *Nature* **256**, 463–468. (doi:10.1038/256463a0)
97. Bonen L, Doolittle WF. 1975 On the prokaryotic nature of red algal chloroplasts. *Proc. Natl Acad. Sci. USA* **72**, 2310–2314. (doi:10.1073/pnas.72.6.2310)
98. Anderson S *et al.* 1981 Sequence and organization of the human mitochondrial genome. *Nature* **290**, 457–465. (doi:10.1038/290457a0)
99. John P, Whatley FR. 1975 *Paracoccus denitrificans* and the evolutionary origin of the mitochondrion. *Nature* **254**, 495–498. (doi:10.1038/254495a0)
100. Woese CR. 1977 Endosymbionts and mitochondrial origins. *J. Mol. Evol*. **10**, 93–96. (doi:10.1007/BF01751802)
101. Van Valen LM, Maiorana VC. 1980 The Archaebacteria and eukaryotic origins. *Nature* **287**, 248–250. (doi:10.1038/287248a0)
102. Doolittle WF. 1980 Revolutionary concepts in evolutionary biology. *Trends Biochem. Sci*. **5**, 146–149. (doi:10.1016/0968-0004(80)90010-9)
103. Margulis L. 1981 *Symbiosis in cell evolution*. San Francisco, CA: Freeman.
104. Vellai T, Vida G. 1998 The origin of the eukaryotes: the difference between prokaryotic and eukaryotic cells. *Proc. R. Soc. Lond. B* **266**, 1571–1577. (doi:10.1098/rspb.1999.0817)
105. Searcy DG. 1992 Origins of mitochondria and chloroplasts from sulphurbased symbioses. In *The origin and evolution of the cell* (eds H Hartman, K Matsuno), pp. 47–78. Singapore: World Scientific.
106. Martijn J, Ettema TJG. 2013 From archaeon to eukaryote: The evolutionary dark ages of the eukaryotic cell. *Biochem. Soc. Trans*. **41**, 451–457. (doi:10.1042/BST20120292)
107. Gray MW. 2014 The pre-endosymbiont hypothesis: a new perspective on the origin and evolution of mitochondria. *Cold Spring Harb. Perspect. Biol*. **6**, a016097. (doi:10.1101/cshperspect.a016097)
108. Baum DA, Baum B. 2014 An inside-out origin for the eukaryotic cell. *BMC Biol*. **12**, 76. (doi:10.1186/s12915-014-0076-2)
109. Davidov Y, Jurkevitch E. 2009 Predation between prokaryotes and the origin of eukaryotes. *Bioessays* **31**, 748–757. (doi:10.1002/bies.200900018)
110. Davidov Y, Huchon D, Koval SF, Jurkevitch E. 2006 A new alpha-proteobacterial clade of *Bdellovibrio*-like predators: implications for the mitochondrial endosymbiotic theory. *Environ. Microbiol*. **8**, 2179–2188. (doi:10.1111/j.1462-2920.2006.01101.x)
111. Pasternak Z *et al.* 2014 In and out: An analysis of epibiotic vs periplasmic bacterial predators. *ISME J*. **8**, 625–635. (doi:10.1038/ismej.2013.164)
112. Goldberg AV *et al.* 2008 Localization and functionality of microsporidian iron-sulphur cluster assembly proteins. *Nature* **452**, 624–628. (doi:10.1038/nature06606)
113. Tsaousis AD, Kunji ERS, Goldberg AV, Lucocq JM, Hirt RP, Embley TM. 2008 A novel route for ATP acquisition by the remnant mitochondria of *Encephalitozoon cuniculi*. *Nature* **453**, 553–556. (doi:10.1038/nature06903)

114. Williams BAP, Hirt RP, Lucocq JM, Embley TM. 2002 A mitochondrial remnant in the microsporidian *Trachipleistophora hominis*. *Nature* **418**, 865–869. (doi:10.1038/nature00949)
115. Brown JR, Doolittle WF. 1997 Archaea and the prokaryote-to-eukaryote transition. *Microbiol. Mol. Biol. Rev*. **61**, 456–502.
116. Martin W, Brinkmann H, Savona C, Cerff R. 1993 Evidence for a chimeric nature of nuclear genomes: Eubacterial origin of eukaryotic glyceraldehyde-3-phosphate dehydrogenase genes. *Proc. Natl Acad. Sci. USA* **90**, 8692–8696. (doi:10.1073/pnas.90.18.8692)
117. Timmis JN, Ayliffe MA, Huang CY, Martin W. 2004 Endosymbiotic gene transfer: Organelle genomes forge eukaryotic chromosomes. *Nat. Rev. Genet*. **5**, 123–135. (doi:10.1038/nrg1271)
118. López-García P, Moreira D. 1999 Metabolic symbiosis at the origin of eukaryotes. *Trends Biochem. Sci*. **24**, 88–93. (doi:10.1016/S0968-0004(98)01342-5)
119. Finlay BJ, Embley TM, Fenchel T. 1993 A new polymorphic methanogen, closely related to *Methanocorpusculum parvum*, living in stable symbiosis within the anaerobic ciliate *Trimyema* sp. *J. Gen. Microbiol*. **139**, 371–378. (doi:10.1099/00221287-139-2-371)
120. von Dohlen CD, Kohler S, Alsop ST, McManus WR. 2001 Mealybug β-proteobacterial endosymbionts contain γ-proteobacterial symbionts. *Nature* **412**, 433–436. (doi:10.1038/35086563)
121. Husnik F *et al.* 2013 Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. *Cell* **153**, 1567–1578. (doi:10.1016/j.cell.2013.05.040)
122. Yutin N, Wolf MY, Wolf YI, Koonin EV. 2009 The origins of phagocytosis and eukaryogenesis. *Biol. Direct* **4**, 9. (doi:10.1186/1745-6150-4-9)
123. Nelson-Sathi S, Dagan T, Landan G, Janssen A, Steel M, McInerney JO, Deppenmeier U, Martin WF. 2012 Acquisition of 1,000 eubacterial genes physiologically transformed a methanogen at the origin of Haloarchaea. *Proc. Natl Acad. Sci. USA* **109**, 20 537–20 542. (doi:10.1073/pnas.1209119109)
124. Nelson-Sathi S *et al.* 2015 Origins of major archaeal clades correspond to gene acquisitions from bacteria. *Nature* **517**, 77–80. (doi:10.1038/nature13805)
125. Say RF, Fuchs G. 2010 Fructose 1,6-bisphosphate aldolase/phosphatase may be an ancestral gluconeogenic enzyme. *Nature* **464**, 1077–1081. (doi:10.1038/nature08884)
126. Reher M, Fuhrer T, Bott M, Schönheit P. 2010 The nonphosphorylative Entner-Doudoroff pathway in the thermoacidophilic euryarchaeon *Picrophilus torridus* involves a novel 2-keto-3-deoxygluconate-specific aldolase. *J. Bacteriol*. **192**, 964–974. (doi:10.1128/JB.01281-09)
127. Bräsen C, Esser D, Rauch B, Siebers B. 2014 Carbohydrate metabolism in Archaea: current insights into unusual enzymes and pathways and their regulation. *Microbiol. Mol. Biol. Rev*. **78**, 89–175. (doi:10.1128/MMBR.00041-13)
128. Hannaert V 2000 Enolase from *Trypanosoma brucei*, from the amitochondriate protist *Mastigamoeba balamuthi*, and from the chloroplast and cytosol of *Euglena gracilis*: pieces in the evolutionary puzzle of the eukaryotic glycolytic pathway. *Mol. Biol. Evol*. **17**, 989–1000. (doi:10.1093/oxfordjournals.molbev.a026395)
129. Lange BM, Rujan T, Martin W, Croteau R. 2000 Isoprenoid biosynthesis: The evolution of two ancient and distinct pathways across genomes. *Proc. Natl Acad. Sci. USA* **97**, 13 172–13 177. (doi:10.1073/pnas.240454797)
130. Waldbauer JR, Newman DK, Summons RE. 2011 Microaerobic steroid biosynthesis and the molecular fossil record of Archean life. *Proc. Natl Acad. Sci. USA* **108**, 13 409–13 414. (doi:10.1073/pnas.1104160108)
131. Poole A, Jeffares D, Penny D. 1999 Early evolution: prokaryotes, the new kids on the block. *Bioessays* **21**, 880–889. (doi:10.1002/(SICI)1521-1878(199910)21:10<880::AID-BIES11>3.0.CO;2-P)
132. Lambowitz AM, Zimmerly S. 2011 Group II introns: Mobile ribozymes that invade DNA. *Cold Spring Harb. Perspect. Biol*. **3**, a003616. (doi:10.1101/cshperspect.a003616)
133. Lambowitz AM, Zimmerly S. 2004 Mobile group II introns. *Annu. Rev. Genet*. **38**, 1–35. (doi:10.1146/annurev.genet.38.072902.091600)
134. Matsuura M *et al.* 1997 A bacterial group II intron encoding reverse transcriptase, maturase, and DNA endonuclease activities: biochemical demonstration of maturase activity and insertion of new genetic information within the intron. *Genes Dev*. **11**, 2910–2924. (doi:10.1101/gad.11.21.2910)
135. Lynch M, Richardson AO. 2002 The evolution of spliceosomal introns. *Curr. Opin. Genet. Dev*. **12**, 701–710. (doi:10.1016/S0959-437X(02)00360-X)
136. Audibert A, Weil D, Dautry F. 2002 *In vivo* kinetics of mRNA splicing and transport in mammalian cells. *Mol. Cell. Biol*. **22**, 6706–6718. (doi:10.1128/MCB.22.19.6706-6718.2002)
137. Collins L, Penny D. 2005 Complex spliceosomal organization ancestral to extant eukaryotes. *Mol. Biol. Evol*. **22**, 1053–1066. (doi:10.1093/molbev/msi091)
138. 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. *Curr. Biol*. **13**, 1512–1517. (doi:10.1016/S0960-9822(03)00558-X)
139. Roy SW, Gilbert W. 2005 Rates of intron loss and gain: implications for early eukaryotic evolution. *Proc. Natl Acad. Sci. USA* **102**, 5773–5778. (doi:10.1073/pnas.0500383102)
140. Mans BJ, Anantharaman V, Aravind L, Koonin EV. 2004 Comparative genomics, evolution and origins of the nuclear envelope and nuclear pore complex. *Cell Cycle* **3**, 1612–1637. (doi:10.4161/cc.3.12.1345)
141. Staub E, Fiziev P, Rosenthal A, Hinzmann B. 2004 Insights into the evolution of the nucleolus by an analysis of its protein domain repertoire. *Bioessays* **26**, 567–581. (doi:10.1002/bies.20032)
142. Ptashne M. 2013 Epigenetics: core misconcept. *Proc. Natl Acad. Sci. USA* **110**, 7101–7103. (doi:10.1073/pnas.1305399110)
143. Tielens AGM, Rotte C, van Hellemond JJ, Martin W. 2002 Mitochondria as we don't know them. *Trends Biochem. Sci*. **27**, 564–572. (doi:10.1016/S0968-0004(02)02193-X)
144. Martin W *et al.* 2002 Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc. Natl Acad. Sci. USA* **99**, 12 246–12 251. (doi:10.1073/pnas.182432999)
145. Rotte C, Stejskal F, Zhu G, Keithly JS, Martin W. 2001 Pyruvate:NADP$^+$ oxidoreductase from the mitochondrion of *Euglena gracilis* and from the apicomplexan *Cryptosporidium parvum*: a fusion of pyruvate:ferredoxin oxidoreductase and NADPH-cytochrome P450 reductase. *Mol. Biol. Evol*. **18**, 710–720. (doi:10.1093/oxfordjournals.molbev.a003853)
146. Hoffmeister M, van der Klei A, Rotte C, van Grinsven KWA, van Hellemond JJ, Henze K, Tielens AGM, Martin W. 2004 *Euglena gracilis* rhodoquinone:ubiquinone ratio and mitochondrial proteome differ under aerobic and anaerobic conditions. *J. Biol. Chem.* **279**, 22 422–22 429. (doi:10.1074/jbc.M400913200)
147. Atteia A, van Lis R, Gelius-Dietrich G, Adrait A, Garin J, Joyard J, Rolland N, Martin W. 2006 Pyruvate:formate lyase and a novel route of eukaryotic ATP-synthesis in anaerobic *Chlamydomonas* mitochondria. *J. Biol. Chem*. **281**, 9909–9918. (doi:10.1074/jbc.M507862200)
148. Atteia A *et al.* 2009 A proteomic survey of *Chlamydomonas reinhardtii* mitochondria sheds new light on the metabolic plasticity of the organelle and on the nature of the α-proteobacterial mitochondrial ancestor. *Mol. Biol. Evol.* **29**, 1533–1548. (doi:10.1093/molbev/msp068)
149. Atteia A, van Lis R, Tielens AGM, Martin WF. 2013 Anaerobic energy metabolism in unicellular photosynthetic eukaryotes. *Biochim. Biophys. Acta* **1827**, 210–223. (doi:10.1016/j.bbabio.2012.08.002)
150. Price DC *et al.* 2012 *Cyanophora paradoxa* genome elucidates origin of photosynthesis in algae and plants. *Science* **335**, 843–847. (doi:10.1126/science.1213561)
151. Ku C, Roettger M, Zimorski V, Nelson-Sathi S, Sousa FL, Martin WF. 2014 Plastid origin: who, when and why? *Acta Soc. Bot. Pol.* **83**, 281–289. (doi:10.5586/asbp.2014.045)
152. Domman D, Horn M, Embley TM, Williams TA. 2015 Plastid establishment did not require a chlamydial partner. *Nat. Comm*. **6**, 6421. (doi:10.1038/ncomms7421)
153. Deusch O, Landan G, Roettger M, Gruenheit N, Kowallik KV, Allen JF, Martin W, Dagan T. 2008 Genes of cyanobacterial origin in plant nuclear genomes point to a heterocyst-forming plastid ancestor. *Mol. Biol. Evol*. **25**, 748–761. (doi:10.1093/molbev/msn022)

154. Dagan T *et al.* 2013 Genomes of stigonematalean cyanobacteria (Subsection V) and the evolution of oxygenic photosynthesis from prokaryotes to plastids. *Genome Biol. Evol.* **5**, 31–44. (doi:10.1093/gbe/evs117)
155. Dolezal P, Likic V, Tachezy J, Lithgow T. 2006 Evolution of the molecular machines for protein import into mitochondria. *Science* **313**, 314–318. (doi:10.1126/science.1127895)
156. Schleiff EE, Becker TT. 2011 Common ground for protein translocation: Access control for mitochondria and chloroplasts. *Nat. Rev. Mol. Cell Biol.* **12**, 48–59. (doi:10.1038/nrm3027)
157. Gould SB, Waller RF, McFadden GI. 2008 Plastid evolution. *Annu. Rev. Plant Biol.* **59**, 491–517. (doi:10.1146/annurev.arplant.59.032607.092915)
158. Curtis BA *et al.* 2012 Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs. *Nature* **492**, 59–65. (doi:10.1038/nature11681)
159. Gould SB. 2012 Evolutionary genomics: algae's complex origins. *Nature* **492**, 46–48. (doi:10.1038/nature11759)
160. Ball SG *et al.* 2013 Metabolic effectors secreted by bacterial pathogens: essential facilitators of plastid endosymbiosis. *Plant Cell* **25**, 7–21. (doi:10.1105/tpc.112.101329)
161. Moreira D, Deschamps P. 2014 What was the real contribution of endosymbionts to the eukaryotic nucleus? Insights from photosynthetic eukaryotes. *Cold Spring Harb. Perspect. Biol.* **6**, a016014. (doi:10.1101/cshperspect.a016014)
162. Deschamps P. 2014 Primary endosymbiosis: have cyanobacteria and chlamydiae ever been roommates? *Acta Soc. Bot. Pol.* **83**, 291–302. (doi:10.5586/asbp.2014.048)
163. Degli-Esposti M. 2014 Bioenergetic evolution in proteobacteria and mitochondria. *Genome Biol. Evol.* **6**, 3238–3251. (doi:10.1093/gbe/evu257)
164. Allen JF. 1993 Control of gene-expression by redox potential and the requirement for chloroplast and mitochondrial genomes. *J. Theor. Biol.* **165**, 609–631. (doi:10.1006/jtbi.1993.1210)
165. Allen JF. 2003 The function of genomes in bioenergetic organelles. *Phil. Trans. R. Soc. Lond. B* **358**, 19–37. (doi:10.1098/rstb.2002.1191)
166. Lane N. 2015 *The vital question: why is life the way it is?* London, UK: Profile Books.
167. Von Heijne G. 1986 Why mitochondria need a genome. *FEBS Lett.* **198**, 1–4. (doi:10.1016/0014-5793(86)81172-3)
168. Koonin EV, Yutin N. 2014 The dispersed archaeal eukaryome and the complex archaeal ancestor of eukaryotes. *Cold Spring Harb. Perspect. Biol.* **6**, a016188. (doi:10.1101/cshperspect.a016188)
169. Williams TA, Foster PG, Nye TMW, Cox CJ, Embley TM. 2012 A congruent phylogenomic signal places eukaryotes within the Archaea. *Proc. R. Soc. B* **279**, 4870–4879. (doi:10.1098/rspb.2012.1795)
170. Petitjean C, Deschamps P, López-García P, Moreira D. 2014 Rooting the domain archaea by phylogenomic analysis supports the foundation of the new kingdom Proteoarchaeota. *Genome Biol. Evol.* **7**, 191–204. (doi:10.1093/gbe/evu274)
171. Liu YC, Beer LL, Whitman WB. 2012 Methanogens: a window into ancient sulphur metabolism. *Trends Micobiol.* **20**, 251–258. (doi:10.1016/j.tim.2012.02.002)
172. Ueno Y, Yamada K, Yoshida N, Maruyama S, Isozaki Y. 2006 Evidence from fluid inclusions for microbial methanogenesis in the early Archaean era. *Nature* **440**, 516–519. (doi:10.1038/nature04584)
173. Decker K, Jungermann K, Thauer RK. 1970 Energy production in anaerobic organisms. *Angew. Chem. Int. Ed. Engl.* **9**, 138–158. (doi:10.1002/anie.197001381)
174. Proskurowski G, Lilley MD, Seewald JS, Fruh-Green GL, Olson EJ, Lupton JE, Sylva SP, Kelley DS. 2008 Abiogenic hydrocarbon production at Lost City hydrothermal field. *Science* **319**, 604–607. (doi:10.1126/science.1151194)
175. McCollom TM, Seewald JS. 2013 Serpentinites, hydrogen, and life. *Elements* **9**, 129–134. (doi:10.2113/gselements.9.2.129)
176. Schrenk MO, Brazelton WJ, Lang SQ. 2013 Serpentinization, carbon, and deep life. *Rev. Mineral. Geochem.* **75**, 575–606. (doi:10.2138/rmg.2013.75.18)
177. Russell MJ, Hall AJ, Martin W. 2010 Serpentinization as a source of energy at the origin of life. *Geobiology* **8**, 355–371. (doi:10.1111/j.1472-4669.2010.00249.x)
178. Martin W, Russell MJ. 2007 On the origin of biochemistry at an alkaline hydrothermal vent. *Phil. Trans. R. Soc. B* **362**, 1887–1925. (doi:10.1098/rstb.2006.1881)
179. Lane N, Martin WF. 2012 The origin of membrane bioenergetics. *Cell* **151**, 1406–1416. (doi:10.1016/j.cell.2012.11.050)
180. Martin WF, Sousa FL, Lane N. 2014 Energy at life's origin. *Science* **344**, 1092–1093. (doi:10.1126/science.1251653)
181. Sousa FL, Martin WF. 2014 Biochemical fossils of the ancient transition from geoenergetics to bioenergetics in prokaryotic one carbon compound metabolism. *Biochim. Biophys. Acta* **1837**, 964–981. (doi:10.1016/j.bbabio.2014.02.001)

Publication II

Bacterial vesicle secretion and the evolutionary origin of the eukaryotic endomembrane system

Sven B. Gould[1], **Sriram G. garg**[1] and William F. Martin[1]

1 - Institute for Molecular Evolution, Heinrich-Heine Universität Düsseldorf, Universitätsstraße 1, Düsseldorf 40225, Germany

The presented manuscript was peer-reviewed and published in **Trends in Microbiology** with an impact factor of **9.808** on 31st March 2016. It was also featured as the cover article of the Journal and had video abstract associated with it

Contribution as **second author**

33.33%

Conceived portions of the hypothesis including mining the literature for relevant citations. Wrote parts of the text and discussion. Designed and compiled a video for *Trends in Microbiology* to be featured in their YouTube channel “cellvideoabstracts”

CellPress

Opinion

Bacterial Vesicle Secretion and the Evolutionary Origin of the Eukaryotic Endomembrane System

Sven B. Gould,[1,*] Sriram G. Garg,[1] and William F. Martin[1,*]

Eukaryotes possess an elaborate endomembrane system with endoplasmic reticulum, nucleus, Golgi, lysosomes, peroxisomes, autophagosomes, and dynamic vesicle traffic. Theories addressing the evolutionary origin of eukaryotic endomembranes have overlooked the outer membrane vesicles (OMVs) that bacteria, archaea, and mitochondria secrete into their surroundings. We propose that the eukaryotic endomembrane system originated from bacterial OMVs released by the mitochondrial ancestor within the cytosol of its archaeal host at eukaryote origin. Confined within the host's cytosol, OMVs accumulated naturally, fusing either with each other or with the host's plasma membrane. This matched the host's archaeal secretory pathway for cotranslational protein insertion with outward bound mitochondrial-derived vesicles consisting of bacterial lipids, forging a primordial, secretory endoplasmic reticulum as the cornerstone of the eukaryotic endomembrane system.

Eukaryogenesis: A Matter of Compartmentalisation

Among the many traits that distinguish eukaryotic from prokaryotic cells, none is more conspicuous or significant than the eukaryotic **endomembrane system** (see Glossary). Like other eukaryotic-specific traits, such as mitosis and sex, its evolutionary origin remains obscure. The compartments of the endomembrane system are present throughout the major eukaryotic groups, as are the proteins that are specific to them [1]. Hence both were present in the eukaryote common ancestor [2], for which reason thoughts on the origin of the endomembrane system are linked to thoughts on the origin of eukaryotes themselves.

Despite many differences in their mechanistic details, theories for the origin of the endomembrane system traditionally derive it from inward invaginations of the plasma membrane, such that the endoplasmic reticulum (ER) lumen is topologically homologous to the environment [1,3–6]. This is true for theories that posit autogenous (nonsymbiotic) eukaryote origins [7] and for theories that posit eukaryotes to descend from symbiotic associations of prokaryotes [8]. Though most current theories now posit that mitochondria arose in an archaeal host through endosymbiosis (Box 1), the question of how the merger of two prokaryotic cells gave rise to a cell possessing a eukaryotic endomembrane system with elaborate vesicle trafficking (Figure 1) remains unanswered, as does the question of how **archaeal lipids** of the host's plasma membrane came to be replaced by bacterial lipids.

Though prokaryotes do not generate intracellular vesicle traffic of the kind found in eukaryotes, they do indeed generate OMVs, but these are secreted outwardly into the environment, not

Trends

Eukaryogenesis models struggle with explaining the origin of the endomembrane system and the transition from an archaeal plasma membrane based on isoprene ethers to a bacterial-type membrane based on fatty acid esters.

Bacteria and archaea secrete outer menbrane vesicles (OMVs) into their surroundings. If the endosymbiont that became the mitochondrion did so in the archaeal host, it physically generated the first vesicles of the endomembrane system.

Endosymbiont OMVs could only accumulate in the host's cytosol – fusion with each other could have generated compartments, fusion with the archaeal plasma membrane could have converted its chemical composition.

Starting endomembrane origin with outward flux of endosymbiont-derived OMVs integrates mitochondria, their lipids, and their energetics into current models of eukaryote origin, explaining why eukaryotes had a mitochondrion-bearing ancestor.

[1]Institute for Molecular Evolution, University of Düsseldorf, 40225 Düsseldorf, Germany

*Correspondence: gould@hhu.de (S.B. Gould) and bill@hhu.de (W.F. Martin).

Box 1. Endosymbiosis at Eukaryote Origin

The origin of eukaryotes hinges upon endosymbiosis, and eukaryotic cell complexity arose in the wake of mitochondrial origin, not as its prerequisite [57]. From the genomic standpoint a consensus is emerging that the origin of eukaryotes involved only two distinct partners: an archaeal host cell and an α-proteobacterial endosymbiont that became the mitochondrion [29,43,44,57,71,74–76]. This consensus does not touch upon whether the archaeal host bore a nucleus or not, but several issues require consideration concerning this discrepancy. It concerns, in particular, the purpose of a nucleus in an archaeal cell with cotranscriptional translation that remains unanswered in gradual models for eukaryogenesis that place the origin of the nucleus before that of the mitochondrion.

The selective pressures that brought forth the possession of the nuclear envelope (NE) as a permanent fixture of eukaryotic cells are, we suggest, distinct from the OMV-dependent ER origin of the NE itself. The presence of spliceosomes in the eukaryote common ancestor suggests that the initial selective advantage of possessing an NE was the spatiotemporal separation of spliceosomal splicing from translation, with spliceosomal introns stemming from group II introns acquired via endosymbiotic gene transfer from the mitochondrial symbiont [77]. Spliceosomal splicing requires a nucleus to exclude active ribosomes from intron-containing transcripts, because ribosomes operate much more rapidly than spliceosomes, such that cotranscriptional translation on nascent transcripts bearing spliceosomal introns would lead to defective polypeptides only. The physical exclusion of ribosomes from active chromatin via membranes would allow the slow process of splicing to go to completion before translation sets in. Similar to the intron hypothesis for the origin of the nucleus [77], our present suggestion for the origin of the endomembrane system requires a non-nucleated archaeal host with cotranscriptional translation at the origin of mitochondria.

inwardly into the cytosol. Decades ago, microbiologists observed that Gram-negative bacteria can secrete lipopolysaccharide (LPS) complexes [9] that presumably stem from the outer membrane [10] into the environment. As explained in the next section, quite a bit is now known about prokaryotic OMVs, but less about the proteins involved, which are, in some cases, homologous to those germane to vesicle scission into eukaryotic **multivesicular bodies (MVBs)** for example. Moreover, even mitochondria themselves are known to secrete **mitochondria-derived vesicles** (MDVs; Figure 1) into the cytosol [11–14]. No previous theory for the origin of the eukaryotic endomembrane system, however, incorporates the observations available for prokaryotic OMVs. Here we close that gap with an evolutionary inference that accounts for the origin of the eukaryotic endomembrane system in a novel and natural manner.

Prokaryotic Vesicle Secretion

As Deatherage and Cookson [15] write, it has long been known, but underappreciated, that bacteria and archaea generate OMVs. Both Gram-negative [16] and Gram-positive [17] bacteria secrete OMVs that stem from their outer membrane (Figure 2). In addition, some bacteria form nanowires, long tube-like protrusions of the outer membrane [18]. Bacterial OMV cargo ranges from outer membrane proteins to the content of the periplasmic space, which can be specifically apportioned for inclusion into OMVs [19]. OMVs are also clinically important as they can include key toxins associated with bacterial virulence and toxicity [20,21]. The rate of OMV secretion and the nature of their content can vary according to nutrient availability, stress, host–pathogen interactions, and exposure to antibiotics such as gentamicin [9,20]. The mechanistic details behind OMV release are still poorly understood, but in Gram-negative bacteria the release of OMVs is thought to result from the interplay of peptidoglycan, surface proteins, and the LPS complexes themselves [10,15,16,21,22].

Archaea also secrete OMVs [15,23], which contain proteins of the S-layer, components of the outer membrane [24], and in some cases also toxins [25]. The release of archaeal OMVs involves the Cdv (cell division) proteins A, B, and C [24,26], which are homologous to members of the eukaryotic **ESCRT** III protein family involved in membrane vesicle scission [27]. In addition to their role in OMV secretion, archaeal Cdv proteins are involved in cell division (Figure 2). While bacteria require FtsZ for cell division, many archaea lack FtsZ, with the formation of the division ring and the final scission of the daughter cells being mediated by Cdv proteins [26]. Similar to their role in cell division [26,27], Cdv proteins could aid in the tethering and scission of the membraneous neck that leads to the release of the nascent OMV from the archaeal plasma

Glossary

Archaeal lipids: membrane lipids composed of isoprenoid hydrocarbon side chains linked via an ether bond to glycerol-1-phosphate.
Autophagosomes: double-membrane-bound compartments involved in the degradation of intracellular proteins and organelles through autophagy. Outer membrane fuses with the lysosome to form the autolysosome.
Bacterial lipids: membrane lipids composed of a glycerol-3-phosphate linked to fatty acid side chains via an ester linkage.
Coatomer: class of proteins involved in vesicle coat formation. Many share a similar domain architecture uniting a β-propeller and an α-solenoid domain.
Endomembrane system: elaborate membrane system unique to eukaryotes; it includes the nucleus, the endoplasmic reticulum, the Golgi apparatus, the lysosome, the peroxisome, autophagosomes, and the myriad vesicle-trafficking processes that interconnect them with each other and the plasma membrane.
Endosomal sorting complex required for transport (ESCRT): multicomponent machinery subdivided into ESCRT-0, I, II, III; it facilitates membrane vesicle budding 'away' from the cytoplasm.
Flagellar pore complex (FPC): also known as the ciliary pore complex, a structure composed of many proteins that share a high degree of homology with the nuclear pore complex (NPC) and regulates transport into the flagellum.
Glyoxysome: specialized type of peroxisome found in plants and some fungi.
Golgi apparatus: highly dynamic structure of ordered stacks that act as a sorting station for vesicular trafficking from ER to the plasma membrane and other compartments.
Lokiarchaea: recently discovered archaeal phylum that monophyletically branches with eukaryotes.
Lysosome: acidified compartment and final destination for the degradation of proteins and particles coming from multivesicular bodies (MVBs).
Mitochondria-derived vesicles (MDVs): vesicles that originate from the mitochondria and fuse with

membrane into the environment [24]. Importantly, prokaryotic OMV flux, whether bacterial or archaeal, is not inward but outward (Figure 2) in cases reported to date.

A Bacterial Vesicle Model for the Origin of the ER

The essence of our proposal is that the α-proteobacterial ancestor of mitochondria was also able to produce OMVs, that it did so as it became an endosymbiont in its archaeal host, and that those OMVs provided the initial seed of the eukaryotic endomembrane system. This suggestion is compatible with the widespread production of OMVs among modern prokaryotes and with the more recent observation that mitochondria themselves generate MDVs within eukaryotic cells today [11–14]. Upon endosymbiosis, the archaeal secretion system (**SecY/Sec61p**) and its associated **N-glycosylation** machinery integrated readily into the endosymbiont's OMVs, giving rise to a primordial ER that provided the founding stock from which all other endomembrane compartments, including the nucleus, arose (Figure 3, Key Figure).

In terms of the number and nature of evolutionary innovations required to evolve a basic endomembrane system with selectable ER function, our minimal premises can hardly be underbid. We require that the eukaryote common ancestor possessed a mitochondrial symbiont [28], which is now consensus among evolutionary cell biologists [2,29]. We also require that the host was a normal archaeon [28], lacking both a nucleus and its own pre-existing endomembrane system at mitochondrial acquisition, its archaeal chromosomes located in the cytosol and subject to cotranscriptional translation [30]. Thus, our model requires a mechanism of entry for the endosymbiont that is independent of **phagocytosis** and thus demands that one prokaryote can become an endosymbiont within another prokaryote. Clear examples for such symbioses do indeed abound [31,32]. Phagocytosis is not a prerequisite to endosymbiosis and in light of archaeal physiology is problematic for reasons discussed below (and in Box 2).

If the OMV-producing ancestor of mitochondria continues with its natural activity of producing OMVs consisting of bacterial lipids in an archaeal host with cytosolic chromosomes, what happens? Quite a lot happens, and a quite sudden transition appears possible, too, without even requiring evolutionary inventions, merely spatial reorientation of pre-existing host (archaeal) and symbiont (bacterial) components by virtue of endosymbiosis (Figure 3). Notably, prior to endosymbiosis, the symbiont's OMVs diffuse into the environment. In the closed quarters of an archaeal cytosol, the membrane vesicles have no place to go. They can fuse, either with themselves to generate larger vesicular compartments, or with the plasma membrane to export their contents to the cell exterior. The former generates a basic ER topology. The latter constitutes, we propose, the ancestral outward state of eukaryotic membrane flux, and furthermore converts the chemical composition of the host's plasma membrane from isoprene ethers to bacterial fatty acid esters. Importantly, these three salient eukaryotic traits – cytosolic vesicles, outward membrane flux, and the accumulation of bacterial lipids in the archaeal plasma membrane – arise without need for any evolutionary invention (Figure 3). They arise as the result of an OMV-producing bacterium living as an endosymbiont within an archaeon.

The presence of bacterial OMVs in the archaeal cytosol provide a fundamentally new and continuously arising membrane system in the host's cytosol. The consequence is that proteins once destined to the host's plasma membrane via the secretory pathways now have an additional, alternative target: cytosolic OMVs. Accordingly, the host's SecY/Sec61p system, which, prior to symbiosis, facilitated the cotranslational insertion of membrane proteins into the only membrane facing the cytosol – the plasma membrane – now has a new target: OMVs of endosymbiotic origin. This formed a primordial ER membrane architecture, which matches exactly the topology of the modern eukaryotic SEC complex at the rough ER (Figures1 and 3). Initially, this primitive vesicle flux to the plasma membrane required not a single invention, only endosymbiosis, and it provided ground for natural selection (e.g., of **coatomer** proteins and

various other compartments such as peroxisome and MVBs.

Multivesicular bodies (MVBs): membrane-bound compartments containing cytoplasm-derived vesicles destined for degradation at the lysosome.

N-glycosylation: adds oligosaccharide side chains to certain asparagines in proteins. Typically occurs in the ER lumen of eukaryotes.

Nuclear pore complex (NPC): multiprotein complex that spans the nuclear envelope and regulates transport. Many NPC proteins share similarities with proteins of the flagellar pore complex (FPC) and vesicle coat.

Peroxisome: compartment involved in the catabolism of fatty acids, polyamines, and hydrogen peroxide.

Phagocytosis: uptake of large particles such as entire bacterial cells by macrophages or amoebae. Food-particle-containing phagosomes fuse with MVBs and ultimately the lysosome for degradation.

Ribophorin I: protein of the rough ER that binds to the SEC complex, promoting N-glycosylation by serving as a substrate-specific chaperone.

Sarcoplasmic/endoplasmic reticulum Ca^{2+} ATPase (SERCA): a P-type ATPase found in the ER that regulates Ca^{2+} storage in the ER lumen.

SecY/Sec61p: main translocon of the SEC complex involved in translocation of nascent polypeptides from ribosomes into the ER lumen.

V-ATPase: a type of proton pump that acidifies compartments, commonly found in vacuoles and lysosomes. The eukaryotic V-ATPase shares significant homology with the archaeal plasma membrane ATPase (or the A-ATPase).

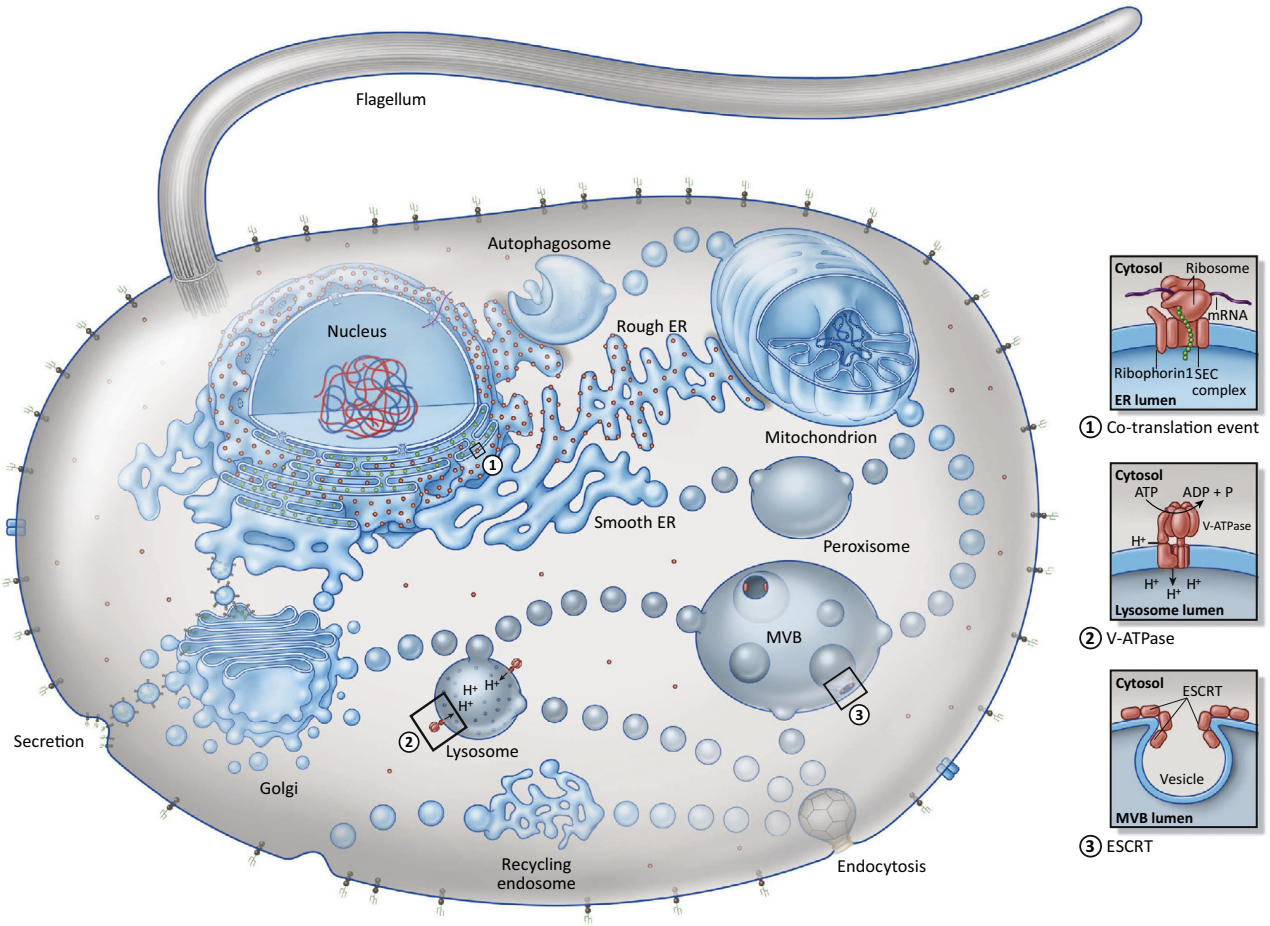


Figure 1. The Eukaryotic Cell. The one decisive trait of the eukaryotic cell is its elaborate endomembrane system. At its centre stands the smooth and rough endoplasmic reticulum (ER) [78], the latter being studded with ribosomes that cotranslationally transport proteins across the SEC complex [41]. For N-glycosylation, **ribophorin I** associates with the Sec61 translocon and serves as a substrate-specific chaperone [79]. Vesicles that bud from the ER can transverse the Golgi – for further modification of cargo and lipids and subsequent sorting – or generate, and continuously supply, other compartments, such as the peroxisome and phagosome [50]. Mitochondria-derived vesicles (MDVs) help to form autophagosomes that originate at the ER–mitochondria contact sites [40] and peroxisomes [12]. Multivesicular bodies (MVBs) represent specialised endosome-associated compartments that contain internal vesicles [80]. They also receive MDVs for subsequent degradation at the lysosome [13]. ESCRT proteins mediate the scission of membranes to release vesicles into the MVBs [80]. The endomembrane system of the eukaryotic cell is a merger of host (red) and endosymbiont (blue) components.

targeted flux) to work on. Because our proposal posits the endosymbiont's OMVs to physically generate the primordial ER, it directly accounts for the observation that only eukaryotes, the cells descended from a mitochondrion-bearing ancestor, possess a *bona fide* endomembrane system. In the following, we briefly consider the properties and components of the eukaryotic endomembrane system, and their homologies.

Connections of the ER with Mitochondria and Other Compartments

Ostensibly, relics of the ER's origin from endosymbiotic OMVs are still visible today. Prokaryotes synthesize their lipids directly at the plasma membrane, but not so in eukaryotes. Eukaryotic lipid synthesis – which is similar to bacterial membrane lipid synthesis and not to the archaeal lipid synthesis pathway – occurs mainly at the ER and involves considerable exchange with the mitochondria [33,34]. In traditional invagination models for the origin of the endomembrane

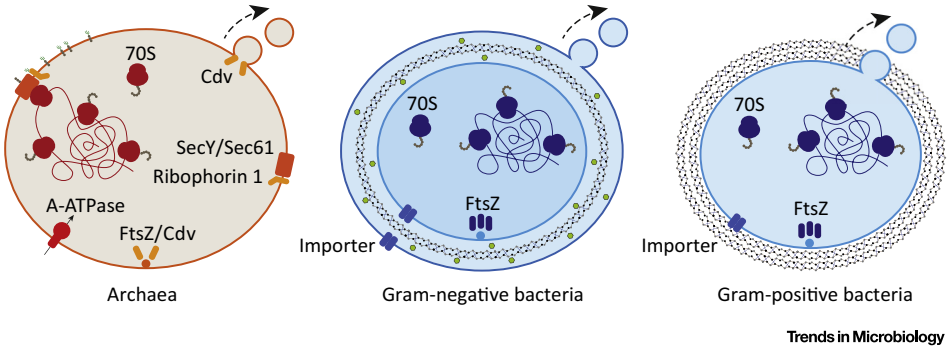


Figure 2. Prokaryotic Membrane Vesicle Secretion. Both bacteria and archaea release outer membrane vesicles (OMVs) into the environment that bud from the outer membrane. In archaea the Cdv proteins, which are involved in cell division and are homologous to proteins of the eukaryotic ESCRT machinery, mediate vesicle budding. All prokaryotes use 70S ribosomes for protein translation, and all bacteria, but only some archaea, make use of FtsZ for cell division. The illustration focuses on components discussed in the context of the proposal for the origin of the eukaryotic endomembrane system, such as the storage of Ca^{2+} (green hexagons) in the periplasmic space of Gram-negative bacteria or ribophorin 1, a protein involved in N-glycosylation.

system, the ER lumen is homologous to the environment. In our model, the ER lumen is homologous to the periplasm of the mitochondrial endosymbiont (Figure 3), the mitochondrial intermembrane space. The main sites of Ca^{2+} storage (mediated by **SERCA**) and signalling in eukaryotes are the ER and mitochondria [35], and indeed *Escherichia coli* concentrates Ca^{2+} between its inner and outer membrane under certain conditions [36]. The ER furthermore temporarily connects to other compartments and mitochondria through dedicated 10–30 nm long contact sites [37,38], from which autophagosomes arise and expand through MDV secretion [39]. Vesicle transport between mitochondria and peroxisomes, mediated by Vps35, has been recently observed [12], as well to MVBs [13]. Such contact sites are crucial for lipid biosynthesis, ion exchange and storage, signaling, and a range of membrane dynamics [34].

In addition, eukaryotic protein secretion commences at the rough ER through the binding of the 80S ribosome to the SEC complex that receives the nascent polypeptide chain to initiate cotranslational targeting into the ER lumen [40]. The entire eukaryotic SEC machinery is of archaeal origin [41,42], as is the ribosome [43,44]. The same is true for Rpn1 that stems from archaea [44]. By inference, the entire archaeal SecY/Sec61p/Rpn1 system for cotranslational N-glycosylation [45] was seamlessly integrated early into mitochondrial OMVs, which subsequently fused with the host's plasma membrane (Figure 3B). One might interject that the archaeal SecY/Sec61p system would be unlikely to integrate into a bacterial lipid bilayer, but the crystal structure of the archaeal Sec61 complex was determined from proteins expressed in *E. coli* [46].

N-glycosylation itself is not as unique to eukaryotes as once believed, being widespread among prokaryotes [47,48]. A continuous flow of bacterial lipid OMVs to the archaeal plasma membrane, also for the release of N-glycosylated proteins, would have naturally transformed the lipid composition of the archaeal plasma membrane from ether-linked isoprenes to ester-linked fatty acids (Figure 3B). Considering the diameter of bacterial OMVs (10–300 nm; [15]) and the diameter of an average archaeon (1 μm; [49]), less than 10 000 OMVs can generate enough membrane material to transform a surface area equal to that of an average archaeon. A clear implication of our proposal is that the lipid transition in eukaryotes could have occurred very rapidly in evolutionary terms, without precisely defining rapid in specific years or generations.

The secretion of vesicles to the plasma membrane is hardly the sole function of the ER. Vesicles that originate from the ER fuse with, and form, all other endomembrane-bounded cell

Key Figure

A Model for the Evolutionary Origin of the Eukaryotic Endomembrane System

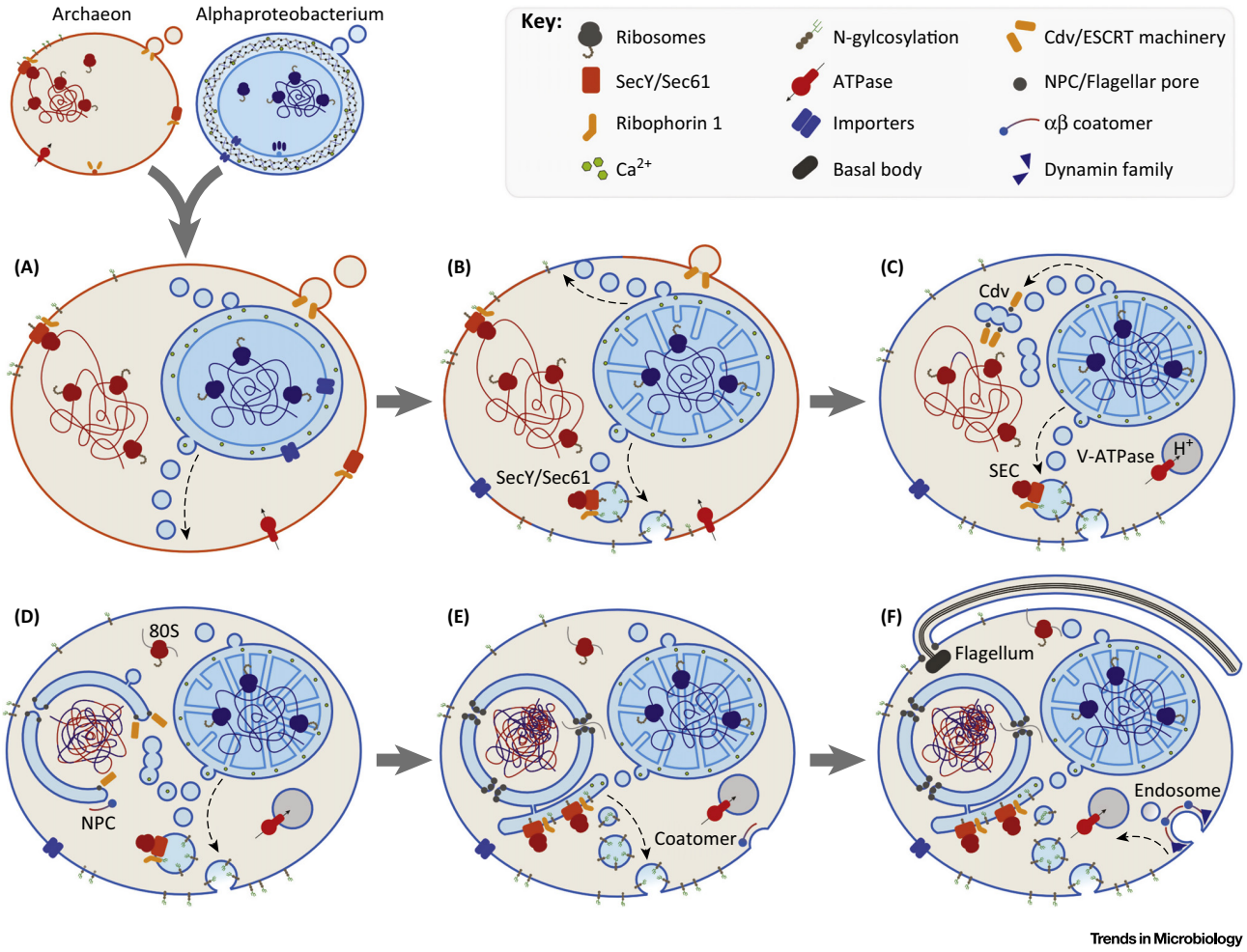


Figure 3. (A) After endosymbiosis, the endosymbiont continues with OMV secretion, thereby generating the first vesicular flux inside the archaeal cytosol. (B) Cytoplasmic OMVs provided an alternative target for the archaeal SecY/Sec61 and N-glycosylation machinery, generating a primitive endoplasmic reticulum (ER) and stabilizing OMV flux towards the plasma membrane (PM), where N-glycosylated proteins are released. This simultaneously initiated the conversion of the PM from archaeal ether-linked isoprenes (orange) to bacterial ester-linked fatty acids (blue). (C) The invasion of group II introns through endosymbiotic gene transfer drove the formation of the nucleus, may be with the aid of Cdv proteins. The presence of a primitive ER allowed the evolution of additional compartments, including a vacuole acidified by the archaeal A-ATPase (today the V-ATPase). (D) Together with the nucleus, the nuclear pore complex (NPC) was formed using proteins that may be originated from the fusion of genes encoding β-propeller folds and α-solenoid domains, ribosomes were now excluded from the nucleus. (E) A fully functional endomembrane system, in which the nuclear envelope (NE) is continuous with the ER, was in place, and coatomer proteins initiated the emergence of the first inbound vesicle budding (endocytosis). (F) Simultaneously, the flagellum evolves from proteins that originate from the NPC and/or the endosome. A virtue of this model is that the ancestral direction of membrane flux is a natural consequence of a fully formed pre-existing property of the mitochondrial endosymbiont that continues until today (Figure 1).

compartments, and these include the **Golgi apparatus**, the lysosome, the peroxisome, the **glyoxysome**, and the autophagosome [50]. Analyses of peroxisomal function and proteins, such as those for fatty acid β-oxidation, indicate that mitochondria predated peroxisomes in evolution [51,52], which is consistent with our model. In our proposal, the ER is the central hub upon which the biogenesis of all other nonendosymbiotic compartments depends, including the nucleus.

Box 2. The Energetic Price of Phagocytosis

If the archaeal host that acquired the mitochondrion were phagocytotic, it must have had all the parts required for phagocytosis, namely, a fully developed cytoskeleton, food vacuoles, and an endomembrane system. In other words, assuming a phagocytotic origin of mitochondria [3–6,65] means assuming that the host had evolved eukaryotic complexity without the participation of mitochondria. That rekindles the archezoa theory, which was rejected over a decade ago [81] because its predictions failed. Mitochondria are tied to eukaryote origin, hence to the origin of complexity. Why? The origin of eukaryote complexity required energy, mitochondrial energy. Eukaryotic cell complexity emerges from massive amounts of proteins that constitute and modulate the cytoskeleton and membrane flux in the eukaryotic cytosol. Protein synthesis is 75% of a cell's energy budget; mitochondria afforded eukaryotes internalized bioenergetic membranes that scale freely with increasing cell volume [57], and that covered the costs of that protein synthesis.

Prokaryotes synthesize ATP at the plasma membrane. Eukaryote origin witnessed the loss of all ATP synthesis at the plasma membrane and the transition to compartmentalized energy metabolism with glycolytic ATP synthesis in the cytosol and chemiosmotic ATP synthesis in mitochondria [28,57,70]. The bioenergetic transition at eukaryote origin was evolutionarily rapid. How so? The archaeal host's plasma membrane ATPase did not become a pseudogene. It remained under functional constraint, was targeted via the Sec pathway to a novel endomembrane, and reversed function to acidify food vacuoles at cytosolic ATP expense.

Phagocytosis first theories fail to account for the source of cytosolic ATP required to acidify food vacuoles. One might counter that fermentation was the source, but archaeal fermenters are chemiosmotic, using their ATPase at the plasma membrane [82]. Provided that carbon was supplied to the mitochondrion through importers integrated into the archaeal plasma membrane, only one key innovation was required to transform the mitochondrial endosymbiont into an ATP-exporting organelle – the ADP/ATP carrier in the mitochondrial inner membrane [83]. This and other independent lines of evidence [84,85] all suggest rapid eukaryote origin.

Phagocytosis demands a fully functional endomembrane system, in turn requiring proteins that facilitate vesicle flux and the membrane vesicles themselves. Key proteins of the endomembrane system have homologous domains in prokaryotes, they underwent duplication, diversification and functional specialization at eukaryote origin [1,2]. Their diversification required the intracellular vesicles that afforded these proteins their selectable functions. Mitochondria not only supplied the energy needed to evolve eukaryotic endomembrane proteins, they physically provided the vesicles and source for their natural selection. Similar to energetics, bacterial outer membrane vesicles (OMVs) place mitochondria before phagocytosis in the sequence of events at eukaryote origin.

On the Origin of the Nucleus

The ER is contiguous with the nuclear envelope (NE), which, similar to the ER, is a single folded membrane with two leaves (Figure 1). In eukaryotes with open mitosis, the NE arises from ER vesicles, which store proteins of the NE when the nucleus disintegrates during cell division [53,54]. In eukaryotes with closed mitosis, the NE increases in size during cell growth via lipids supplied via the ER [55]. In other words, the nuclear envelope can be viewed as a functionally distinct extension of the ER. As with the emergence of the eukaryotic SEC system, the archaeal host added crucial components to cytosolic membranes of endosymbiont origin (Figure 3). ESCRT III and the p97 AAA-ATPase control annular fusion of the newly forming NE [56], a process topologically resembling membrane fusion at the end of cytokinesis. Archaeal cell division and OMV secretion depends on CdvB and CdvC that are homologous to eukaryotic ESCRT III proteins and Vps4, respectively [24,26,27]. The formation of a primordial ER, and later an NE, could have involved archaeal precursors of ESCRT proteins and small GTPases of the archaeal host, such as those identified in **Lokiarchaea** [44]. Consistent with their function during cell division, we suggest that, during evolution, the NE emerged from the ER (Figure 3C), while the ER emerged from the mitochondrion. This order is independent of, but fully compatible with, bioenergetic considerations that identify the origin of the mitochondrial endosymbiont in an archaeal host (Box 1) as the rate-limiting event at eukaryote origin from which all other processes of eukaryogenesis unfold [57].

Some proteins of the **nuclear pore complex (NPC)**, the **flagellar pore complex (FPC)**, intraflagellar transport, and some coatomer proteins that form vesicle coats share similar structural properties (or even entire proteins) and have hence been suggested to be evolutionarily

linked [1,58–60]. Many of these proteins are characterised by a WD-domain containing β-propellers followed by an α-solenoid domain. This architecture allows for membrane interaction and bending, two important requirements for re-shaping membranes. Although WD-domains and solenoid-like proteins are widespread among prokaryotes [61,62], their domains are rarely organized as in eukaryotic cells [63], although the eukaryotic organization could result from simple gene fusion events [64]. Phylogenetic analysis suggests that these eukaryotic WD-domain proteins arose at roughly the same time [1]. In light of the present considerations, the nuclear envelope and pore came first, later followed by the emergence of the flagella and endocytosis, and not *vice versa* as according to recent suggestions [3,6,65].

Phagocytosis and Energetics

Many traditional models for the origin of the endomembrane system posit that phagocytosis arose prior to mitochondrial origin [3–5,65]. Phagocytosis requires a multiprotein machinery that forms the phagosomal cavity, which might have evolved multiple times independently during eukaryote evolution [66]. Scission of endosomes and phagosomes from the plasma membrane involve dynamins, a family of large GTPases [67]. Phylogenomic analysis of the dynamin segments suggests that the ancient version responsible for mitochondrial division was also the one mediating scission of early endomembrane vesicles [68]. Dynamin-related proteins such as DynA are common in bacteria, and the only three dynamin-like protein encoding genes found in archaea are of bacterial origin [69]. It appears that eukaryotic dynamins evolved from endosymbiotic dynamin-like proteins, speaking in favor of mitochondria proceeding phagocytosis.

Phagocytosis also requires a process that acidifies the food vacuole. Food vacuoles are useless if their contents cannot be broken down by proteases, which in eukaryotes are acid-activated [70]. Eukaryotic vacuole acidification requires, in turn, an important archaeal component, the **V-ATPase**, which consumes ATP to acidify vacuoles rather than synthesizing ATP from redox-generated ion gradients [71]. During eukaryote origin, the host's plasma membrane rotor–stator archaeal A-type ATPase (the eukaryotic version is called the vacuolar or V-type ATPase [72]) was re-targeted to a new compartment of OMV origin (now the lysosome), concomitant with a functional reversal of direction from ATP synthesis in archaea to ATP consumption in eukaryotes. This raises an interesting question seldom, if ever, asked in the context of eukaryote endomembrane origin: if phagocytosis preceded the mitochondrion, where was the cytosolic ATP coming from that allowed the V-ATPase to run backwards? And given that all the enzymes of chemiosmotic energy harnessing in eukaryotes stem from bacteria [69,73], why did the host lose all components of membrane bioenergetics other than the A-type ATPase? We suggest that the answer to both questions is mitochondria, for two reasons (Box 2).

Concluding Remarks

The present proposal for the origin of the ER and derived membranes differs in premises and substance from previous suggestions in the literature, in that it is based on outward vesicle flux, not inward. Our proposal requires almost no innovations, exceptional or unique evolutionary processes in either the mitochondrial ancestor or the archaeal host in order to bring forth a basic ER function with outward vesicle flux. Our proposal raises new questions (see Outstanding Questions), while directly accounting (i) for the archaeal ancestry, localisation and orientation of the secretory machinery that performs cotranslational insertion of proteins into eukaryotic membranes, (ii) for the circumstance that eukaryotes store Ca^{2+} in the ER lumen, which is, in this model, homologous to the ancestral mitochondrial periplasmic space, (iii) for the ancestral ground state and bacterial lipid composition of eukaryotic endomembranes, (iv) for the archaeal nature of eukaryotic ribosomes and N-glycosylation at the ER, (v) for the finding that eukaryotic lipid synthesis occurs predominantly at the ER and mitochondria, not at the plasma membrane, (vi) for the transitional mechanism that converted the composition of eukaryotic membranes from archaeal to bacterial lipids, (vii) for the formation of the nucleus from the ER during cell

Outstanding Questions

As new archaeal lineages become characterized that are, by the measure of ribosomal phylogeny, more closely related to the host that acquired the mitochondrion, will we find large, complex, eukaryote-like cells that never harboured mitochondria (like the archezoa theory once predicted), or will they be morphologically normal archaeal cells?

To what extent do mitochondria-derived vesicles exist among eukaryotic supergroups, and what functions do they perform?

Given technologies to insert OMV-producing bacteria into archaeal cells or suitable synthetic analogues, will it be possible to generate analogues of primitive endomembrane systems?

Could the fusion of prokaryotic β-propeller and α-solenoid domains generate protocoatomer proteins that facilitate positive membrane curvature?

Do mitochondria-derived vesicles contribute to the rebuilding of the ER and nucleus after mitosis in some eukaryotic lineages?

development, not vice versa, and (viii) for the archaeal ancestry, localisation, and orientation of the eukaryotic V-ATPase in food vacuoles. From our proposal, a natural evolutionary order in the origin of several key characters of eukaryotic cells unfolds in that, during eukaryogenesis, the ER represented the first autogenous (nonendosymbiotic) cell compartment, formed from OMVs secreted by the mitochondrion, subsequently giving rise to both the nuclear envelope and an ancestrally outward endomembrane flux.

Acknowledgments

The authors acknowledge financial support through the Deutsche Forschungsgemeinschaft to SBG (GO1825/4-1) and the European Research Council to WFM (grant 666053). SBG, SG, and WFM conceived, wrote, read, and approved the manuscript. We thank Debbie Maizels (Zoobotanica Scientific Illustration: www.scientific-art. com) for her outstanding help with generating Figure 1 and Ursula Goodenough for sending the right email at the right time.

References

1. Field, M.C. *et al.* (2011) On a bender-BARs, ESCRTs, COPs, and finally getting your coat. *J. Cell Biol.* 193, 963–972
2. Koumandou, V.L. *et al.* (2013) Molecular paleontology and complexity in the last eukaryotic common ancestor. *Crit. Rev. Biochem. Mol. Biol.* 48, 373–396
3. Jékely, G. (2008) *Eukaryotic Membranes and Cytoskeleton, Origins and Evolution,* Springer Science+Businness Media, LLC
4. de Duve, C. (2007) The origin of eukaryotes: a reappraisal. *Nat. Rev. Genet.* 8, 395–403
5. Cavalier-Smith, T. (2014) The neomuran revolution and phagotrophic origin of eukaryotes and cilia in the light of intracellular coevolution and a revised tree of life. *Cold Spring Harb. Perspect. Biol.* 6, a016006
6. Mast, F.D. *et al.* (2014) Evolutionary mechanisms for establishing eukaryotic cellular complexity. *Trends Cell Biol.* 24, 435–442
7. Baum, D.A. (2015) A comparison of autogenous theories for the origin of eukaryotic cells. *Am. J. Bot.* 102, 1954–1965
8. Martin, W.F. *et al.* (2015) Endosymbiotic theories for eukaryote origin. *Phil. Trans. Royal. Soc. B* 370, 20140330
9. Knox, K.W. *et al.* (1966) Relation between excreted lipopolysaccharide complexes and surface structures of a lysine-limited culture of *Escherichia coli*. *J. Bacteriol.* 92, 1206–1217
10. Hoekstra, D. *et al.* (1976) Release of outer membrane fragments from normally growing *Escherichia coli*. *Biochim. Biophys. Acta* 455, 889–899
11. Benaim, G. *et al.* (1990) Ca^{2+} transport in isolated mitochondrial vesicles from *Leishmania braziliensis* promastigotes. *Mol. Biochem. Parasitol.* 39, 61–68
12. Braschi, E. *et al.* (2010) Vps35 mediates vesicle transport between the mitochondria and peroxisomes. *Curr. Biol.* 20, 1310–1315
13. Soubannier, V. *et al.* (2012) A vesicular transport pathway shuttles cargo from mitochondria to lysosomes. *Curr. Biol.* 22, 135–141
14. Cook, K.L. *et al.* (2014) Mitochondria directly donate their membrane to form autophagosomes during a novel mechanism of parkin-associated mitophagy. *Cell Biosci.* 4, 1
15. Deatherage, B.L. and Cookson, B.T. (2012) Membrane vesicle release in bacteria, eukaryotes, and archaea: a conserved yet underappreciated aspect of microbial life. *Infect. Immun.* 80, 1948–1957
16. Schwechheimer, C. and Kuehn, M.J. (2015) Outer-membrane vesicles from Gram-negative bacteria: biogenesis and functions. *Nat. Rev. Micro.* 13, 605–619
17. Lee, E-Y. *et al.* (2009) Gram-positive bacteria produce membrane vesicles: proteomics-based characterization of *Staphylococcus aureus*-derived membrane vesicles. *Proteomics* 9, 5425–5436
18. Chun, A.L. (2014) Bacterial nanowires: an extended membrane. *Nat. Nanotechnol.* 9, 750
19. Haurat, M.F. *et al.* (2011) Selective sorting of cargo proteins into bacterial membrane vesicles. *J. Biol. Chem.* 286, 1269–1276
20. Kadurugamuwa, J.L. and Beveridge, T.J. (1997) Natural release of virulence factors in membrane vesicles by *Pseudomonas aeruginosa* and the effect of aminoglycoside antibiotics on their release. *J. Antimicrob. Chemother.* 40, 615–621
21. Kulp, A. and Kuehn, M.J. (2010) Biological functions and biogenesis of secreted bacterial outer membrane vesicles. *Annu. Rev. Microbiol.* 64, 163–184
22. Mashburn, L.M. and Whiteley, M. (2005) Membrane vesicles traffic signals and facilitate group activities in a prokaryote. *Nature* 437, 422–425
23. Grimm, R. *et al.* (1998) Electron tomography of ice-embedded prokaryotic cells. *Biophys. J.* 74, 1031–1042
24. Ellen, A.F. *et al.* (2010) Shaping the archaeal cell envelope. *Archaea* 2010, 1–13
25. Prangishvili, D. *et al.* (2000) Sulfolobicins, specific proteinaceous toxins produced by strains of the extremely thermophilic archaeal genus *Sulfolobus*. *J. Bacteriol.* 182, 2985–2988
26. Lindås, A-C. *et al.* (2008) A unique cell division machinery in the Archaea. *Proc. Natl. Acad. Sci. U.S.A.* 105, 18942–18946
27. Makarova, K.S. *et al.* (2010) Evolution of diverse cell division and vesicle formation systems in Archaea. *Nat. Rev. Microbiol.* 8, 731–741
28. Martin, W. and Muller, M. (1998) The hydrogen hypothesis for the first eukaryote. *Nature* 392, 37–41
29. McInerney, J.O. *et al.* (2014) The hybrid nature of the Eukaryota and a consilient view of life on Earth. *Nat. Rev. Microbiol.* 12, 449–455
30. French, S.L. *et al.* (2007) Transcription and translation are coupled in Archaea. *Mol. Biol. Evol.* 24, 893–895
31. Dohlen and von, C.D. *et al.* (2001) Mealybug beta-proteobacterial endosymbionts contain gamma-proteobacterial symbionts. *Nature* 412, 433–436
32. Husnik, F. *et al.* (2013) Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. *Cell* 153, 1567–1578
33. Tatsuta, T. *et al.* (2014) Mitochondrial lipid trafficking. *Trends Cell Biol.* 24, 44–52
34. Murley, A. and Nunnari, J. (2016) The emerging network of mitochondria-organelle contacts. *Mol. Cell.* 61, 648–653
35. Clapham, D.E. (2007) Calcium signaling. *Cell* 131, 1047–1058
36. Jones, H.E. *et al.* (2002) Direct measurement of free Ca^{2+} shows different regulation of Ca^{2+} between the periplasm and the cytosol of *Escherichia coli*. *Cell Calcium* 32, 183–192
37. Phillips, M.J. and Voeltz, G.K. (2015) Structure and function of ER membrane contact sites with other organelles. *Nat. Rev. Mol. Cell. Biol.* 17, 69–82
38. Lang, A. *et al.* (2015) ER–mitochondria contact sites in yeast: beyond the myths of ERMES. *Curr. Opin. Cell Biol.* 35, 7–12
39. Hailey, D.W. *et al.* (2010) Mitochondria supply membranes for autophagosome biogenesis during starvation. *Cell* 141, 656–667
40. Rapoport, T.A. (2007) Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. *Nature* 450, 663–669
41. Kinch, L.N. *et al.* (2002) Sec61β – a component of the archaeal protein secretory system. *Trends Biochem. Sci.* 27, 170–171

42. Cao, T.B. and Saier, M.H., Jr (2003) The general protein secretory pathway: phylogenetic analyses leading to evolutionary conclusions. *BBA – Biomembranes* 1609, 115–125
43. Raymann, K. *et al.* (2015) The two-domain tree of life is linked to a new root for the archaea. *Proc. Natl. Acad. Sci. U.S.A.* 112, 6670–6675
44. Spang, A. *et al.* (2015) Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* 521, 173–179
45. Ruiz-Canada, C. *et al.* (2009) Cotranslational and posttranslational N-glycosylation of polypeptides by distinct mammalian OST isoforms. *Cell* 136, 272–283
46. van den Berg, B. *et al.* (2004) X-ray structure of a protein-conducting channel. *Nature* 427, 36–44
47. Nothaft, H. and Szymanski, C.M. (2013) Bacterial protein N-glycosylation: new perspectives and applications. *J. Biol. Chem.* 288, 6912–6920
48. Jarrell, K.F. *et al.* (2014) N-linked glycosylation in archaea: a structural, functional, and genetic analysis. *Microbiol. Mol. Biol. Rev.* 78, 304–341
49. Portillo, M.C. *et al.* (2013) Cell size distributions of soil bacterial and archaeal taxa. *Appl. Environ. Microbiol.* 79, 7610–7617
50. Desjardins, M. (2003) ER-mediated phagocytosis: a new membrane for new functions. *Nat. Rev. Immunol.* 3, 280–291
51. Bolte, K. *et al.* (2015) The evolution of eukaryotic cells from the perspective of peroxisomes. *Bioessays* 37, 195–203
52. Speijer, D. (2010) Oxygen radicals shaping evolution: Why fatty acid catabolism leads to peroxisomes while neurons do without it. *Bioessays* 33, 88–94
53. Webster, M. *et al.* (2009) Sizing up the nucleus: nuclear shape, size and nuclear-envelope assembly. *J. Cell. Sci.* 122, 1477–1486
54. Lu, L. *et al.* (2011) Formation of the postmitotic nuclear envelope from extended ER cisternae precedes nuclear pore assembly. *J. Cell Biol.* 194, 425–440
55. Arnone, J.T. *et al.* (2013) The dynamic nature of the nuclear envelope: Lessons from closed mitosis. *Nucleus* 4, 261–266
56. Olmos, Y. *et al.* (2015) ESCRT-III controls nuclear envelope reformation. *Nature* 522, 236–239
57. Lane, N. and Martin, W. (2010) The energetics of genome complexity. *Nature* 467, 929–934
58. Devos, D. *et al.* (2004) Components of coated vesicles and nuclear pore complexes share a common molecular architecture. *PLoS Biol.* 2, 2085–2093
59. Dishinger, J.F. *et al.* (2010) Ciliary entry of the kinesin-2 motor KIF17 is regulated by importin-β2 and RanGTP. *Nat. Cell Biol.* 12, 703–710
60. Kee, H.L. *et al.* (2012) A size-exclusion permeability barrier and nucleoporins characterize a ciliary pore complex that regulates transport into cilia. *Nat. Cell Biol.* 14, 431–437
61. Chaudhuri, I. *et al.* (2008) Evolution of the β-propeller fold. *Proteins* 71, 795–803
62. Fournier, D. *et al.* (2013) Functional and genomic analyses of alpha-solenoid proteins. *PLoS ONE* 8, e79894
63. Jékely, G. and Arendt, D. (2006) Evolution of intraflagellar transport from coated vesicles and autogenous origin of the eukaryotic cilium. *Bioessays* 28, 191–198
64. Enright, A.J. *et al.* (1999) Protein interaction maps for complete genomes based on gene fusion events. *Nature* 402, 86–90
65. Koonin, E.V. (2015) Origin of eukaryotes from within archaea, archaeal eukaryome and bursts of gene gain: eukaryogenesis just made easier? *Phil. Trans. Royal Soc. B* 370, 20140333
66. Yutin, N. and Koonin, E.V. (2012) Archaeal origin of tubulin. *Biol. Direct.* 7, 10
67. Huynh, K.K. and Grinstein, S. (2008) Phagocytosis: dynamin's dual role in phagosome biogenesis. *Curr. Biol.* 18, R563–R565
68. Purkanti, R. and Thattai, M. (2015) Ancient dynamin segments capture early stages of host-mitochondrial integration. *Proc. Natl. Acad. Sci. U.S.A.* 112, 2800–2805
69. Ku, C. *et al.* (2015) Endosymbiotic origin and differential loss of eukaryotic genes. *Nature* 524, 427–432
70. Braulke, T. and Bonifacino, J.S. (2009) Sorting of lysosomal proteins. *BBA – Mol. Cell. Res.* 1793, 605–614
71. Mindell, J.A. (2012) Lysosomal acidification mechanisms. *Annu. Rev. Physiol.* 74, 69–86
72. Grüber, G. and Marshansky, V. (2008) New insights into structure–function relationships between archeal ATP synthase (A_1A_0) and vacuolar type ATPase (V_1V_0). *Bioessays* 30, 1096–1109
73. Müller, M. *et al.* (2012) Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. *Microbiol. Mol. Biol. Rev.* 76, 444–495
74. Cox, C.J. *et al.* (2008) The archaebacterial origin of eukaryotes. *Proc. Natl. Acad. Sci. U.S.A.* 105, 20356–20361
75. Kelly, S. *et al.* (2011) Archaeal phylogenomics provides evidence in support of a methanogenic origin of the archaea and a thaumarchaeal origin for the eukaryotes. *Proc. R. Soc. B* 278, 1009–1018
76. Williams, T.A. *et al.* (2013) An archaeal origin of eukaryotes supports only two primary domains of life. *Nature* 504, 231–236
77. Martin, W. and Koonin, E.V. (2006) Introns and the origin of nucleus–cytosol compartmentalization. *Nature* 440, 41–45
78. Voeltz, G.K. *et al.* (2002) Structural organization of the endoplasmic reticulum. *EMBO Rep.* 3, 944–950
79. Wilson, C.M. and High, S. (2007) Ribophorin I acts as a substrate-specific facilitator of N-glycosylation. *J. Cell. Sci.* 120, 648–657
80. Schmidt, O. and Teis, D. (2012) The ESCRT machinery. *Curr. Biol.* 22, R116–R120
81. Embley, T.M. and Martin, W. (2006) Eukaryotic evolution, changes and challenges. *Nature* 440, 623–630
82. Schönheit, P. *et al.* (2016) On the origin of heterotrophy. *Trends Microbiol.* 24, 12–25
83. Radzvilavicius, A.L. and Blackstone, N.W. (2015) Conflict and cooperation in eukaryogenesis: implications for the timing of endosymbiosis and the evolution of sex. *J. R. Soc. Interface* 12, 20150584
84. Dacks, J.B. and Field, M.C. (2007) Evolution of the eukaryotic membrane-trafficking system: origin, tempo and mode. *J. Cell. Sci.* 120, 2977–2985
85. Speijer, D. (2015) Birth of the eukaryotes by a set of reactive innovations: New insights force us to relinquish gradual models. *Bioessays* 37, 1268–1276

Publication III

Conservation of transit peptide-independent protein import into the mitochondrial and hydrogenosomal matrix

Sriram Garg[1], Verena Zimorski[1], Jan Stölting[1], Petr Rada[2], Jan Tachezy[2], William F. Martin[1] and Sven B. Gould[1]

1 - Institute for Molecular Evolution, Heinrich-Heine Universität Düsseldorf, Universitätsstraße 1, Düsseldorf 40225, Germany

2 - Department of Parasitology, Charles University in Prague, Faculty of Science, Prague, Czech Republic.

The presented manuscript was peer-reviewed and published in **Genome Biology and Evolution** with impact factor of **4.257** on 2nd September 2015

Contribution as **first author**

80%

Designed and conducted experiments leading to Figure 1B,2,3,4 and Table 1. Designed and illustrated Figure 5. Wrote majority of the text and mined the literature for relevant citations.

Genome Biology and Evolution

SMBE

Conservation of Transit Peptide-Independent Protein Import into the Mitochondrial and Hydrogenosomal Matrix

Sriram Garg[1], Jan Stölting[1], Verena Zimorski[1], Petr Rada[2], Jan Tachezy[2], William F. Martin[1], and Sven B. Gould[1,*]

[1]Institute for Molecular Evolution, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

[2]Department of Parasitology, Charles University in Prague, Faculty of Science, Czech Republic

*Corresponding author: E-mail: gould@hhu.de.

Accepted: August 26, 2015

Abstract

The origin of protein import was a key step in the endosymbiotic acquisition of mitochondria. Though the main translocon of the mitochondrial outer membrane, TOM40, is ubiquitous among organelles of mitochondrial ancestry, the transit peptides, or N-terminal targeting sequences (NTSs), recognised by the TOM complex, are not. To better understand the nature of evolutionary conservation in mitochondrial protein import, we investigated the targeting behavior of *Trichomonas vaginalis* hydrogenosomal proteins in *Saccharomyces cerevisiae* and vice versa. Hydrogenosomes import yeast mitochondrial proteins even in the absence of their native NTSs, but do not import yeast cytosolic proteins. Conversely, yeast mitochondria import hydrogenosomal proteins with and without their short NTSs. Conservation of an NTS-independent mitochondrial import route from excavates to opisthokonts indicates its presence in the eukaryote common ancestor. Mitochondrial protein import is known to entail electrophoresis of positively charged NTSs across the electrochemical gradient of the inner mitochondrial membrane. Our present findings indicate that mitochondrial transit peptides, which readily arise from random sequences, were initially selected as a signal for charge-dependent protein targeting specifically to the mitochondrial matrix. Evolutionary loss of the electron transport chain in hydrogenosomes and mitosomes lifted the selective constraints that maintain positive charge in NTSs, allowing first the NTS charge, and subsequently the NTS itself, to be lost. This resulted in NTS-independent matrix targeting, which is conserved across the evolutionary divide separating trichomonads and yeast, and which we propose is the ancestral state of mitochondrial protein import.

Key words: mitochondria, hydrogenosomes, mitosomes, protein import, TOM/TIM.

Introduction

The origin of mitochondria marked the emergence of eukaryotes (Williams et al. 2013; McInerney et al. 2014), whose increased cellular complexity over prokaryotes is founded in the compartmentalization of chemiosmotic ATP (adenosine triphosphate) synthesis in the organelle (Martin and Koonin 2006; Lane and Martin 2010). All known eukaryotes possess, or possessed in their past, mitochondria or organelles derived thereof—hydrogenosomes and mitosomes (Van der Giezen et al. 2002). The family of mitochondrial organelles underwent different trajectories of specialization in different eukaryotic lineages. Aerobic mitochondria use O_2 as the terminal electron acceptor, anaerobic mitochondria use other terminal acceptors such as fumarate. Hydrogenosomes generate ATP via H_2-producing fermentations, while mitosomes consume ATP, rather than generating it (Muller et al. 2012). Hydrogenosomes are evolutionarily reduced in that they have lost the respiratory chain and the electrochemical gradient ($\Delta\psi$) and instead generate ATP through substrate-level phosphorylation only (Lindmark and Muller 1973). Mitosomes are the most highly reduced forms of mitochondria, their only known functions involving Fe–S cluster assembly (Lill and Neupert 1996; Goldberg et al. 2008) and sulfur metabolism (Mi-ichi et al. 2009). Despite this specialization, mitochondrial protein import is conserved. Mitosomes of *Encephalitozoon cuniculi* might import only as few as 22 proteins (Katinka et al. 2001; Waller et al. 2009), yet like any other eukaryote studied so far, they depend on a mitochondrial translocon machinery consisting of components conserved in the canonical TIM and TOM complexes (Translocase of the Outer/Inner Mitochondrial membrane) of yeast and human mitochondria to do so (Doležal et al. 2006; Neupert

and Herrmann 2007; Chacinska et al. 2009; Endo and Yamano 2009; Schleiff and Becker 2010).

Early in mitochondrial evolution, the invention of a protein import machinery allowed the organelle to relinquish genes to the nucleus (Timmis et al. 2004), but in order for the organelle to maintain its biochemical identity, and hence fulfill its bioenergetic functions, a mechanism that selectively discriminated between proteins germane to the organelle and pre-existing host proteins in the cytosol must have been in place. Today, this discrimination is provided by the TOM and TIM complexes, which comprise the core of the mitochondrial protein import machinery (Doležal et al. 2006; Chacinska et al. 2009; Schleiff and Becker 2010; Neupert 2015). In the oxygen-respiring mitochondria of yeast and humans, hundreds of matrix proteins enter the organelle via the TOM receptor platform that interacts with mitochondrial N-terminal targeting sequences (mNTSs) (Neupert and Herrmann 2007; Chacinska et al. 2009; Schleiff and Becker 2010). Anaerobic organelles of mitochondrial origin, hydrogenosomes and mitosomes, import fewer proteins than classical mitochondria but still make use of the same core components of the TOM and TIM machinery (Waller et al. 2009).

The main TOM component, Tom40, shuttles the unfolded preproteins into the inner membrane space, where they are received by the TIM23 complex that translocates proteins into the matrix in a process that in yeast requires both ATP and an electrochemical gradient ($\Delta\psi$) across the inner membrane (Martin et al. 1991). Proteins targeted to the mitochondrial matrix harbor N-terminal targeting sequences (mNTSs) that can readily arise from random sequences (Baker and Schatz 1987) and that are present naturally in bacterial genomes (Lucattini et al. 2004). Although the translocases of the mitochondrial outer and inner membranes are ubiquitous among organelles of mitochondrial ancestry, positively charged NTSs that direct proteins to the organellar matrix are not (Regoes et al. 2005; Goldberg et al. 2008; Šmíd et al. 2008; Waller et al. 2009; Zimorski et al. 2013).

The two membranes that surround hydrogenosomes harbor many homologs of the TOM/TIM machinery. Proteins present include TOM40, TIM23, and proteins of the SAM and PAM complex, but they appear to lack many of the peripheral components of the mitochondrial targeting machinery as proteomic profiling has shown (Rada et al. 2011). *Trichomonas* hydrogenosomes lack a genome and therefore import all of the 200–500 proteins that exist in the organelle from the cytosol (Burstein et al. 2012). The *Trichomonas* genome encodes 226 proteins that harbor a short N-terminal motif with conserved features thought to represent the hydrogenosomal N-terminal targeting sequence or hNTS (Carlton et al. 2007; Burstein et al. 2012). This hNTS, while short, has been shown in some cases to be sufficient to target marker proteins to mitochondria of yeast (Häusler et al. 1997). Surprisingly, though, the deletion of the hNTS had only a marginal, if any, impact on the targeting efficiency of at least eight *Trichomonas* matrix proteins to hydrogenosomes (Mentel et al. 2008; Burstein et al. 2012; Zimorski et al. 2013), raising questions about the role and essentiality of the hNTS in hydrogenosomal protein import. To investigate the extent to which N-terminal independent targeting is conserved across the evolutionary divide that separates excavates and opisthokonts, we analyzed the targeting behavior of *Trichomonas* hydrogenosomal proteins in yeast and, reciprocally, the targeting of yeast mitochondrial proteins in *Trichomonas* with and without their NTSs.

Materials and Methods

Cultivation and Cloning

Trichomonas vaginalis T1 (J-H Tai, Institute of Biomedical Sciences, Taipei, Taiwan) was cultivated in TYM (Tryptone-Yeast extract-Maltose) medium at 37 °C as described previously (Gorrell et al. 1984). *Saccharomyces cerevisiae* INVSc1 was obtained from Invitrogen (Cat.No C810-00) and cultivated in YPD (Yeast extract-Peptone-Dextrose) (2% [w/v] glucose, 1% [w/v] yeast extract, and 2% [w/v] peptone) at 30 °C. Transfected yeast strains were cultivated in SC (Synthetic Complete) minimal medium (0.67% [w/v] yeast nitrogen base, 0.96% (w/v) yeast synthetic dropout medium without uracil) supplemented with 1% (w/v) raffinose. Open reading frames of the genes were cloned using genomic DNA of *T. vaginalis* T1 or genomic DNA of INVSc1 as template using gene-specific primers containing appropriate sites for the respective restriction enzymes as listed in supplementary table S2, Supplementary Material online. pTagVag2 was used for expression of genes in *T. vaginalis* which contains a C-terminal di-hemagglutinin (HA) tag and the promoter of the *Tv*SCSα1 gene. For expression in *S. cerevisiae*, an inducible expression vector pYES2/eGFP (pYES2/CT with a C-terminal eGFP) was used whereby fusion constructs could be induced by the addition of 4% (w/v) galactose. *Trichomonas vaginalis* T1 cells were transfected as described before (Land et al. 2003) with 50 μg of plasmid and selected with 100 μg/ml G418. Transformation of *S. cerevisiae* cells was carried out using the protocol described in the manufacturer's manual.

Cell Fractionation and Organelle Isolation

Isolations of hydrogenosomes were performed exactly as described before in Zimorski et al. (2013) except for an additional isopycnic centrifugation in 45% (v/v) Percoll density gradient with two intermediary washing steps to remove contaminating fractions. The isolation of mitochondria was carried out according to the protocol detailed in Gregg et al. (2009) with transfected cells grown in SC minimal medium. The total lysate fraction was collected immediately after the homogenization of cells. The supernatant of the pelleted mitochondria represents the cytoplasmic fraction.

Western Blotting, Immunofluorescence, and Imaging

Protein samples were separated through standard SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) procedures and blotted onto a nitrocellulose membrane. The membranes were blocked in 5% (w/v) dried milk powder in Tris-buffered saline with 0.1% (v/v) Tween20 (TBS-T) (tris-buffered-saline-tween20) for 30 min. The blots were incubated with primary antibodies at a concentration of 1:1,000 in blocking buffer overnight at 4 °C or 60 min at RT (room temperature) and then washed with TBS-T followed by incubation with secondary antibodies at a concentration of 1:5,000 in blocking buffer with 1% (w/v) dried milk powder for 60 min at RT and subsequent washes before imaging the blots directly in a Bio-Rad ChemiDoc™ XRS system.

For immunofluorescent labeling cells fixed in 1% (v/v) paraformaldehyde were deposited on cover slides coated with 0.01% polylysine and permeabilized for 15 min in 0.5% (v/v) Triton-X100. Permeabilized cells were blocked using a blocking buffer containing 1% (w/v) bovine serum albumin for 60 min followed by incubation in mouse anti-HA monoclonal antibody (Sigma) and rabbit anti-SCSα polyclonal serum in blocking buffer at a 1:1,000 dilution overnight at 4 °C. The cells were then washed three times in PBS before incubation with donkey antimouse Alexa 488 and donkey antirabbit Alexa 594 antibodies at a 1:5,000 dilution in blocking buffer for 60 min at RT. After final washes cells were mounted using Fluroshield™ containing DAPI (4′,6-diamidino-2-phenylindole) (Sigma) and observed using the Zeiss LSM710 confocal microscopy system.

Induction of eGFP (enhanced green fluroscent protein) fusion construct expressing yeast cells was carried out by growing log-phase yeast transformants in the presence of 4% (v/v) galactose for 4 h followed by incubation with 1 nM MitoTracker Red CMXRos (Invitrogen) and then mounted on silane coated slides in a solution of 1.2% (w/v) agarose to immobilize the cells and visualized in the Zeiss LSM710 confocal microscope. All images were analyzed using ImageJ software (Pérez and Pascau 2013).

Analysis of Hydrogenosomal Proteins by Liquid Chromatography-Electrospray Ionization MS/MS

Samples were digested and analyzed using liquid chromatography (LC)-electrospray ionization (ESI) mass spectrometry. Protein lysates (5 µg) were focused on a 4–12% polyacrylamide bis-tris gel (Life Technologies). After silver staining, protein bands were cut, destained (15 mM $Na_2S_2O_3$, 50 mM $K_3[Fe(CN)_6]$), reduced (10 mM DTT (dithiothreitol), 50 mM $(NH_4)HCO_3$), alkylated (50 mM C_2H_4INO, 50 mM NH_4HCO_3), and proteins were digested overnight in 50 mM NH_4HCO_3, with 0.1 µg trypsin (Serva) or 0.1 µg GluC (Promega). Alternatively, digestion with 0.1 µg ArgC (Promega) was carried out in ArgC digestion buffer (50 mM Tris [pH 7.6], 2 mM ethylenediaminetetraacetic acid, 5 mM DTT, and 4.5 mM $CaCl_2$). For LC-MS/MS (mass spectrometry) analyses, peptides were extracted from the gel with 1:1 (v/v) 0.1% TFA (trifluroacetic acid)/acetonitrile and after removal of acetonitrile 500 ng peptides were subjected to LC.

An Ultimate 3000 Rapid Separation liquid chromatography system (Dionex/Thermo Scientific) was used for peptide separation. After injection, peptides were preconcentrated on an Acclaim PepMap100 trap column (3 µm C18 particle size, 100 Å pore size, 75 µm inner diameter, 2 cm length; Dionex/Thermo Scientific) at a flow rate of 6 µl/min using 0.1% (v/v) TFA as mobile phase. After 10 min, peptides were separated on an analytical column (Acclaim PepMapRSLC, 2 µm C18 particle size, 100 Å pore size, 75 µm inner diameter, 25 cm length; Dionex/Thermo Scientific) at 60 °C using a 2-h gradient from 4% to 40% solvent B (solvent A: 0.1% (v/v) formic acid in water′ solvent B: 0.1% (v/v) formic acid, 84% (v/v) acetonitrile in water) at a flow rate of 300 nl/min.

Mass spectrometry was carried out on an Orbitrap Elite high resolution instrument (Thermo Scientific) operated in positive mode and equipped with a nano ESI source. Capillary temperature was set to 275 °C and source voltage to 1.4 kV. Survey scans were carried out in the orbitrap analyzer over a mass range from 350 to 1,700 m/z at a resolution of 60,000 (at 400 m/z). The target value for the automatic gain control was 1,000,000 and the maximum fill time was 200 ms. The 20 most intense 2^+ and 3^+ charged peptide ions (minimal signal intensity 500) were isolated, transferred to the linear ion trap (LTQ [linear trap quadrupole]) part of the instrument, and fragmented using collision induced dissociation. Peptide fragments were analyzed with a maximal fill time of 300 ms and automatic gain control target value of 10,000. The available mass range was 200–2,000 m/z at a resolution of 5,400 (at 400 m/z). Already fragmented ions were excluded from fragmentation for 45 s.

Analysis of Mass Spectrometric Data

Raw files were processed with MaxQuant (version 1.4.1.2, Max Planck Institute for Biochemistry, Munich, Germany) for protein and peptide identification and quantification with default parameters if not otherwise stated. Searches were carried out using *T. vaginalis* protein sequences from TrichDB (release 1.3 from 26.5.2011 including 59,672 protein entries; Aurrecoechea et al. 2008) applying the following parameters: Mass tolerance precursor (Orbitrap): 20 ppm first search, 4.5 ppm second search; mass tolerance fragment spectra (linear ion trap): 0.5 Da (linear ion trap); fixed modification: Carbamidomethyl (C), nicotin (K); variable modifications: Mthionine oxidation. Searches with protease-specific cleavage (depending on the used enzyme GluC, ArgC, maximum of two missed cleavage sites) were used with specific cleavage and in an alternative setting with N-terminal semispecific cleavage.

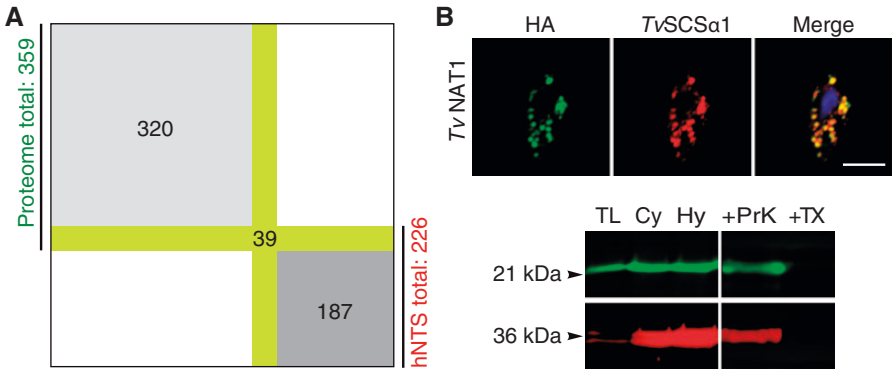


Fig. 1.—The majority of hydrogenosomal proteins lack an hNTS. (*A*) Of the 226 proteins predicted to have an hNTS, only 39 were found in the hydrogenosomal proteome of 359 proteins identified in total. (*B*) Immunofluorescent colocalization of the HA-tagged TVAG_270750 that harbors no hNTS and the hydrogenosomal marker enzyme *Tv*SCSα (succinyl coenzyme A synthetase). Below a multiplex western blot of the protease protection assay on isolated hydrogenosomes with green representing anti-HA and red anti-*Tv*SCSα. TL, total lysate; Cy, cytosol; Hy, hydrogenosomes; TX, 0.1% Triton X-100; ProtK, Triton X-100 + 100 μg/μl proteinase K. Numbers to the left indicate the approximate molecular weights of the constructs in kilo Daltons (kDa).

For peptide and protein acceptance, the false discovery rate (FDR) was set to 1%, and only proteins with at least two identified peptides were used for protein assembly. Quantification was carried out using the label-free quantification algorithm implemented in MaxQuant using a minimal ratio count of 2 and the "match between runs" option enabled. Alternatively, raw files were further processed for protein and peptide identification using Proteome Discoverer (version 1.4.1.14, Thermo Scientific) connected to a Mascot server (version 2.4.1, Matrix Sciences, London, UK) with default parameters for spectrum selection. Searches were carried out using 59,672 protein entries and protein sequences from TrichDB (release 1.3 from 26.5.2011) applying the following parameters: Mass tolerance precursor (analyzed in the Orbitrap part of the instrument) 10 ppm, mass tolerance fragment spectra (analyzed in the linear ion trap) 0.4 Da, enzyme-specific cleavage with a maximum of one missed cleavage site and N- and C-terminal semispecific cleavage specificity, carbamidomethyl at cysteines and nicotine at lysines as fixed modification and methionine oxidation. For peptide and protein acceptance, the "Percolator" function with a Target FDR set to 1% and validation based on *q*-value was used. Only peptides with high confidence (FDR < 1%) were used for protein assembly. Protein grouping was enabled. Net charge of peptides was analyzed using the EMBOSS package pepstats (Rice et al. 2000).

Results

The Majority of Hydrogenosomal Proteins Do Not Harbor an N-Terminal Targeting Sequence

A previous proteomic investigation of isolated hydrogenosomes from *T. vaginalis* identified 536 proteins, including 99 proteins for which only one peptide was identified (Schneider et al. 2011). Hydrogenosomes thus harbor on the order of 4–500 proteins, which is about half as many proteins as are predicted to localize to yeast mitochondria (Meisinger et al. 2008), but about twice the number of *Trichomonas* proteins (226) predicted by Burstein et al. (2012) to contain an hydrogenosomal NTS. A subsequent proteomic study of only the membrane associated proteins (Rada et al. 2011) revealed another 102 proteins that were not identified by Schneider et al. (2011). This prompted us to reinvestigate the *Trichomonas* hydrogenosomal proteome. Using biological triplicates of highly purified isolated hydrogenosomes and proteolytic digestion of the isolated proteins by two independent proteases, ArgC and GluC, for analysis by mass spectrometry (supplementary table S1, Supplementary Material online), 359 proteins were common to the three separate LC-MS/MS runs and detected with at least two peptides per protein. Of these 359 proteins, only 39 proteins have an hNTS based on previous predictions (Burstein et al. 2012; fig. 1*A*). The rest of the 320 proteins lacked a predictable hNTS altogether, including proteins that were previously not identified (supplementary table S2, Supplementary Material online).

One of those proteins lacking an hNTS was TVAG_270750 (*Tv*NAT1), which shares ~50% amino acid identity with bacterial acetyltransferases from the GNAT family. To confirm that *Tv*NAT1 is a hydrogenosomal protein in vivo, we expressed *Tv*NAT1 as an HA-tagged construct. It colocalized with *Tv*SCSα1 (fig. 1*B*), a marker enzyme of the hydrogenosomal matrix (Zimorski et al. 2013). Matrix localization was further supported by a protease protection assay (PPA) on isolated hydrogenosomes (fig. 1*B*). *Tv*NAT1 is thus yet one more in a growing list of proteins that localize to the *Trichomonas* hydrogenosomal matrix in the absence of an NTS (Mentel et al. 2008; Burstein et al. 2012; Zimorski et al. 2013). This prompted us to undertake a broader and more systematic investigation of NTS-independent targeting to hydrogenosomes.

Reciprocal Targeting of Mitochondrial and Hydrogenosomal Proteins with and without NTSs

We investigated the targeting behavior of *T. vaginalis* hydrogenosomal proteins in *S. cerevisiae* with and without their hNTSs. Four different hydrogenosomal matrix proteins of *T. vaginalis* (*Tv*SCSα1, *Tv*Fdx, *Tv*ME, *Tv*ISCA1), whose hNTS had been previously shown to be nonessential in *Trichomonas* (Zimorski et al. 2013), were fused to the N-terminus of eGFP and localized in yeast. All four proteins of the parasite carrying their hNTS were targeted to yeast mitochondria (fig. 2*A*–*D*). These four proteins were also targeted to the mitochondria of *S. cerevisiae* when the proteins were expressed without their NTS (fig. 2*A*–*D*). As further support, mitochondria of the transformed strains were isolated and the subcellular fractions investigated in multiplex western blots using an anti-eGFP antibody and an antibody against CoxIV, a protein of the inner mitochondrial membrane (fig. 2*A*–*E*). The western blots confirmed the localization observed by immunofluorescent microscopy, that is, the fusion proteins were exclusively detected in the fractions containing either total protein or the proteins of the isolated yeast mitochondria, and that no matter of whether the hNTS was present or not.

We also tested the reciprocal case. To determine if the converse was true for yeast mitochondrial proteins, we expressed four canonical and abundant yeast mitochondrial proteins (*Sc*LSC1, *Sc*COXIV, *Sc*IDH1, and *Sc*KGD2) in the parasite *T. vaginalis*. *Sc*LSC1 is the yeast homolog of *Tv*SCSα1, *Sc*COXIV is part of the mitochondrial respiratory chain, and *Sc*IDH1 and *Sc*KGD2 are metabolic enzymes of the mitochondrial matrix for which *Trichomonas* encodes no homologs. An earlier study of the N-proteome of the mitochondria of yeast identified all four proteins to be present in the mitochondria (Vögtle et al. 2009). Moreover, these proteins were present in a processed form having their mNTS cleaved by a peptidase, which strongly suggests that their mNTS is necessary for mitochondrial import. All four yeast proteins localized to *T. vaginalis* hydrogenosomes independent of the presence or absence of their mNTS (fig. 3*A*–*D*). In addition, multiplex western blots of purified hydrogenosomes and subsequent PPAs were also performed and demonstrated the proteins to be present in the organellar fractions (fig. 3*A*–*D*).

This targeting and localization is restricted to proteins of organellar origin in both organisms. As controls for yeast, we localized eGFP alone and additionally fused the eGFP to the C-terminus of an actin gene of *T. vaginalis* (*Tv*Actin). Both constructs remained in the cytosol and did not colocalize with MitoTracker® Red (fig. 4*A*). Three cytosolic yeast proteins (the glycolytic enzymes *Sc*GAPDH, *Sc*Actin, and *Sc*Rab5) were expressed in the parasite as a control using the same expression vector that is pTagVag2. The fusion proteins did not associate with the hydrogenosomes of *Trichomonas* (fig. 4*B*). This further demonstrates that the recognition of import substrate at the hydrogenosomal and mitochondrial outer membrane is specific, even in the absence of an NTS. The data indicate that some internal targeting information must exist in these yeast mitochondrial and *Trichomonas* hydrogenosomal proteins recognized by both the yeast and *Trichomonas* organelle protein import machinery that can discriminate between cytosolic proteins and proteins of the organelle. The nature of that targeting information remains obscure.

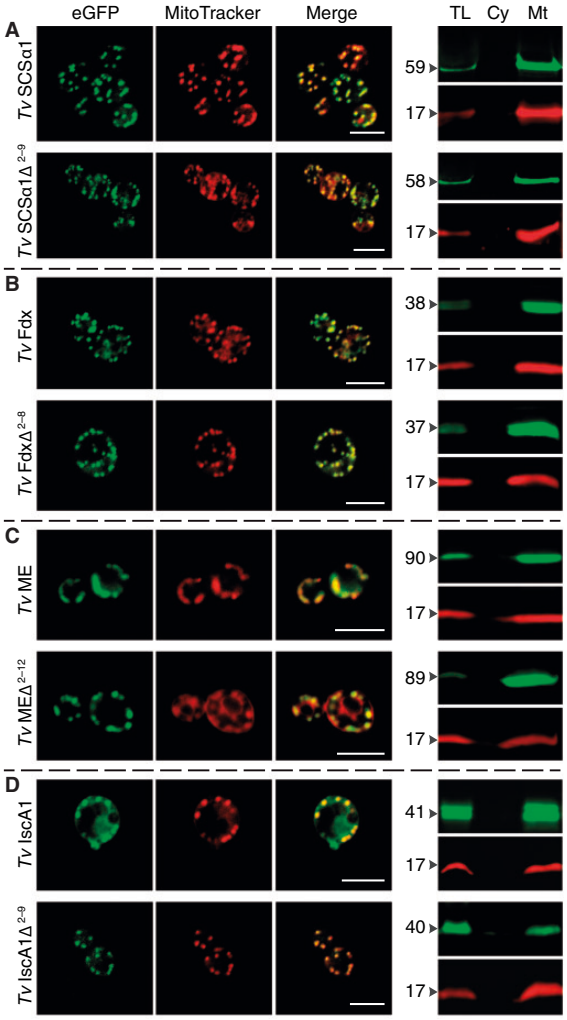


Fig. 2.—Hydrogenosomal proteins are targeted to yeast mitochondria with and without N-terminal leaders. (*A*–*D*) All four hydrogenosomal proteins of *Trichomonas vaginalis* analyzed (*Tv*SCSα1, succinyl coenzyme A synthetase; *Tv*Fdx, ferredoxin; *Tv*ME, malic enzyme; *Tv*ISCA1; iron–sulfur assembly protein 1) are targeted to the mitochondria of yeast, even in the absence of their NTSs. Δ indicates the positions of the N-terminal amino acids (i.e., the hNTS) that were deleted. (*E*) *Trichomonas* actin was used as a control next to the transfection of the empty vector that expresses GFP alone (supplementary fig. S2, Supplementary Material online). Scale indicates 5 μm. On the right, the multiplex western blots of isolated mitochondria probed with anti-GFP antibody (green) and the mitochondrial marker protein anti-COXIV (red). TL, total lysate; Cy, cytosol; Mt, mitochondria. Numbers to the left indicate the approximate molecular weights of the constructs in kilo Daltons (kDa).

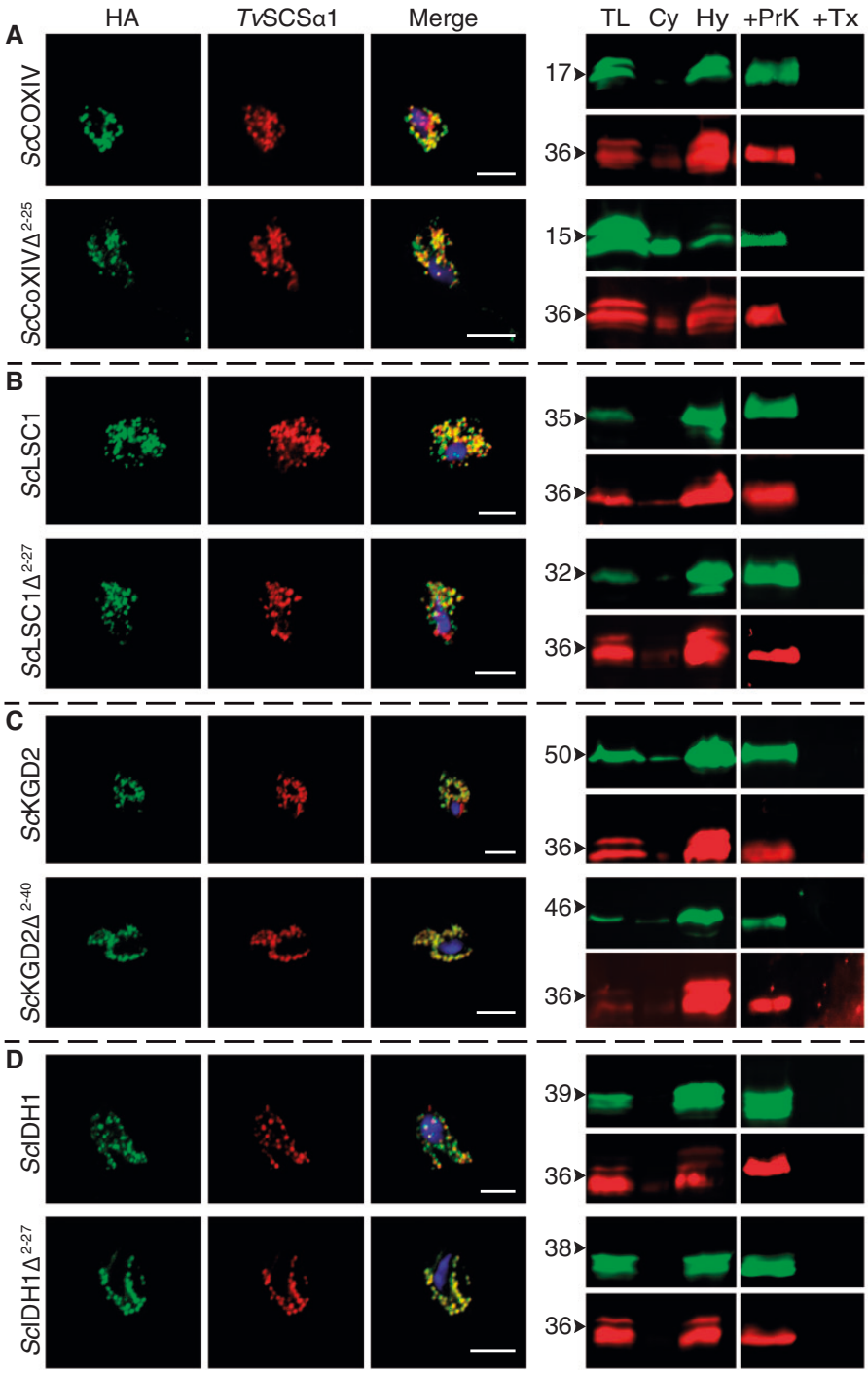


Fig. 3.—Yeast mitochondrial proteins are targeted to *Trichomonas vaginalis* hydrogenosomes with and without N-terminal leaders. (*A–D*) Expression of four mitochondrial proteins of *Saccharomyces cerevisiae* (*Sc*LSC1, succinate coenzyme A ligase; *Sc*COXIV, cytochrome C oxidase; *Sc*IDH1, isocitrate dehydrogenase; *Sc*KGD2, alpha-ketoglutarate dehydrogenase) tagged with HA demonstrate that they are all targeted to hydrogenosomes in the presence and absence of their mNTS. Δ indicates the positions of the N-terminal amino acids (i.e., the mNTS) that were deleted. In the merge the DNA is stained through DAPI. Scale bar indicates 5 μm. On the right, multiplex western blots on isolated hydrogenosomes (TL, total lysate; Cy, Cytosol, Hy; hydrogenosomes) are shown along with a protease protection assay (TX, 0.1% TritonX-100; ProtK, Triton X-100 + 100 μg/μl proteinase K). Green bands represent the HA-tag, red bands *Tv*SCSα1. Numbers to the left indicate approximate molecular weights of the constructs in kilo Daltons (kDa).

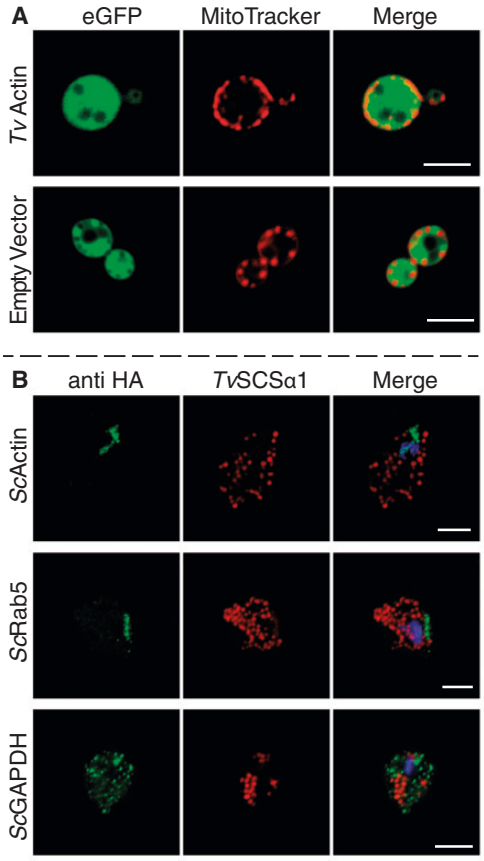


Fig. 4.—Cytosolic proteins do not colocalize with the organelles. (*A*) Expression of *Tv*Actin (*Trichomonas* actin) and the empty vector expressing the eGFP tag alone demonstrates that they do not localize to the mitochondria stained with MitoTracker® Red. (*B*) Three cytosolic proteins from yeast (*Sc*GAPDH, glyceraldehyde-3-phosphate dehydrogenase; *Sc*Actin, actin; *Sc*Rab5, rab family GTPase) were expressed using the same expression vector and yet do not colocalize with the hydrogenosomal marker *Tv*SCSα1.

Discussion

Organelles of mitochondrial origin share a common ancestry (Muller et al. 2012; Makiuchi and Nozaki 2014). In some eukaryotes, such as *Trichomonas* and *Giardia*, the organelles have undergone reduction to become hydrogenosomes and mitosomes, respectively. This process is accompanied by the loss of oxidative phosphorylation coupled with a loss of the electrochemical gradient, $\Delta\psi$, across the inner membrane, loss of the organellar genome and translation machinery, and a reduction in the number of proteins that are targeted to the organelles (Muller 1993; Goldberg et al. 2008; Jedelský et al. 2010; Schneider et al. 2011). Accompanying that biochemical and functional reduction, the protein import machinery has also undergone reduction from a very complex receptor platform in mitochondria to a more minimalistic import machinery in mitosomes (Doležal et al. 2006; Waller et al. 2009; Schleiff and Becker 2010). Targeting of matrix proteins to mitochondria is initiated through NTSs that are recognized by receptors, which are associated with the outer membrane of the organelle (Schleiff and Becker 2010). Although this process is conserved across all eukaryotes, the nature of the translocon machinery operating in the very earliest eukaryotes is still obscure.

Current views have it that the core translocons of TOM and TIM trace back to prokaryotic membrane proteins (Hewitt et al. 2011) and were hence present in the ancestor of mitochondria. A general analysis of eukaryotic porins revealed that Tom40 shares a significant structural homology with the beta-barrel structure of bacterial porins (Zeth and Thein 2010) and that the main translocation pores of the TIM complex, Tim23 and Tim17, evolved from common bacterial transporters (Rassow et al. 1999). In contrast to Tom40, which is highly conserved, Tom20 is far more variable than the core translocases and might have even evolved several times independently (Perry et al. 2006). That in turn suggests that the receptors for the NTS, although ubiquitous among eukaryotes, evolved after the origin of translocation pores of the two import complexes TOM and TIM, which were thus ancestral.

Essential components of TOM and TIM are conserved in hydrogenosomes and mitosomes (Regoes et al. 2005; Doležal et al. 2006; Rada et al. 2011). Mitosomes are even more reduced than hydrogenosomes (Waller et al. 2009; Heinz and Lithgow 2013), the organelles of *E. cuniculi* import only a few dozen proteins (Katinka et al. 2001; Waller et al. 2009), but also employ conserved TOM and TIM components (Waller et al. 2009). Similar to the situation with *Trichomonas* hydrogenosomes, mitosomal NTSs, when present, are short, with the majority of proteins targeted to microsporidian mitosomes lacking N-terminal extensions altogether (Katinka et al. 2001; Waller et al. 2009). Earlier findings that *Trichomonas* proteins localize to hydrogenosomes independent of their short hNTSs (Mentel et al. 2008; Burstein et al. 2012; Zimorski et al. 2013), along with similar observations for *Giardia* (Regoes et al. 2005) and more recently *Trypanosoma* (Hamilton et al. 2014), indicate that internal motifs of yet unknown nature can interact with the TOM translocon and mediate subsequent translocation of the organellar proteins without the need for an NTS. For those hydrogenosomal and mitosomal proteins, which have retained an NTS, the net positive charge—a conserved hallmark of mitochondrial NTSs (von Heijne 1986)—is lost (fig. 5*B*).

That the hNTS is not required for hydrogenosomal targeting is supported by our proteome analysis (supplementary table S1, Supplementary Material online). In 6 separate LC-MS/MS runs that were based on biological triplicates, 359 proteins were identified with a minimum of 2 peptides per protein. One hundred eighty-seven of the proteins that harbor a predicted NTS—and thus were good candidates to

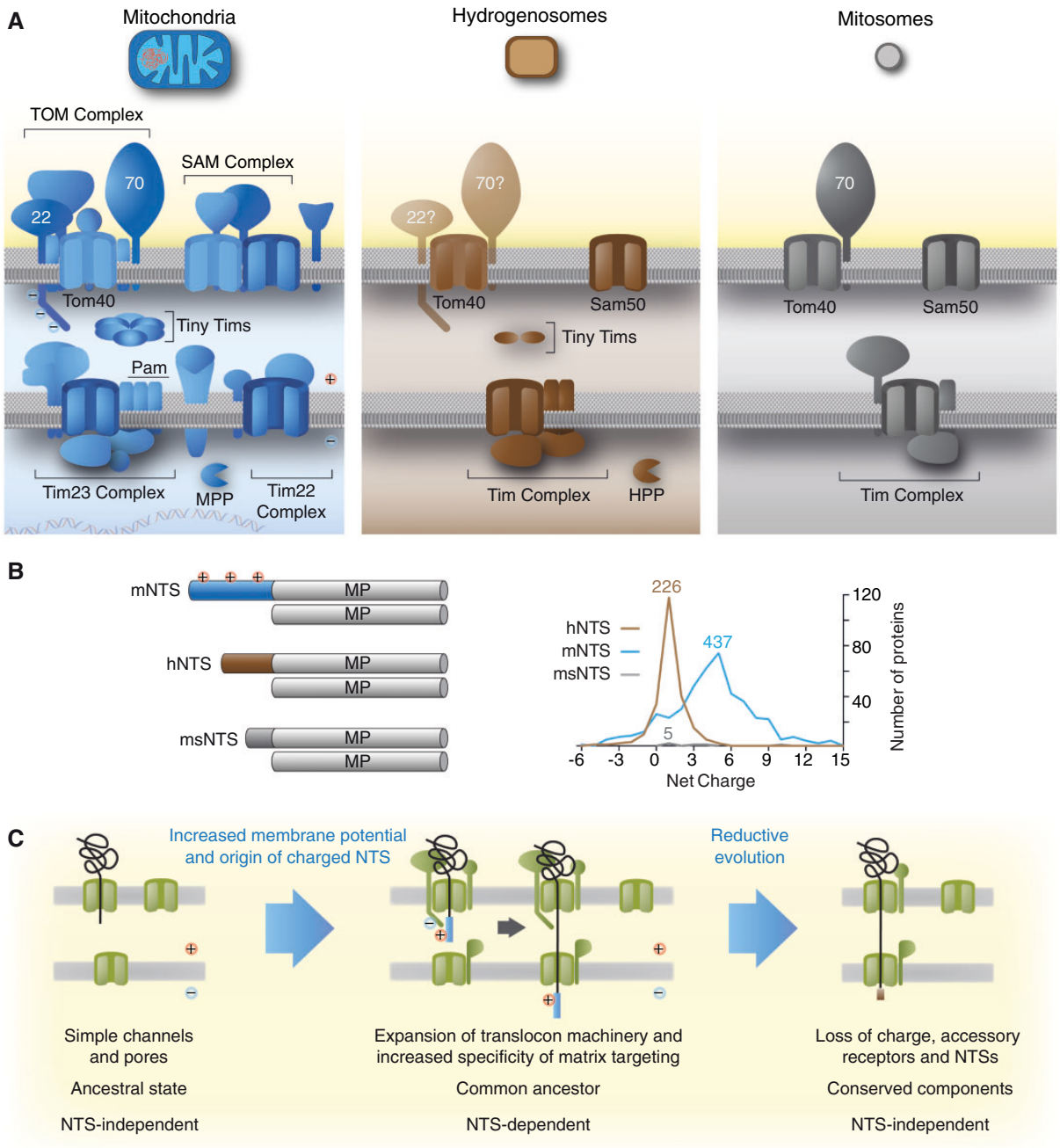


Fig. 5.—The complexity of the mitochondrial targeting machinery reduces together with the evolutionarily reduction of the organelle's biochemistry. (*A*) An illustration of the organellar import machineries from mitochondria, hydrogenosomes, and mitosomes. It is estimated that classical mitochondria import between 1,000 and 1,500 proteins, while hydrogenosomes of *Trichomonas* import only 200–500 proteins. (*B*) Mitochondrial matrix proteins carry long NTSs that are usually positively charged, but a few lack a recognizable mNTS. If present, the hydrogenosomal NTS is shorter and no longer charged. Mitosomal proteins of *Encephalitozoon cuniculi* completely lost the N-terminal targeting information and solely rely on internal motifs. Mitosomes of *E. cuniculi* are streamlined to such a degree, where they might require to import as little as 22 proteins from the cytosol; proteins that are largely responsible for the last known remaining biochemical pathway, which is iron–sulfur cluster (Fe–S) biosynthesis. (*C*) Early during mitochondrial evolution, transmembrane protein import was NTS independent. Import complexity (amount of proteins involved and the required targeting information) evolved downstream. During reductive evolution of the organelle, for instance in anaerobic parasites, import complexity decreased back, toward NTS-independent recognition of organellar proteins, translated in the cytosol.

Table 1
Summary of the Localization Studies

	Gene	NTS (a.a.)	Localization in *Trichomonas vaginalis*		Localization in *S. cerevisiae*	
			Full	Δm/hNTS	Full	Δm/hNTS
S. cerevisiae	*Sc*KGD2	40	Hy	Hy	Mt	Uk
	*Sc*IDH1	11	Hy	Hy	Mt	Uk
	*Sc*COXIV	25	Hy	Hy	Mt	Cy
	*Sc*LSC1	27	Hy	Hy	Mt	Uk
	*Sc*ACT1	—	Ct	—	Ct	—
	*Sc*GAPDH	—	Cy	—	Cy	—
	*Sc*Rab5	—	Cy/En	—	En	—
T. vaginalis	*Tv*SCSα1	9	Hy	Hy	Mt	Mt
	*Tv*ME	12	Hy	Hy	Mt	Mt
	*Tv*ISCA1	9	Hy	Hy	Mt	Mt
	*Tv*Fdx	8	Hy	Hy	Mt	Mt
	*Tv*Actin	—	Ct	—	Cy	—
	eGFP	—	—	—	Cy	—

Note.—The length of the NTS was determined using TargetP (Emanuelsson et al. 2000). Hy, hydrogenosomes; Mt, mitochondria; Cy, cytosol; Ct, cytoskeleton; En, endosomes; Uk, unknown.

constitute the core hydrogenosomal proteome—were absent from our analysis. They might be of low abundance, not expressed at all, or further evidence that the correct targeting to hydrogenosomes does not hinge upon the presence of that very hNTS. Three hundred twenty proteins (~90% of the hydrogenosomal proteins identified) lacked an hNTS (fig. 1). The lack of NTS is not without precedent in mitochondria. A global N-proteome of yeast mitochondria identified 400 proteins with processed N-termini of at least 10 a.a. (amino acids) out of 585 proteins identified in total (Vögtle et al. 2009). Although many yeast mitochondrial proteins are directed to the inner and outer membranes as well as to the intermembrane space (IMS) without an NTS, targeting to the yeast mitochondrial matrix appears to be NTS dependent in cases reported so far (Neupert and Herrmann 2007; Chacinska et al. 2009; Schleiff and Becker 2010; Neupert 2015).

Our present data (summarized in table 1) provide more evidence that in *Trichomonas* hydrogenosomal targeting works in the absence of an NTS, albeit some might still require it (Mentel et al. 2008; Zimorski et al. 2013). In any case, this mode of NTS-independent targeting is—at least for the proteins tested—conserved in yeast, because hydrogenosomal matrix proteins lacking an NTS are directed to the yeast mitochondrion (fig. 2). Moreover, the converse is true of yeast matrix proteins in *Trichomonas* hydrogenosomes (fig. 3). Vector-caused localization artifacts can be ruled out. Our controls (empty vector and fusion proteins involving nonorganellar proteins of *Trichomonas* and yeast) never colocalized with hydrogenosomal markers (fig. 4). In addition, the vector used for the transfection of *Trichomonas* was previously used to analyze surface proteins (Noël et al. 2010), nuclear proteins (Zubácová et al. 2012), and cytoskeletal proteins (Kusdian et al. 2013). None of these fusion proteins associated with the hydrogenosomes. The same is true for yeast and the pYES2/CT plasmid (Donahue et al. 2001; Todisco et al. 2014). In summary, this indicates that proteins of mitochondrial ancestry have yet unspecified properties that mediate interactions with the Tom40 translocon, which was present in the earliest eukaryotes. The nature or identity of these properties remains so far unidentified.

This raises a curious question: If Tom40 can recognize its own substrates, why did NTS-dependent targeting evolve in the first place, and more intriguingly, why is it preferentially lost in hydrogenosomes and mitosomes? One possible rationale for the origin of NTS-dependent targeting is specificity. The presence of a dedicated receptor-ligand (TOM-NTS) pair for recognition and import would allow increased specificity of TOM interactions and thus channel substrates to the TIM complex. Although the origin of a sophisticated receptor platform including Tom20, Tom22, and Tom70 (fig. 4*A*) might have been selected for NTS recognition and specificity, import of proteins lacking an NTS is also specific (Regoes et al. 2005; Goldberg et al. 2008; Mentel et al. 2008; Šmíd et al. 2008; Waller et al. 2009; Burstein et al. 2012; Zimorski et al. 2013; Hamilton et al. 2014). Hence, receptor interactions at the TOM complex alone cannot explain the presence of an NTS. We suggest that conservation of NTS-independent targeting of yeast and trichomonad proteins to the organelle constitute conserved, not convergent properties, and that they reflect the ancestral state of mitochondrial protein recognition and import from the cytosol.

It is well established that mitochondrial membrane potential electrophoretically directs the NTS to the TIM channel via the negatively charged tail of Tom22 (Pfanner and Neupert 1986; Martin et al. 1991; Esaki et al. 2004). In accordance with the "increasing affinity" model (Schleiff and Becker 2010), Tom22 binds the positively charged NTS within the IMS and recruits the TIM complex and TOM and TIM form a continuous pore across both membranes (Schleiff and Becker 2010). Noncleavable internal motifs target proteins to the mitochondrial IMS and the membranes (Chacinska et al. 2009; Schleiff and Becker 2010). In case of IMS proteins that have a charged mNTS, like cytochrome b2, the mNTS needs to traverse the matrix first and manipulation of the charged region of the mNTS decreases import efficiency (Geissler et al. 2000). Indeed in some cases, cytochrome b2 destined to the IMS lacks an NTS altogether (Hewitt et al. 2012).

We propose that in the eukaryotic common ancestor, a positively charged NTS was initially selected at the termini of matrix proteins for their electrophoretic import via the membrane potential across the inner membrane (fig. 5*C*), providing

specific targeting to their designated compartment—the matrix. The evolution of the positive charge on the NTS also allowed discrimination between mitochondrial matrix and nonmatrix destinations. The ease with which a functional NTS can be obtained through random DNA sequences (Baker and Schatz 1987) indicates that the evolutionary origin of transit peptides was facile, requiring virtually no innovation at all (Lucattini et al. 2004), merely selection for accrual of positive charges on the N-termini of matrix-specific proteins and for subsequent proteolytic processing via the conserved mitochondrial processing peptidase (Šmíd et al. 2008).

The simpler nature of protein import in hydrogenosomal and mitosomal evolution has often, and rightly, been attributed to the general process of reductive evolution (Van der Giezen et al. 2002; Doležal et al. 2006; Šmíd et al. 2008). Our proposal that positive charge on the NTS arose as a matrix-specific targeting signal suggests what, exactly, was lost first (the charge), while uncovering the existence and conservation—though not the nature—of NTS-independent import signals in *Trichomonas* and yeast. Loss of the electron transport chain in the inner membrane in hydrogenosomes (and mitosomes) led to loss of $\Delta\psi$, rendering positive charge on the NTS superfluous, hence readily lost through mutation. This accounts for the conspicuous lack of charge in hydrogenosomal and mitosomal NTSs. In the absence of charge, the NTS itself could however only become expendable in the event that either 1) a novel NTS-independent import pathway arose in the inner membrane in a lineage specific manner or 2) a conserved import pathway pre-existed that accommodated NTS-independent import. Conservation of NTS-independent targeting in *Trichomonas* and yeast indicate that the latter was the case. Our results bring into question the prevalence, evolutionary conservation, and antiquity of internal or cryptic signals in proteins targeted to mitochondrial organelles.

Supplementary Materials

Supplementary tables S1 and S2 are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

Financial support through the Deutsche Forschungsgemeinschaft to S.B.G. (GO1825/4-1), an ERC grant to W.F.M. (666053) and the Czech Science Foundation (13-09208J) to J.T. is gratefully acknowledged. We thank Kai Stühler and Gereon Poschmann of the BMFZ HHU-Düsseldorf for their excellent support and help analyzing the proteome data.

Literature Cited

Aurrecoechea C, et al. 2008. GiardiaDB and TrichDB: integrated genomic resources for the eukaryotic protist pathogens *Giardia lamblia* and *Trichomonas vaginalis*. Nucleic Acids Res. 37:D526–D530.

Baker A, Schatz G. 1987. Sequences from a prokaryotic genome or the mouse dihydrofolate reductase gene can restore the import of a truncated precursor protein into yeast mitochondria. Proc Natl Acad Sci U S A. 84:3117–3121.

Burstein D, et al. 2012. A machine learning approach to identify hydrogenosomal proteins in *Trichomonas vaginalis*. Eukaryot Cell. 11:217–228.

Carlton JM, et al. 2007. Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. Science 315:207–212.

Chacinska A, Koehler CM, Milenkovic D, Lithgow T, Pfanner N. 2009. Importing mitochondrial proteins: machineries and mechanisms. Cell 138:628–644.

Doležal P, Likic V, Tachezy J, Lithgow T. 2006. Evolution of the molecular machines for protein import into mitochondria. Science 313:314–318.

Donahue SL. 2001. Mitochondrial DNA ligase function in *Saccharomyces cerevisiae*. Nucleic Acids Res. 29:1582–1589.

Emanuelsson O, Nielsen H, Brunak S, von Heijne G. 2000. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol. 300:1005–1016.

Endo T, Yamano K. 2009. Transport of proteins across or into the mitochondrial outer membrane. Biochim Biophys Acta. 1803:706–714.

Esaki M, et al. 2004. Mitochondrial protein import. J Biol Chem. 279:45701–45707.

Geissler A, et al. 2000. Membrane potential-driven protein import into mitochondria. The sorting sequence of cytochrome b(2) modulates the deltapsi-dependence of translocation of the matrix-targeting sequence. Mol Biol Cell. 11:3977–3991.

Goldberg AV, et al. 2008. Localization and functionality of microsporidian iron–sulphur cluster assembly proteins. Nature 452:624–628.

Gorrell TE, Yarlett N, Müller M. 1984. Isolation and characterization of *Trichomonas vaginalis* ferredoxin. Carlsberg Res Commun. 49:259–268.

Gregg C, Kyryakov P, Titorenko VI. 2009. Purification of mitochondria from yeast cells. J Vis Exp. 24.

Hamilton V, Singha UK, Smith JT, Weems E, Chaudhuri M. 2014. Trypanosome alternative oxidase possesses both an N-terminal and internal mitochondrial targeting signal. Eukaryot Cell. 13:539–547.

Häusler T, Stierhof YD, Blattner J, Clayton C. 1997. Conservation of mitochondrial targeting sequence function in mitochondrial and hydrogenosomal proteins from the early-branching eukaryotes Crithidia, Trypanosoma and Trichomonas. Eur J Cell Biol. 73:240–251.

Heinz E, Lithgow T. 2013. Back to basics: a revealing secondary reduction of the mitochondrial protein import pathway in diverse intracellular parasites. Biochim Biophys Acta. 1833:295–303.

Hewitt V, Alcock F, Lithgow T. 2011. Minor modifications and major adaptations: the evolution of molecular machines driving mitochondrial protein import. Biochim Biophys Acta. 1808:947–954.

Hewitt VL, et al. 2012. A model system for mitochondrial biogenesis reveals evolutionary rewiring of protein import and membrane assembly pathways. Proc Natl Acad Sci U S A. 109:E3358–E3366.

Jedelský PL, et al. 2010. The minimal proteome in the reduced mitochondrion of the parasitic protist *Giardia intestinalis*. PLoS One 6:e17285–e17285.

Katinka MD, et al. 2001. Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. Nature 414:450–453.

Kusdian G, Woehle C, Martin WF, Gould SB. 2013. The actin-based machinery of *Trichomonas vaginalis* mediates flagellate-amoeboid

transition and migration across host tissue. Cell Microbiol.15: 1707–1721.
Land KM, et al. 2003. Targeted gene replacement of a ferredoxin gene in *Trichomonas vaginalis* does not lead to metronidazole resistance. Mol Microbiol. 51:115–122.
Lane N, Martin W. 2010. The energetics of genome complexity. Nature 467:929–934.
Lill R, Neupert W. 1996. Mechanisms of protein import across the mitochondrial outer membrane. Trends Cell Biol. 6:56–61.
Lindmark DG, Muller M. 1973. Hydrogenosome, a cytoplasmic organelle of the anaerobic flagellate *Tritrichomonas foetus*, and its role in pyruvate metabolism. J Biol Chem. 248:7724–7728.
Lucattini R, Likic VA, Lithgow T. 2004. Bacterial proteins predisposed for targeting to mitochondria. Mol Biol Evol. 21:652–658.
Makiuchi T, Nozaki T. 2014. Highly divergent mitochondrion-related organelles in anaerobic parasitic protozoa. Biochimie. 100:3–17.
Martin J, Mahlke K, Pfanner N. 1991. Role of an energized inner membrane in mitochondrial protein import. Delta psi drives the movement of presequences. J Biol Chem. 266:18051–18057.
Martin W, Koonin EV. 2006. Introns and the origin of nucleus—cytosol compartmentalization. Nature 440:41–45.
Mclnerney JO, O'Connell MJ, Pisani D. 2014. The hybrid nature of the Eukaryota and a consilient view of life on Earth. Nat Rev Microbiol. 12:449–455.
Meisinger C, Sickmann A, Pfanner N. 2008. The mitochondrial proteome: from inventory to function. Cell 134:22–24.
Mentel M, Zimorski V, Haferkamp P, Martin W, Henze K. 2008. Protein import into hydrogenosomes of *Trichomonas vaginalis* involves both N-terminal and internal targeting signals: a case study of thioredoxin reductases. Eukaryot Cell. 7:1750–1757.
Mi-ichi F, Abu Yousuf M, Nakada-Tsukui K, Nozaki T. 2009. Mitosomes in *Entamoeba histolytica* contain a sulfate activation pathway. Proc Natl Acad Sci U S A. 106:21731–21736.
Muller M. 1993. The hydrogenosome. J Gen Microbiol. 139:2879–2889.
Muller M, et al. 2012. Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. Microbiol Mol Biol Rev. 76:444–495.
Neupert W. 2015. A perspective on transport of proteins into mitochondria: a myriad of open questions. J Mol Biol. 427:1135–1158.
Neupert W, Herrmann JM. 2007. Translocation of proteins into mitochondria. Annu Rev Biochem. 76:723–749.
Noël CJ, et al. 2010. Trichomonas vaginalis vast BspA-like gene family: evidence for functional diversity from structural organisation and transcriptomics. BMC Genomics 11:99–125.
Pérez JMM, Pascau J. 2013. Image processing with ImageJ. Packt Publishing Ltd.
Perry AJ, Hulett JM, Likic VA, Lithgow T, Gooley PR. 2006. Convergent evolution of receptors for protein import into mitochondria. Curr Biol. 16:221–229.
Pfanner N, Neupert W. 1986. Transport of F1-ATPase subunit beta into mitochondria depends on both a membrane potential and nucleoside triphosphates. FEBS Lett. 209:152–156.
Rada P, et al. 2011. The core components of organelle biogenesis and membrane transport in the hydrogenosomes of *Trichomonas vaginalis*. PLoS One 6:e24428.
Rassow J, Dekker PJT, van Wilpe S, Meijer M, Soll J. 1999. The preprotein translocase of the mitochondrial inner membrane: function and evolution. J Mol Biol. 286:105–120.
Regoes A, et al. 2005. Protein import, replication, and inheritance of a vestigial mitochondrion. J Biol Chem. 280:30557–30563.
Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. 16:276–277.
Schleiff EE, Becker TT. 2010. Common ground for protein translocation: access control for mitochondria and chloroplasts. Nat Rev Mol Cell Biol. 12:48–59.
Schneider RE, et al. 2011. The *Trichomonas vaginalis* hydrogenosome proteome is highly reduced relative to mitochondria, yet complex compared with mitosomes. Int J Parasitol. 41:1421–1434.
Šmíd O, et al. 2008. Reductive evolution of the mitochondrial processing peptidases of the unicellular parasites *Trichomonas vaginalis* and *Giardia intestinalis*. PLoS Pathog. 4:e1000243.
Timmis JN, Ayliffe MA, Huang CY, Martin W. 2004. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Rev Genet. 5:123–135.
Todisco S, et al. 2014. The *Saccharomyces cerevisiae* gene YPR011c encodes a mitochondrial transporter of adenosine 5′-phosphosulfate and 3′-phospho-adenosine 5′-phosphosulfate. Biochim Biophys Acta. 1837:326–334.
Van der Giezen M, et al. 2002. Conserved properties of hydrogenosomal and mitochondrial ADP/ATP carriers: a common origin for both organelles. EMBO J. 21:572–579.
Vögtle FN, et al. 2009. Global analysis of the mitochondrial N-proteome identifies a processing peptidase critical for protein stability. Cell 139:428–439.
von Heijne G. 1986. Mitochondrial targeting sequences may form amphiphilic helices. EMBO J. 5:1335–1471.
Waller RF, et al. 2009. Evidence of a reduced and modified mitochondrial protein import apparatus in microsporidian mitosomes. Eukaryot Cell. 8:19–26.
Williams TA, Foster PG, Cox CJ, Embley TM. 2013. An archaeal origin of eukaryotes supports only two primary domains of life. Nature 504:231–236.
Zeth K, Thein M. 2010. Porins in prokaryotes and eukaryotes: common themes and variations. Biochem J. 431:13–22.
Zimorski V, et al. 2013. The N-terminal sequences of four major hydrogenosomal proteins are not essential for import into hydrogenosomes of *Trichomonas vaginalis*. J Eukaryot Microbiol. 60:89–97.
Zubácová Z, Hostomská J, Tachezy J. 2012. Histone H3 variants in *Trichomonas vaginalis*. Eukaryot Cell. 11:654–661.

Associate editor: Takashi Gojobori

Publication IV

N-Terminal presequence-independent import of phosphofructokinase into hydrogenosomes of *Trichomonas vaginalis*

Petr Rada[1], Abhijith Radhakrishna Makki[1], Verena Zimorski[2], **Sriram Garg**[2], Vladimír Hampl[1], Ivan Hrdý[1], Sven B. Gould[2] and Jan Tachezy[1]

1 - Department of Parasitology, Charles University in Prague, Faculty of Science, Prague, Czech Republic

2 - Institute for Molecular Evolution, Heinrich-Heine Universität Düsseldorf, Universitätsstraße 1, Düsseldorf 40225, Germany

The presented manuscript was peer-reviewed and published in **Eukaryotic Cell** with an impact factor of **3.328** on 16th October 2015

Contribution as **fourth author**

15%

Assisted in formulating the discussion for the manuscript and corrected the written text including mining the literature for relevant citations

EC
Journals.ASM.org

N-Terminal Presequence-Independent Import of Phosphofructokinase into Hydrogenosomes of *Trichomonas vaginalis*

Petr Rada,[a] Abhijith Radhakrishna Makki,[a] Verena Zimorski,[b] Sriram Garg,[b] Vladimír Hampl,[a] Ivan Hrdý,[a] Sven B. Gould,[b] Jan Tachezy[a]

Department of Parasitology, Charles University in Prague, Faculty of Science, Prague, Czech Republic[a]; Institute for Molecular Evolution, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany[b]

Mitochondrial evolution entailed the origin of protein import machinery that allows nuclear-encoded proteins to be targeted to the organelle, as well as the origin of cleavable N-terminal targeting sequences (NTS) that allow efficient sorting and import of matrix proteins. In hydrogenosomes and mitosomes, reduced forms of mitochondria with reduced proteomes, NTS-independent targeting of matrix proteins is known. Here, we studied the cellular localization of two glycolytic enzymes in the anaerobic pathogen *Trichomonas vaginalis*: PP_i-dependent phosphofructokinase (*Tv*PP_i-PFK), which is the main glycolytic PFK activity of the protist, and ATP-dependent PFK (*Tv*ATP-PFK), the function of which is less clear. *Tv*PP_i-PFK was detected predominantly in the cytosol, as expected, while all four *Tv*ATP-PFK paralogues were imported into *T. vaginalis* hydrogenosomes, although none of them possesses an NTS. The heterologous expression of *Tv*ATP-PFK in *Saccharomyces cerevisiae* revealed an intrinsic capability of the protein to be recognized and imported into yeast mitochondria, whereas yeast ATP-PFK resides in the cytosol. *Tv*ATP-PFK consists of only a catalytic domain, similarly to "short" bacterial enzymes, while *Sc*ATP-PFK includes an N-terminal extension, a catalytic domain, and a C-terminal regulatory domain. Expression of the catalytic domain of *Sc*ATP-PFK and short *Escherichia coli* ATP-PFK in *T. vaginalis* resulted in their partial delivery to hydrogenosomes. These results indicate that *Tv*ATP-PFK and the homologous ATP-PFKs possess internal structural targeting information that is recognized by the hydrogenosomal import machinery. From an evolutionary perspective, the predisposition of ancient ATP-PFK to be recognized and imported into hydrogenosomes might be a relict from the early phases of organelle evolution.

The transition of the mitochondrion into an ATP-producing organelle was the crucial event at the eukaryote origin (1). ATP synthesis in eukaryotes is typically compartmentalized, with glycolysis in the cytosol and pyruvate oxidation in the mitochondria, which is linked to highly efficient oxidative phosphorylation (1, 2). In protists, however, there are notable exceptions to the usual scheme regarding both glycolysis and pyruvate oxidation. In *Trichomonas vaginalis* and other eukaryotes that possess an anaerobic form of mitochonria called hydrogenosomes, pyruvate is oxidized within the organelle via less efficient anaerobic fermentation (3). *Giardia intestinalis*, *Entamoeba histolytica*, and other eukaryotes possess a reduced form of mitochondria called mitosomes that do not produce ATP at all (4). In these organisms, pyruvate oxidation takes place exclusively in the cytosol (1). In kinetoplastids, glycolysis is compartmentalized in specialized microbodies called glycosomes (5). In some green algae, the first half of the glycolytic pathway is localized in the chloroplast (6, 7), while in the diatom *Phaeodactylum tricornutum* and other stramenopiles, several glycolytic enzymes are targeted to multiple compartments, such as the cytosol, plastids, and mitochondria (8, 9).

A particularly vexing case of compartmentalization involves *T. vaginalis* phosphofructokinase (PFK). In *Trichomonas*, glycolysis proceeds via a pyrophosphate (PP_i)-dependent phosphofructokinase (PP_i-PFK) (10), an enzyme that is generally rare in eukaryotes, albeit typical in plants (11). Therefore, it was surprising that genes for ATP-dependent phosphofructokinase (ATP-PFK) turned up in the *Trichomonas* genome (12). Furthermore, peptides of the expressed protein were found in the hydrogenosomal proteome (13–15), although the exact topology of hydrogenosome-associated *T. vaginalis* ATP-PFK (*Tv*ATP-PFK) remains unclear (13, 15). PP_i-PFK and ATP-PFK share an evolutionary origin (16, 17). In bacteria, ATP-PFK is a homo-oligomeric enzyme that is formed by ~35-kDa subunits (18). In opisthokonts, ATP-PFK underwent gene duplication and fusion events, resulting in an ~90-kDa protein with an N-terminal catalytic domain and a C-terminal regulatory domain (19). The PP_i-PFK protein forms homo- or, in plants, heterotetramers of ~40- to 60-kDa subunits, and in Apicomplexa, the two subunits are fused to a protein of ~140 kDa (20). The advantage of using PP_i-PFK rather than ATP-PFK in glycolysis lies in the increased yield of ATP due to the replacement of ATP with PP_i as a phosphate donor in the phosphorylation of fructose-6-phosphate (3). This is particularly important for *T. vaginalis* and other anaerobes with energy metabolism based mainly on glycolysis (10).

In most eukaryotes, the N-terminal targeting sequences (NTS) are required for the delivery of nuclear-encoded proteins into the mitochondrial matrix, whereas the NTS-independent pathway is mainly involved in the routing of proteins into the outer and inner mitochondrial membranes and the intermembrane space. NTS are typically 15 to 55 residues in length and form a positively

Received 4 July 2015 **Accepted** 8 October 2015

Accepted manuscript posted online 16 October 2015

Citation Rada P, Makki AR, Zimorski V, Garg S, Hampl V, Hrdý I, Gould SB, Tachezy J. 2015. N-terminal presequence-independent import of phosphofructokinase into hydrogenosomes of *Trichomonas vaginalis*. Eukaryot Cell 14:1264–1275. doi:10.1128/EC.00104-15.

Address correspondence to Jan Tachezy, tachezy@natur.cuni.cz.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/EC.00104-15.

charged amphipathic α-helix (21). Upon preprotein delivery into the matrix by the outer (TOM) and inner (TIM) membrane translocases, the NTS is removed by a heterodimeric zinc-dependent mitochondrial processing peptidase (MPP) (22). Proteins routed by the NTS-independent pathway possess either a single or multiple internal targeting signals (ITS) (23). In *Saccharomyces cerevisiae* and human mitochondria, the components and mechanisms of protein import via the NTS-dependent pathway are well characterized (23), whereas less is known about protein import in hydrogenosomes. The NTS-dependent mechanism is present in hydrogenosomes and mitosomes (4, 24, 25), but a few studies have also reported NTS-independent import into the hydrogenosomes of *T. vaginalis* (26, 27, 58).

Interestingly, there are four ~35-kDa *Tv*ATP-PFK proteins encoded in the *T. vaginalis* genome, none of which possesses an NTS. The multiple copies preclude the generation of *Tv*ATP-PFK knockouts with current *Trichomonas* tools to study their functions, which remain mysterious. To clarify the localization and exact organellar topology of *Tv*ATP-PFK, we investigated the targeting of products encoded by *Tv*ATP-PFK genes when expressed in transformed *T. vaginalis* cells using immunofluorescence microscopy and cell fractionation, characterized the ATP dependence of *Tv*ATP-PFK import into isolated hydrogenosomes, and tested whether *Tv*ATP-PFK could be recognized as a substrate for NTS-independent import into yeast mitochondria. Conversely, we assessed whether the homologous catalytic domain of yeast ATP-PFK, as well as ~35-kDa *Escherichia coli* ATP-PFK (*Ec*ATP-PFK), showed a tendency to be imported into hydrogenosomes when expressed in *T. vaginalis*.

MATERIALS AND METHODS

T. vaginalis strain T1 (provided by J.-H. Tai, Institute of Biomedical Sciences, Taipei, Taiwan) was grown in Diamond's tryptone-yeast extract-maltose (TYM) medium supplemented with 10% (vol/vol) heat-inactivated horse serum. *S. cerevisiae* strain INVSc1 (Invitrogen) was grown in yeast extract-peptone-dextrose (YPD) medium or minimal medium devoid of uracil when transfected.

Phylogenetic analyses. The sequences of ATP-PFK and PP_i-PFK in a wide diversity of prokaryotes and eukaryotes were downloaded from the protein and EST database of GenBank release 200.0 and aligned with the *T. vaginalis* sequences with MAFFT (28; http://mafft.cbrc.jp/alignment/server/) using an L-INS-i strategy. The alignment was manually edited using BioEdit 7.0.9.0 (29), and 340 well-aligned positions were used for the subsequent analyses. The phylogenetic tree was constructed by the maximum-likelihood method in RAxML version 7.2.8 (30) using the PROTGAMMALGF model on the RAxML black box server (31). The statistical support was assessed by bootstrapping with 100 repetitions in RAxML. Bayesian posterior probabilities were calculated in Phylobayes (32) on the CIPRES Science Gateway v. 3.3 (http://www.phylo.org/index.php/). Two chains of Markov chain Monte Carlo were run under the CAT GTR model with a sampling frequency of 1,800. The run was terminated when the discrepancy observed across all bipartitions (maxdiff) dropped below 0.3 and effective sizes were larger than 50. The first 500 trees were discarded as burn in, and a consensus tree with posterior probabilities was calculated from the sample of 14,080 trees.

Gene cloning and transformation. Selected genes (*Tv*ATP-PFK1, TVAG_293770; *Tv*PP_i-PFK1, TVAG_430830; *T. vaginalis* ferredoxin 1 [Fdx1], TVAG_003900; *S. cerevisiae*ATP-PFK [*Sc*ATP-PFK], DAA08331; and *E. coli Ec*ATP-PFK, EFJ85506.1) were amplified by PCR from *T. vaginalis* and *S. cerevisiae* genomic DNA and cloned into the plasmids (i) pTagVag2, enabling the expression of the inserted genes with a C-terminal dihemagglutinin (di-HA) tag in trichomonads (33), and (ii) a self-modified version of plasmid pYES2/CT that allows the expression of the inserted genes with C-terminal green fluorescent protein (GFP) in yeasts. Transformed trichomonads and *S. cerevisiae* cells were selected as previously described (33, 34). The primers that were used for amplification and cloning of the selected genes into the pTagVag2 and pYES2/CT plasmids are shown in the supplemental material.

The pTagVag2 plasmid allows expression of the inserted genes under the control of the *T. vaginalis* hydrogenosomal α-subunit succinyl-coenzyme A (CoA) synthetase (SCSα) gene promoter (33). Alternatively, we used native promoters of selected genes instead of the SCSα promoter. The selected genes were amplified by PCR with 300 bp of upstream noncoding sequences and inserted into the pTagVag2 plasmid with a deleted SCSα promoter (pTagVagN). The primers used to amplify and clone the selected genes with their native promoters are shown in the supplemental material.

Immunofluorescence microscopy. Episomally expressed recombinant proteins were detected in trichomonads using a monoclonal mouse anti-HA antibody (35). In double-labeling experiments, hydrogenosomal malic enzyme was detected using a rabbit polyclonal antibody (36). A secondary Alexa Fluor 488 (green) donkey anti-mouse antibody and Alexa Fluor 594 (red) donkey anti-rabbit antibody were used for visualization of target proteins. The cells were examined using an Olympus Cell-R IX81 microscope system. The acquired images were processed using ImageJ software (version 1.4d) (http://rsbweb.nih.gov/ij/). In *S. cerevisiae* cells, episomally expressed recombinant proteins with GFP were detected and examined as described above. In double-labeling experiment, mitochondria were detected with MitoTracker dye (Invitrogen).

Enzyme assays. ATP-PFK activity was determined in the glycolytic direction using a continuous spectrophotometric assay according to the method of Chi et al. (37) with some modifications. The assay mixture for ATP-PFK consisted of 2 ml of 100 mM HEPES, 50 mM KCl, 3 mM $MgCl_2$, 1 mM EDTA, pH 7.0, buffer; 1 mM ATP; 20 mM fructose-6-phosphate; 0.15 to 0.20 mM NADH; 2 to 3 U each of aldolase, triosephosphate isomerase, and glycerol-3-phosphate dehydrogenase (Sigma-Aldrich); and 0.05% (vol/vol) Triton X-100 (ATP-PFK assay buffer). The assay was performed in 1-cm anaerobic cuvettes. The reaction was started by alternatively adding ATP, fructose-6-phosphate, auxiliary enzymes, or protein sample to the assay mixture, and the reaction was monitored as a decrease in the absorbance of NADH at 340 nm using a Shimadzu UV-2600 spectrophotometer. PP_i-PFK activity was determined as previously described (38). The protein concentrations in the subcellular fractions of *T. vaginalis* were determined by the Lowry protein assay.

Preparation of cellular fractions. Highly purified hydrogenosomes were obtained from *T. vaginalis* total cell lysates by differential and Percoll gradient centrifugation as described previously (35). The cytosolic fraction was isolated according to the method of Sutak et al. (35) and subsequently centrifuged at 190,000 × *g* (the high-speed cytosolic fraction). Mitochondria of *S. cerevisiae* were isolated from the yeast according to the method of Gregg et al. (39).

Protease protection assay. Aliquots of intact hydrogenosomes (3 mg) were resuspended in 1 ml of 1× ST buffer (250 mM sucrose, 10 mM Tris, pH 7.8, 0.5 mM KCl) supplemented with protease inhibitor cocktail tablets (Roche Complete, EDTA free). Trypsin (Sigma) was added to a final concentration of 200 μg/ml, and the samples were incubated at 37°C for 30 min. After incubation, the trypsin activity was stopped by the addition of soybean inhibitors (5 mg/ml), and the samples were analyzed by immunoblotting with a monoclonal mouse anti-HA antibody.

Aliquots of intact mitochondria (1 mg) were resuspended in 1 ml of SEM buffer (1 mM MOPS [morpholinepropanesulfonic acid]-KOH, pH 7.2, 250 mM sucrose, 1 mM EDTA). Proteinase K (Sigma) was added to a final concentration of 50 μg/ml, and the samples were incubated at 37°C for 30 min. After incubation, the proteinase K activity was stopped by the addition of 250 μl of trichloroacetic acid. The samples were analyzed by immunoblotting with a monoclonal anti-GFP antibody (Pierce).

Preparation of radiolabeled precursor proteins. The *Tv*ATP-PFK1 gene was cloned into the modified psp64 poly(A) plasmid, which enables

in vitro mRNA synthesis from the inserted genes (Promega). The primers designed for PCR and cloning into the psp64 plasmid are described in the supplemental material. *In vitro* transcription was performed using the mMachine kit (Ambion). [^{35}S]methionine-radiolabeled precursor protein was synthesized *in vitro* using the Flexi Rabbit Reticulocyte Lysate System (Promega).

In vitro import. Each *in vitro* import assay was performed in a reaction mixture that included 100 μl of import buffer (10 mM HEPES, pH 7.4, 250 mM sucrose, 2 mM KP_i, pH 7.4, 25 mM KCl, 10 mM $MgCl_2$, 0.5 mM EDTA, pH 8.0, 1 mM dithiothreitol [DTT], 10 mM ATP), 50 μl of cytosolic extract, 5 μl of radiolabeled precursor protein, and 5 mg of isolated hydrogenosomes. Apyrase (20 U/ml) was used for the import assay, which was conducted in the absence of ATP. The organelles were preincubated for 10 min at 25°C in import buffer with cytosolic extract, after which radiolabeled precursor protein was added to the assay mixture, and the mixture was incubated for 1, 10, and 60 min at 25°C. At each time point, the *in vitro* import was stopped by the addition of 100 μg/ml of proteinase K and placed on ice for 20 min. After incubation, the activity of proteinase K was inhibited by adding 2 mM phenylmethylsulfonyl fluoride (PMSF) (Sigma). The hydrogenosomes were then washed in import buffer and solubilized in SDS loading buffer. To test the activity of proteinase K, after a 60-min incubation of the protein import reaction mixture, the hydrogenosomes were dissolved with 0.5% (vol/vol) Triton X-100, followed by the addition of 100 μg/ml of proteinase K. Proteins in the supernatant were precipitated with methanol-chloroform and solubilized in SDS loading buffer. All of the samples were subjected to SDS-PAGE in a 13.5% separating gel. The gels were vacuum dried and exposed to X-ray films.

RESULTS

Phylogenetic analysis reveals the presence of PP_i-PFK and the short type of ATP-PFK in *T. vaginalis* and other parabasalids. The *T. vaginalis* genome possesses 11 genes encoding phosphofructokinases, four of which encode "short" (~35-kDa)-type ATP-dependent PFKs (*Tv*ATP-PFK1 to -4 [TVAG_293770, TVAG_496160, TVAG_462920, and TVAG_391760]) and seven of which encode PP_i-dependent PFKs (*Tv*PP_i-PFK1 to -7 [TVAG_430830, TVAG_077440, TVAG_281070, TVAG_364620, TVAG_079260, TVAG_263690, and TVAG_335880]). A phylogenetic analysis of ATP-PFKs and PP_i-PFKs revealed that *Trichomonas* *Tv*ATP-PFK1 to -4 fall into the single robust clade T2, together with PFKs from other parabasalids (Fig. 1). The closest eukaryotic relatives of this clade are tandem-fusion PFKs from opisthokonts and amoebozoans (clade E), as well as enzymes from prokaryotes (clades B1 and B2). The *Trichomonas* homologues *Tv*PP_i-PFK1 to -7 also form a clade with parabasalian sequences (Fig. 1). This parabasalian clade (clade T1) branches with enzymes from jakobids, heteroloboseans, and prokaryotes. The presence of both versions of the enzyme in other parabasalids suggests that both PP_i-PFK and ATP-PFK were present in the parabasalid ancestor. The branching of the *T. vaginalis* sequences in several unrelated positions in both clades T1 and T2 indicates that the genes have undergone gene duplications and possibly gene losses within parabasalids. The specificity of both types of PFKs for either ATP or PP_i has been ascribed to the amino acid residues at positions 104 and 124 (according to the numbering of the *E. coli Ec*ATP-PFK [40]). The G_{104} ($GGDG_{104}$ motif) and G/K_{124} residues are important for ATP binding, whereas PP_i binding requires residues D_{104} ($GGDD_{104}$ motif) and K_{124} (10, 17). *Tv*ATP-PFK1, -3, and -4 contain glycine at position 104, and *Tv*ATP-PFK1 and -3 contain glycine at position 124, whereas *Tv*ATP-PFK4 contains an alanine residue at the latter position (Fig. 2; see Fig. S2 in the supplemental material). The interchange of the glycine residue with alanine should not affect the interaction with the ATP molecule. The alanine residue possesses a small side chain, and it is unlikely that the residue creates steric hindrance to prevent binding of the ATP molecule. However, *Tv*ATP-PFK2 contains threonine and serine residues at positions 104 and 124, respectively. Therefore, the ability of *Tv*ATP-PFK2 to bind ATP is uncertain. The expected amino acid residues (D_{104} and K_{124}) are present in *Tv*PP_i-PFK1 and -3 to -6, whereas *Tv*PP_i-PFK2 and -7 contain glutamic acid and alanine residues at position 104, respectively (Fig. 2; see Fig. S2 in the supplemental material). Interestingly, scanning of the alignment of a broad range of sequences that were used for the phylogenetic analysis (Fig. 1) revealed the presence of paralogous genes with canonical G/D_{104} and G/K_{124} amino acid residues and with different residues at these positions in other parabasalids of clade T2 and in members of the Embryophyta, clade P. For example, serine residues at position 124 are also present in the putative ATP-PFKs of *Tritrichomonas foetus* and *Histomonas meleagridis* (see Fig. S2 in the supplemental material). Moreover, the *H. meleagridis* protein contains asparagine at position 104. These sequences, together with *Tv*ATP-PFK2 and -4, form the upper branch of clade T2 (Fig. 1). The unusual paralogues of Embryophyta PP_i-PFK-like sequences contain threonine/isoleucine and valine at positions 104 and 124, respectively (see Fig. S2 in the supplemental material), and they are grouped in the upper Embryophyta branch of clade P (Fig. 1). The functions of plant PP_i-PFK-like proteins are unknown (41).

Cellular localization of *Tv*ATP-PFK paralogues. The analysis of *Tv*ATP-PFK1 to -4 revealed an absence of sequence motifs thought to target precursors to hydrogenosomes. The *Tv*ATP-PFK sequences are colinear with their bacterial orthologues, lacking a predictable NTS and the cleavage site for the processing peptidase (Fig. 2). We found no internal motifs for subcellular targeting, and PSORT II predicted *Tv*ATP-PFKs to localize to the cytosol.

The subcellular localization of *Tv*ATP-PFK1 to -4 was investigated by the transient expression of C-terminally HA-tagged proteins in *T. vaginalis*. Immunofluorescence microscopy revealed that recombinant *Tv*ATP-PFK1, -2, and -4 colocalized with malic enzyme, the hydrogenosomal marker protein (Fig. 3; see Fig. S1 in the supplemental material), which suggested that these three proteins were transported into the hydrogenosomal matrix (we were unable to detect any expression of *Tv*ATP-PFK3 after several independent rounds of transfection). The topology of *Tv*ATP-PFK1 was further tested by protease protection assays. The treatment of isolated organelles with trypsin had no effect on the *Tv*ATP-PFK1 signal in the Western blot analysis, and the signal disappeared only in response to treatment with detergent (Fig. 3B). This finding indicates that *Tv*ATP-PFK is imported into *T. vaginalis* hydrogenosomes and is not associated with the organelle surface.

Although the bioinformatics analysis did not predict the presence of a cleavable NTS, we cannot exclude the possibility that a noncleavable "cryptic" NTS signal might direct *Tv*ATP-PFK1 to hydrogenosomes. Therefore, we expressed a truncated version of *Tv*ATP-PFK1 that lacked the first 16 amino acid residues (aa) (double the size of the known NTS in Fdx1). The truncated *Tv*ATP-PFK1 was delivered to the hydrogenosomes as its complete form (Fig. 3). This result confirmed that import of *Tv*ATP-PFK1 into hydrogenosomes is NTS independent. The expression of *Tv*PP_i-PFK revealed a cytosolic localization of the enzyme, as expected (Fig. 3).

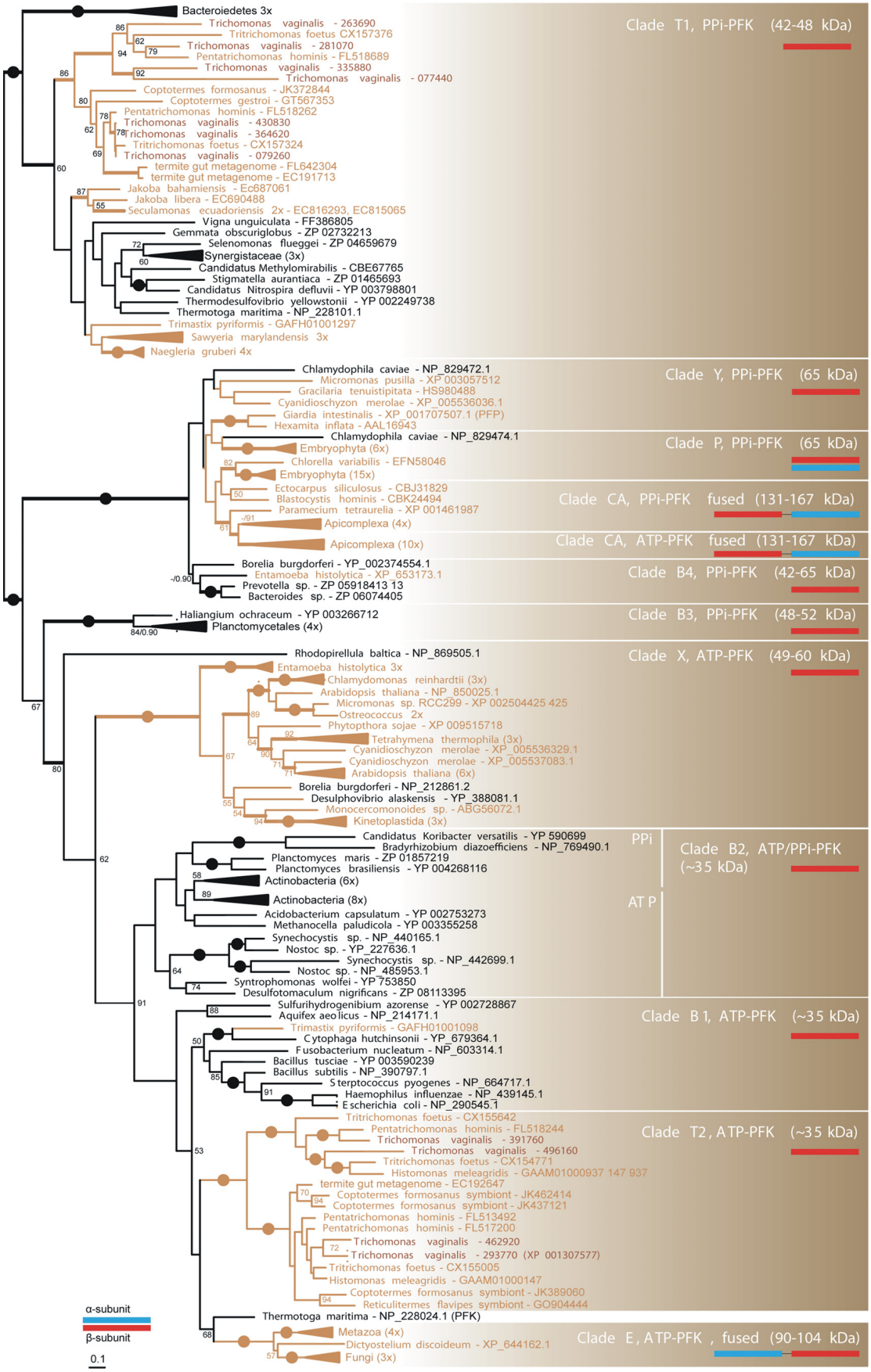
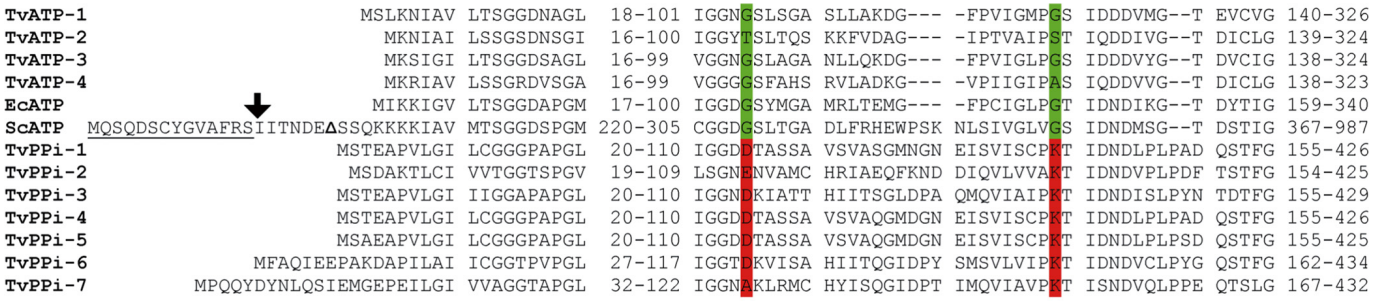


FIG 1 Phylogeny of ATP- and PP_i-dependent PFKs. Shown s a maximum-likelihood (ML) tree of PFK (191 taxa and 340 sites). The numbers at the nodes indicate bootstrap values (BV)/posterior probabilities (PP). Only BV and PP greater than 50% and 0.9, respectively, are shown. Branches with BV of >95% are marked by black circles, and branches with PP of >0.95 are marked by thick lines. Substrate specificity, molecular mass, and subunit composition for clades are indicated. The names of eukaryotes are in brown, and those of prokaryotes are in black.



Δ= 180AA out

FIG 2 Multiple-protein-sequence alignment of the N-terminal portions and ATP/PP_i binding domains of *T. vaginalis* ATP- and PP_i-dependent PFKs. *T. vaginalis* TrichDB accession numbers: *Tv*ATP-PFK1 to -4, TVAG_293770, TVAG_496160, TVAG_462920, and TVAG_391760; *Tv*PP_i-PFK1 to -7, TVAG_430830, TVAG_077440, TVAG_281070, TVAG_364620, TVAG_079260, TVAG_263690, and TVAG_335880. NCBI accession numbers: *E. coli*, NP_418351; *S. cerevisiae*, DAA08331. A PSORT II-predicted NTS in *Sc*ATP-PFK is underlined; the arrow indicates the predicted cleavage site. The amino acid residues that are required for the interaction with ATP are shaded in green, and the residues that are crucial for the interaction with a PP_i molecule are shaded in red.

Next we investigated PP_i- and ATP-dependent PFK activities in cellular fractions of *T. vaginalis*. Under anaerobic conditions, we detected specific PP_i-PFK activity of 0.4 to 0.9 μmol min^{-1} mg $protein^{-1}$ in the high-speed cytosolic fraction. Percoll-purified hydrogenosomes contained a low specific activity (~0.008 to 0.020 μmol min^{-1} mg $protein^{-1}$) of ATP-PFK. PP_i-PFK activity was not associated with the organelles. These results indicate that PP_i- and ATP-dependent PFK activities are present in *T. vaginalis* in two distinct cellular compartments, in the cytosol and in hydrogenosomes, respectively. However, the hydrogenosomal (ATP-dependent) activity is dwarfed by the well-characterized cytosolic PP_i-dependent activity, raising questions about the role of the ATP-dependent activity, if any, in core energy metabolism.

Expression of *Tv*ATP-PFK1 and ferredoxin 1 under the control of native promoters. The *T. vaginalis* SCSα promoter is a strong endogenous promoter for transient expression (42). The unexpected localization of *Tv*ATP-PFK1 when transiently expressed under the control of the SCSα promoter prompted us to test whether the promoter itself could influence the localization of the product. First, we tested SCSα versus the native promoter (NP) by determining the cellular localization of Fdx1, a model hydrogenosomal matrix protein that possesses a typical NTS (24), as well as an ITS (26). Full-length Fdx1 expressed under the control of the SCSα promoter localized to hydrogenosomes (Fig. 4). However, the expression of the same protein with a deleted NTS (ΔFdx1, with deletion of the first 8 amino acids, MLSQVCRF) resulted in a dual localization: the majority of the ΔFdx1 was accumulated in the cytosol, whereas a portion of the ΔFdx1 was targeted to the organelle. The matrix localization of ΔFdx1 was verified by a protease protection assay (Fig. 4). When the SCSα promoter was replaced with the native Fdx1 promoter (300 bp upstream of the coding sequence of the Fdx1 gene), the complete Fdx1 protein was imported into hydrogenosomes; however, Fdx1 with a deleted NTS remained in the cytosol (Fig. 4). It thus appears that the nature of the promoter that is used for protein expression may affect protein localization. In the case of Fdx1, the ITS is apparently not sufficient to deliver the protein into the organelles when the protein is expressed without NTS (ΔFdx1) under the control of the native promoter. Therefore, we also assessed the localization of the recombinant *Tv*ATP-PFK1 expressed in *T. vaginalis* under the control of its native *Tv*ATP-PFK1 promoter (Fig. 3). Immunofluorescence microscopy and Western blot analysis confirmed that under these conditions, *Tv*ATP-PFK1 was targeted into the hydrogenosomal matrix (Fig. 3).

***In vitro* import of *Tv*ATP-PFK1 into hydrogenosomes.** *Tv*ATP-PFK1 import into hydrogenosomes was investigated using an *in vitro* import system. *Tv*ATP-PFK1 labeled with ^{35}S was incubated with hydrogenosomes in import buffer supplemented with ATP and cytosolic extract for 0 to 60 min. After the incubation, the hydrogenosomes were treated with proteinase K to remove labeled proteins that were not imported into the organelles. These experiments revealed the time-dependent accumulation of radiolabeled *Tv*ATP-PFK1 within isolated hydrogenosomes (Fig. 5). Furthermore, we investigated whether ATP was necessary for import. When the import assay was supplemented with apyrase (20 U/ml), which converts ATP to AMP and pyrophosphate, no import of *Tv*ATP-PFK1 was observed (Fig. 5). This result indicates that NTS-independent import of *Tv*ATP-PFK1 requires ATP.

TvATP-PFK is recognized and imported into yeast mitochondria. It has been demonstrated that mitochondria and hydrogenosomes employ a common mode of NTS-dependent protein import (24). Thus, we were curious whether *Tv*ATP-PFK1 possesses an NTS-independent signal that is recognized by the protein import machinery of yeast mitochondria. We expressed *Tv*ATP-PFK1 with a C-terminal GFP tag in *S. cerevisiae*. Immunofluorescence microscopy showed that the GFP fusion protein colocalized with the mitochondrial marker MitoTracker (Fig. 6). A protease protection assay using isolated yeast mitochondria revealed that *Tv*ATP-PFK1 was imported into the organelle and excluded the possibility that the protein was associated with the mitochondrial surface. Cytochrome oxidase subunit VI was used as a control inner membrane protein. *Sc*ATP-PFK consists of an N-terminal extension of 200 aa, a catalytic domain of 359 aa, and a C-terminal regulatory domain (423 aa). When we expressed a full-length *Sc*ATP-PFK and a truncated form that lacked the C-terminal regulatory domain (1/2*Sc*PFK) in yeast, both recombinant proteins remained in the cytosol after translation (Fig. 6). The unique N-terminal extension of *Sc*PFK is rich in negatively charged amino acid residues (pI 4.67), which might prevent the targeting of the protein to mitochondria (43). Thus, we also ex-

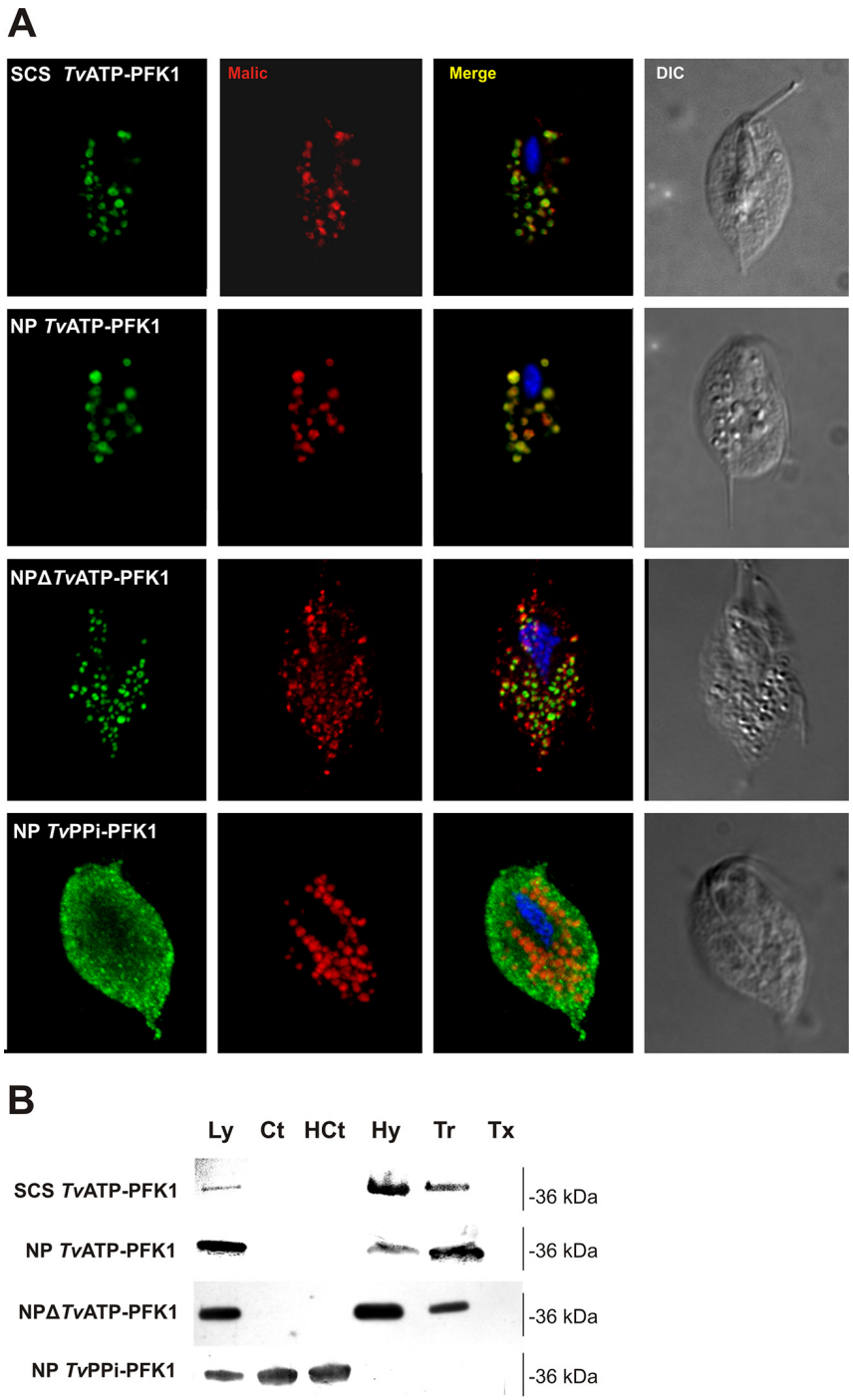


FIG 3 Cellular localization of ATP- and PP_i-dependent PFKs in *T. vaginalis*. (A) Immunofluorescence microscopy. Recombinant HA-tagged proteins were expressed in *T. vaginalis* cells and visualized using a monoclonal anti-HA antibody (green). *Tv*ATP-PFK1 and NP *Tv*ATP-PFK1 were expressed under the control of the strong SCSα promoter and the NP, respectively. NP Δ*Tv*ATP-PFK1 lacks 16 N-terminal amino acid residues. The hydrogenosomal marker protein malic enzyme was stained with a polyclonal rabbit antibody (red). The nucleus was stained using DAPI (4′,6-diamidino-2-phenylindole) (blue). DIC, differential interference contrast. (B) Protein protection assay. Hydrogenosomes were isolated from trichomonads expressing recombinant proteins with the C-terminal HA_2 tag and incubated with trypsin (Tr) or with trypsin and Triton X-100 (Tx). Samples were analyzed by immunoblotting using the monoclonal anti-HA tag antibody. Ly, total cell lysate; Ct, cytosol; HCt, high-speed cytosol; Hy, hydrogenosomes.

pressed the catalytic domain of *Sc*ATP-PFK, which is homologous to that of *Tv*ATP-PFK (ΔN1/2*Sc*PFK) alone. Interestingly, although some ΔN1/2*Sc*PFK signal was still observed in the cytosol, a significant portion was now also associated with the yeast mitochondrial membrane, as demonstrated by a protease protection assay (Fig. 6).

Collectively, these experiments show that *Tv*ATP-PFK1 possesses a targeting signal that is recognized by yeast mitochondria.

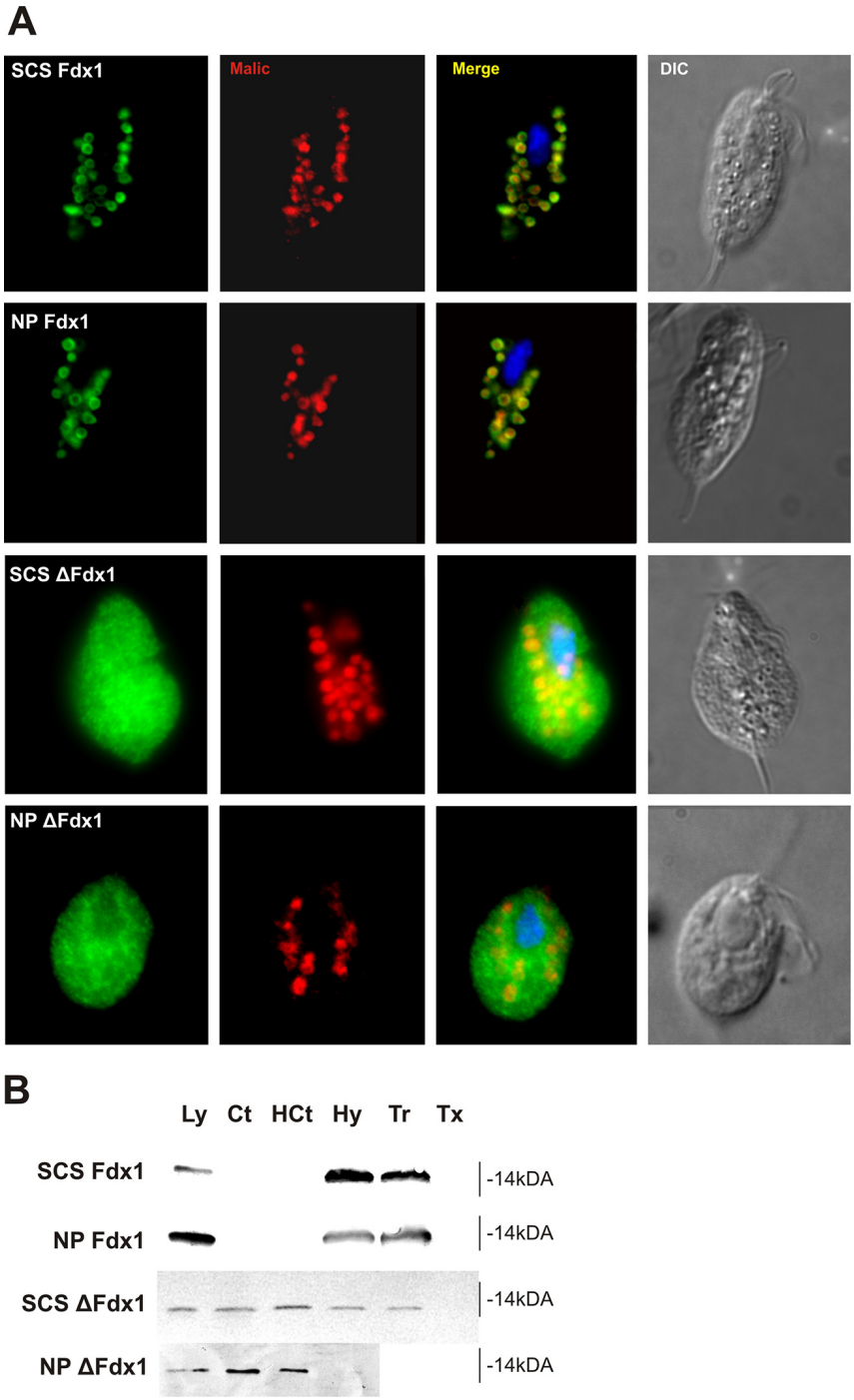


FIG 4 Effects of promoters on the cellular localization of ferredoxin. Fdx1 was used as a model protein with NTS-dependent targeting to test the effect of the SCSα promoter and the native promoter on Fdx1 localization. SCS Fdx1, Fdx1 (TVAG_003900) expressed under the control of the SCSα promoter; NP Fdx1, Fdx1 expressed under its native promoter; SCS ΔFdx1, Fdx1 with a deleted NTS that was expressed under the control of the SCSα promoter; NP ΔFdx1, ΔFdx1 expressed under the control of its native promoter. (A) Immunofluorescence microscopy. Recombinant HA-tagged proteins were expressed in *T. vaginalis* cells and visualized with monoclonal anti-HA antibody (green). The hydrogenosomal marker protein (malic enzyme) was detected using a polyclonal rabbit antibody (red). (B) Immunoblotting of subcellular fractions and protein protection assay. Ly, total cell lysate; Ct, cytosol; HCt, high-speed cytosol; Hy, hydrogenosomes; Tr, hydrogenosomes treated with trypsin; Tx, hydrogenosomal fraction treated with trypsin and Triton X-100.

The complete *Sc*ATP-PFK is retained in the cytosol, but the catalytic portion of *Sc*ATP-PFK displays mitochondrial membrane affinity.

Cellular localization of heterologous ATP-PFKs in *T. vaginalis*. We tested whether the hydrogenosomal protein import machinery can import heterologous ATP-PFKs. When we expressed complete *Sc*ATP-PFK in *T. vaginalis* under the control of the *Tv*ATP-PFK1 promoter, immunofluorescence microscopy revealed predominantly cytosolic localization of the protein, although the protein partially localized to hydrogenosomes (Fig. 7).

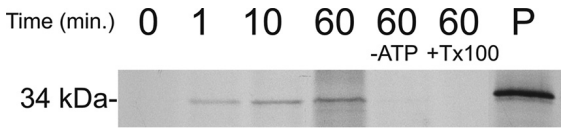


FIG 5 *In vitro* import of *Tv*ATP-PFK1 into hydrogenosomes. *In vitro*-synthesized ^{35}S-radiolabeled *Tv*ATP-PFK1 protein was incubated with isolated hydrogenosomes in import buffer at 25°C for 1, 10, and 60 min. At each time point, surface-associated proteins were degraded with proteinase K. Radiolabeled precursor was not imported in the absence of ATP (−ATP), depleted by addition of apyrase. A control for proteinase K activity was performed by the addition of Triton X-100 to the sample after 60 min of protein import (+Tx100). P, radiolabeled *Tv*ATP-PFK1 precursor protein. The samples were analyzed by SDS-PAGE and autoradiography.

The expression of 1/2*Sc*PFK revealed that the N-terminal half of *Sc*ATP-PFK was mainly associated with hydrogenosomes; however, the hydrogenosomal labeling was rather irregular in comparison to the labeling of malic enzyme, which was used as a control matrix protein. Western blot analysis of cellular fractions confirmed that both *Sc*ATP-PFK and 1/2*Sc*PFK were present in the cytosolic fractions (low- and high-speed cytosolic fractions). Parts of both proteins were also associated with the hydrogenosomal fractions; however, the signals disappeared after trypsin treatment. When we expressed only the catalytic part of the yeast enzyme lacking the negatively charged N-terminal sequence (ΔN1/2*Sc*PFK), a significant portion of the protein appeared inside the hydrogenosomes (Fig. 7). Next, we were interested in whether the targeting information is also present in short *E. coli* ATP-PFK orthologues that display 42% amino acid sequence identity with *Tv*ATP-PFKs. Thus, we expressed *Ec*ATP-PFK under the control of the *Tv*ATP-PFK1 promoter. Under these conditions, the *E. coli* protein was detected in the cytosol, and in part, it was associated with the hydrogenosomal surface (Fig. 7). However, when expressed under the SCSα promoter, a significant part of the protein was imported into the hydrogenosomes.

DISCUSSION

We investigated the cellular localization and NTS-independent import of *Tv*ATP-PFK into *T. vaginalis* hydrogenosomes. The parasite expresses both PP_i- and ATP-dependent enzymes, which are compartmentalized in the cytosol and hydrogenosomes, respectively. The classical PP_i-dependent activity of the parasite is about 50-fold higher than the newly characterized ATP-dependent activity, rendering the metabolic significance of the latter unclear. A phylogenetic analysis revealed that both types of PFKs are present across the parabasalids sampled so far. *Tv*ATP-PFK corresponds to a "short" ~35-kDa form of bacterial PFK that consists of only a catalytic domain, whereas the C-terminal regulatory domain typical of opisthokont ATP-PFKs is lacking. The targeting of *Tv*ATP-PFK1 to hydrogenosomes appears to be a

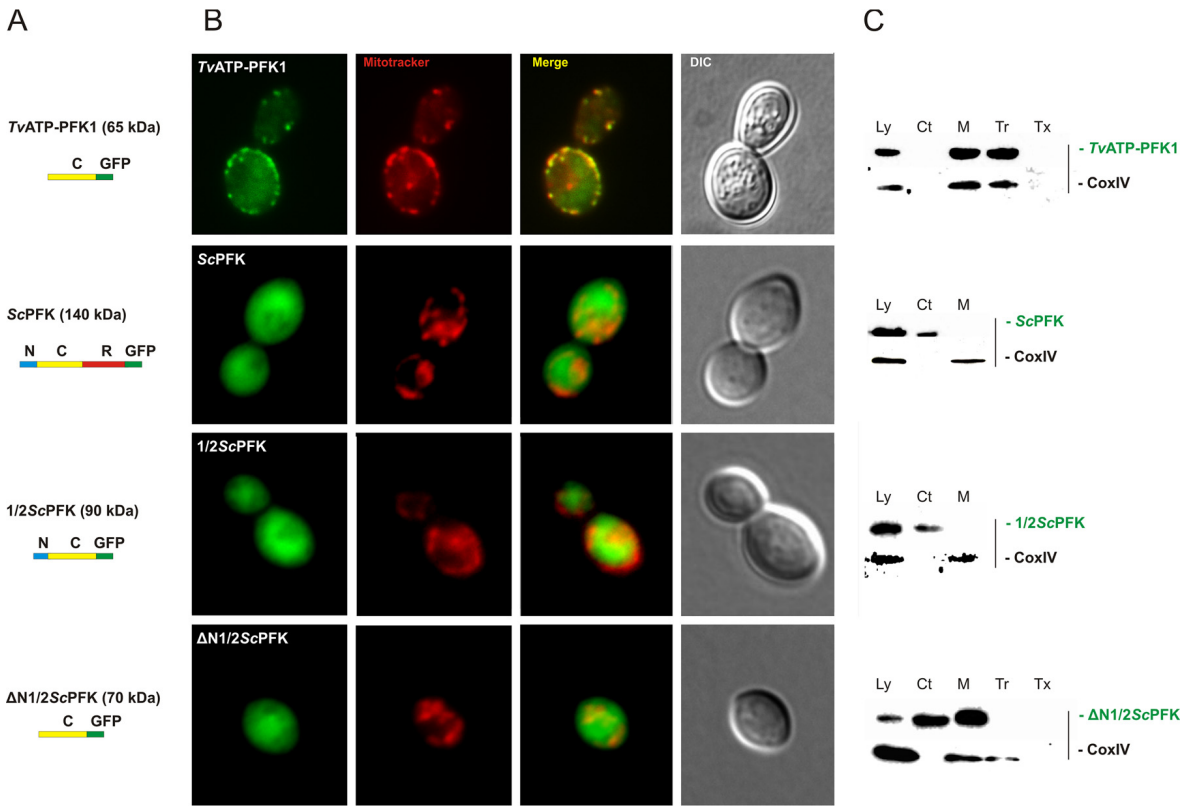


FIG 6 Cellular localization of *Tv*ATP-PFK1 and yeast ATP-PFK in *S. cerevisiae*. (A) Domain structure of the expressed proteins. N, N-terminal extension; C, catalytic domain; R, regulatory domain; GFP, green fluorescent protein tag. (B) Immunofluorescence microscopy. *Tv*ATP-PFK1 was expressed in yeasts with C-terminal GFP (green). Mitochondria were detected using MitoTracker dye (Invitrogen) (red). *Tv*ATP-PFK1, complete short *T. vaginalis* PFK; *Sc*PFK, complete long yeast PFK; 1/2*Sc*PFK, N-terminal extension (205 aa) and catalytic domain (359 aa) of *Sc*PFK; ΔN1/2*Sc*PFK, catalytic domain with deleted N-terminal extension. (C) Immunoblotting of subcellular fractions and protein protection assay. GFP-tagged proteins were detected using an anti-GFP antibody. Cytochrome oxidase subunit IV (CoxIV) was used as a mitochondrial marker, which was detected using a rabbit anti-CoxIV antibody. Ly, total cell lysate; Ct, cytosol; M, mitochondria; Tr, hydrogenosomes treated with trypsin; Tx, hydrogenosomal fraction treated with trypsin and Triton X-100.

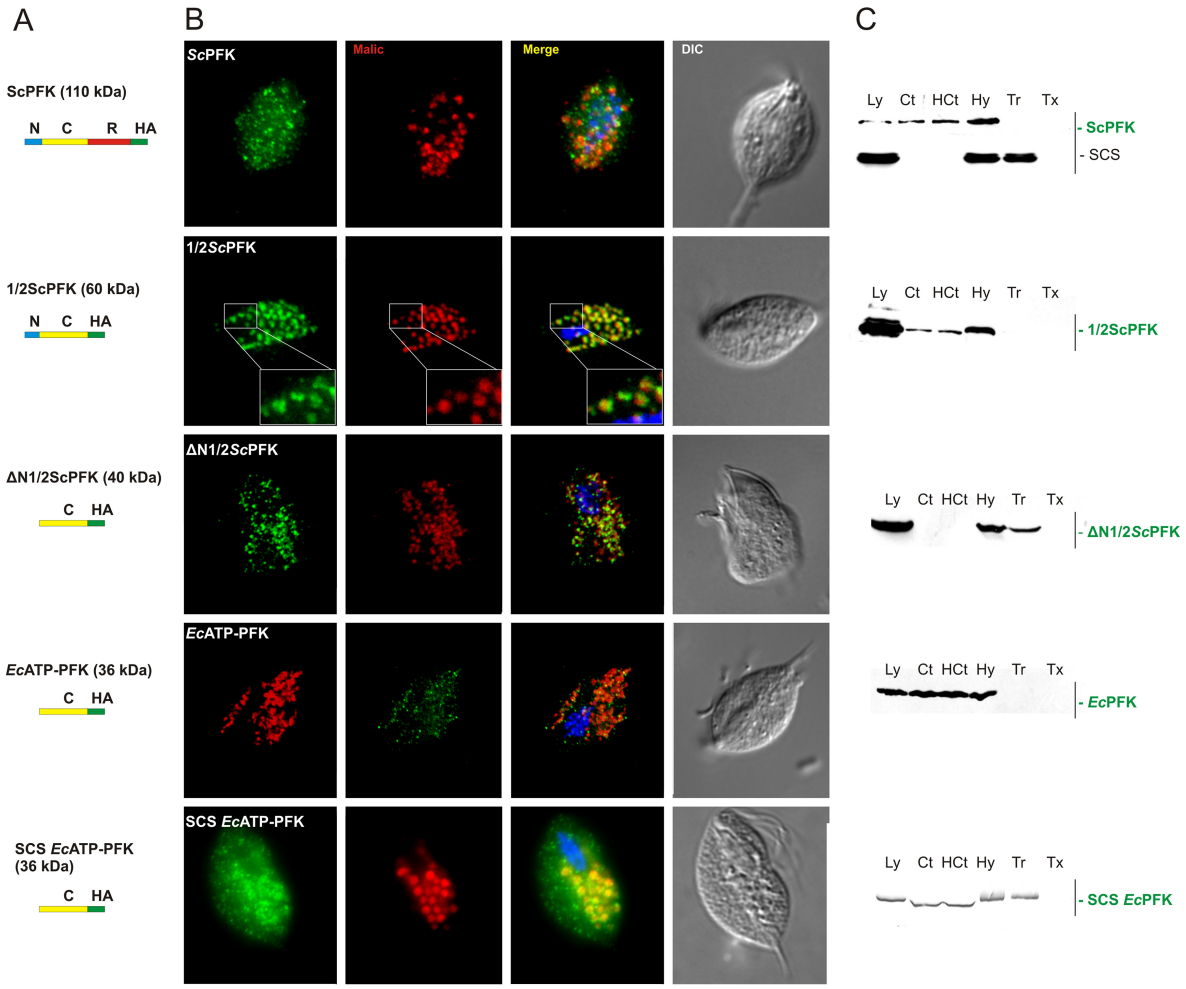


FIG 7 Cellular localization of *S. cerevisiae* *Sc*ATP-PFK and *Ec*ATP-PFK in *T. vaginalis*. (A) Domain structure of the expressed constructs. N, N-terminal extension; C, catalytic domain; R, regulatory domain; HA, hemagglutinin tag. (B) Recombinant HA-tagged proteins were expressed in *T. vaginalis* cells under the control of the *Tv*ATP-PFK1 promoter. SCS *Ec*ATP-PFK was expressed under the control of the SCSα promoter. HA-tagged proteins were visualized with mouse monoclonal anti-HA antibody (green). The hydrogenosomal marker protein (malic enzyme) was detected using a polyclonal rabbit antibody (red). (C) Immunoblotting of subcellular fractions and protein protection assay. Recombinant HA-tagged proteins were detected using monoclonal anti-HA antibody. The hydrogenosomal marker protein SCSα was detected using a rabbit polyclonal antibody. Ly, total cell lysate; Ct, cytosol; HCt, high-speed cytosol; Hy, hydrogenosomes; Tr, hydrogenosomes treated with trypsin; Tx, hydrogenosomal fraction treated with trypsin and Triton X-100.

highly specific and ATP-dependent process, even though the protein is not predicted to possess a cleavable NTS, which is typical of hydrogenosomal matrix proteins (44, 45).

The replacement of ATP with PP_i as a phosphate donor in the phosphorylation of fructose-6-phosphate allows an increased glycolytic ATP yield (3), conceivably a significant feature for a fermenting organism. Examples of organisms that express both PP_i-PFK and ATP-PFK are rare. The actinomycete *Amycolatopsis methanolica* possesses both genes, but their expression depends strictly on the carbon source (46). *Entamoeba histolytica* possesses two genes for PP_i-PFK orthologues; however, one of the gene products has been shown to utilize ATP instead of PP_i, and it has been suggested that the two enzymes might be expressed during different life stages (37). In plants, PP_i-PFK and ATP-PFK are both cytosolic enzymes with reciprocal expression responding to environmental perturbations (47). Whereas the expression of PP_i-PFK is upregulated by anoxia or orthophosphate deficiency, ATP-PFK is downregulated under such conditions. The spatial separation in *T. vaginalis* of PP_i-PFK and ATP-PFK to the cytosol and hydrogenosomes, respectively, could be an alternative solution to avoid interference between the two enzymes.

Specific targeting of *Tv*ATP-PFK to the organelle was demonstrated *in vivo* by episomal expression of tagged *Tv*ATP-PFK1 under SCSα and its native promoters, as well as the *in vitro* import of radiolabeled protein into isolated hydrogenosomes. Through the HA-tagged *Tv*ATP-PFK1, products of four paralogous *Tv*ATP-PFK genes were immunoprecipitated from isolated hydrogenosomes and identified by mass spectrometry. Earlier proteomic studies suggested association of the glycolytic pathway, including *Tv*ATP-PFK, with the hydrogenosome (13, 15), which raises the question of whether glycolytic enzymes form functional protein complexes on the hydrogenosomal outer membrane, as has been shown for mitochondria. For example, in *Arabidopsis thaliana*, 5 to 10% of each glycolytic enzyme is associated with the outer mitochondrial surface. Mammalian and fish heart mitochondria bind hexokinase and ATP-PFK (48), which has been discussed in the context of an increased glycolytic rate under hypoxic conditions (49). However, in *T. vaginalis*, expression of seven glycolytic

enzymes, including PP_i-PFK, showed exclusively cytosolic localization of these proteins (15, 26). Moreover, available cell fractionation studies of glyceraldehyde-3-phosphate dehydrogenase (50) and PP_i-PFK (this study) indicated that the corresponding activities are not associated with the organelle. These data do not support the formation of functional glycolytic complexes at the hydrogenosomal membrane and make the interpretation of previous proteomic analysis problematic, although systematic studies of glycolytic enzyme activities in cellular fractions of *T. vaginalis* are currently lacking. The localization of *Tv*ATP-PFK in the hydrogenosomal matrix, as shown in this study, is new for trichomonads.

Organellar forms of ATP-PFK have been found in glycosomes (51) and chloroplasts (52) thus far, where ATP-PFK operates within a known biochemical context. Kinetoplastids catalyze the "upper" six glycolytic steps in glycosomes, exporting 3-phosphoglycerate to the cytosol. Microalgae, such as *Chlamydomonas reinhardtii*, possess four glycolytic enzymes that convert glucose to glyceraldehyde-3-phosphate in chloroplasts, whereas the rest of glycolysis is localized in the cytosol (7). The most complicated glycolytic network has been found in diatoms, such as *P. tricornutum*, in which the complete set of glycolytic enzymes is present in the cytosol; nine glycolytic enzymes, including ATP-PFK, catalyze the conversion of glucose-1-phosphate to pyruvate in the chloroplast, and five glycolytic enzymes convert glyceraldehyde-3-phospate to pyruvate in the mitochondrion (8). In these organisms, the specific targeting of various glycolytic enzymes into the organelles is mediated by NTS (mitochondria), peroxisomal targeting signals (glycosomes), and plastid targeting signal (chloroplasts). The organellar *Tv*ATP-PFK found in *T. vaginalis* is unique with respect to three features: (i) it is a single glycolytic enzyme that is compartmentalized without apparent distal and proximal partners in the pathway, (ii) it is the only PFK that was observed to be imported into mitochondrion-related organelles, and (iii) the import into hydrogenosomes is mediated by ITS. The overall low hydrogenosomal ATP-PFK activity (approximately 2% of the PP_i-dependent activity), together with the lack of organellar glycolytic partners, raises questions regarding the metabolic role of *Tv*ATP-PFK and whether another function, unrelated to glycolysis, might be a possible alternative. Various moonlighting functions have been suggested for ATP-PFK in eukaryotes and bacteria, such as participation in the microautophagy of peroxisomes (53), RNA processing and degradation (54), and surface binding of plasminogen (55) and mannan (56). In our view, however, none of these functions currently appear likely for *Tv*ATP-PFK.

Heterologous expression of *Tv*ATP-PFK1 in *S. cerevisiae* revealed that the trichomonad enzyme is imported into yeast mitochondria, in addition to hydrogenosomes. This result indicates that *Tv*ATP-PFK1 possesses a targeting signal that is recognized by the hydrogenosomal, as well as the mitochondrial, import machinery. From an evolutionary perspective, these data suggest that the "short" ancient ATP-PFK might be predisposed to being recognized and imported into mitochondria, which might be a relict from the early phases of mitochondrial evolution. If so, the evolving eukaryotic cell had not only to develop a mechanism for retargeting nuclear-encoded proteins to mitochondria, but also to prevent the organellar translocation of some proteins, such as ATP-PFK, that are components of cytosolic pathways. Interestingly, unlike short bacterial ATP-PFK, eukaryotes frequently possess structurally modified long ATP-PFK that consists of catalytic and regulatory domains. In addition, the ATP-PFK of yeast and other fungi is equipped with a negatively charged N-terminal extension that may interfere with organellar import. Indeed, when we expressed the catalytic domain of *Sc*ATP-PFK with the N-terminal extension (1/2*Sc*PFK) in *T. vaginalis*, the protein was not delivered to the hydrogenosomal matrix, indicating that the extension prevents translocation. However, the hydrogenosomal import machinery was able to recognize and partially import truncated yeast *Sc*ATP-PFK, consisting of only the catalytic domain (ΔN1/2ScPFK), and the short proteobacterial *Ec*ATP-PFK, which are both homologous to *Tv*ATP-PFK. These results are consistent with the idea that ancient ATP-PFKs were predisposed to target the organelle. They also support previous analysis of proteins encoded by *E. coli* that predicted the presence of mitochondrial targeting information in about 5% of bacterial proteins (57).

The cell localization studies performed need to be interpreted with caution. Import of *Ec*ATP-PFK was observed when the gene was expressed under a strong SCSα promoter, while expression under the *Tv*ATP-PFK1 promoter resulted in partial association of *Tv*ATP-PFK1 with the outer hydrogenosomal membrane. Similarly, we observed promoter-dependent variation in the cell localization of Fdx, which possesses both NTS and ITS. Although we cannot exclude the possibility that hydrogenosomal localization of proteins expressed under strong promoters reflects protein mislocalization, it has been shown previously that six glycolytic enzymes expressed under the SCSα promoter remained exclusively in the cytosol, as expected, which argues against protein mislocalization (26). Therefore, it is more likely that, in addition to ITS, a suitable level of protein is required for protein translocation into the hydrogenosomes, while proteins without ITS are not targeted to the organelle regardless of the protein level. Importantly, expression of ΔN1/2ScPFK under *Tv*ATP-PFK1 was sufficient for its partial translocation into hydrogenosomes.

In conclusion, we identified ATP-PFK in *T. vaginalis* that is efficiently delivered into mitochondria and hydrogenosomes via NTS-independent mechanisms. Although NTS-independent targeting of membrane proteins is well documented, little is known about NTS-independent targeting of soluble proteins and the characters of multiple inner signals that are embedded within the protein structure (23, 58). The import of ATP-PFK into *T. vaginalis* hydrogenosomes can be used to investigate the molecular mechanisms that facilitate NTS-independent targeting and underpins the importance of internal targeting motifs that, in the case of PFK, are recognized in species spanning different eukaryotic supergroups. Intriguingly, the function of *Tv*ATP-PFK in *T. vaginalis* hydrogenosomes remains mysterious.

ACKNOWLEDGMENTS

This work was supported by the Czech Grant Foundation (13-09208J); the Biomedicine Centre of the Academy of Sciences and Charles University, Prague, Czech Republic (CZ.1.05/1.1.00/02.0109), from the European Regional Development Fund, project no. CZ.1.07/2.3.00/30.0061 (OPVK); and DFG grants to S.B.G. (GO1825/3–1). S.B.G. was additionally supported by a DFG grant to William F. Martin (MA1426/19-1).

REFERENCES

1. **Lane N, Martin W.** 2010. The energetics of genome complexity. Nature **467:**929–934. http://dx.doi.org/10.1038/nature09486.
2. **Hannaert V, Brinkmann H, Nowitzki U, Lee JA, Albert MA, Sensen CW, Gaasterland T, Muller M, Michels P, Martin W.** 2000. Enolase from *Trypanosoma brucei*, from the amitochondriate protist *Mastiga-*

moeba balamuthi, and from the chloroplast and cytosol of *Euglena gracilis*: pieces in the evolutionary puzzle of the eukaryotic glycolytic pathway. Mol Biol Evol **17:**989–1000. http://dx.doi.org/10.1093/oxfordjournals.molbev.a026395.

3. **Müller M, Mentel M, van Hellemond JJ, Henze K, Woehle C, Gould SB, Yu RY, van der Giezen M, Tielens AG, Martin WF.** 2012. Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. Microbiol Mol Biol Rev **76:**444–495. http://dx.doi.org/10.1128/MMBR.05024-11.
4. **Embley TM, Martin W.** 2006. Eukaryotic evolution, changes and challenges. Nature **440:**623–630. http://dx.doi.org/10.1038/nature04546.
5. **Gualdron-Lopez M, Brennand A, Hannaert V, Quinones W, Caceres AJ, Bringaud F, Concepcion JL, Michels PA.** 2012. When, how and why glycolysis became compartmentalised in the Kinetoplastea. A new look at an ancient organelle. Int J Parasitol **42:**1–20. http://dx.doi.org/10.1016/j.ijpara.2011.10.007.
6. **Schnarrenberger C, Pelzer-Reith B, Yatsuki H, Freund S, Jacobshagen S, Hori K.** 1994. Expression and sequence of the only detectable aldolase in *Chlamydomonas reinhardtii*. Arch Biochem Biophys **313:**173–178. http://dx.doi.org/10.1006/abbi.1994.1374.
7. **Johnson X, Alric J.** 2013. Central carbon metabolism and electron transport in *Chlamydomonas reinhardtii*: metabolic constraints for carbon partitioning between oil and starch. Eukaryot Cell **12:**776–793. http://dx.doi.org/10.1128/EC.00318-12.
8. **Kroth PG, Chiovitti A, Gruber A, Martin-Jezequel V, Mock T, Parker MS, Stanley MS, Kaplan A, Caron L, Weber T, Maheswari U, Armbrust EV, Bowler C.** 2008. A model for carbohydrate metabolism in the diatom *Phaeodactylum tricornutum* deduced from comparative whole genome analysis. PLoS One **3:**e1426. http://dx.doi.org/10.1371/journal.pone.0001426.
9. **Nakayama T, Ishida K, Archibald JM.** 2012. Broad distribution of TPI-GAPDH fusion proteins among eukaryotes: evidence for glycolytic reactions in the mitochondrion? PLoS One **7:**e52340. http://dx.doi.org/10.1371/journal.pone.0052340.
10. **Mertens E, Ladror US, Lee JA, Miretsky A, Morris A, Rozario C, Kemp RG, Müller M.** 1998. The pyrophosphate-dependent phosphofructokinase of the protist, Trichomonas vaginalis, and the evolutionary relationships of protist phosphofructokinases. J Mol Evol **47:**739–750. http://dx.doi.org/10.1007/PL00006433.
11. **Winkler C, Delvos B, Martin W, Henze K.** 2007. Purification, microsequencing and cloning of spinach ATP-dependent phosphofructokinase link sequence and function for the plant enzyme. FEBS J **274:**429–438. http://dx.doi.org/10.1111/j.1742-4658.2006.05590.x.
12. **Carlton JM, Hirt RP, Silva JC, Delcher AL, Schatz M, Zhao Q, Wortman JR, Bidwell SL, Alsmark UC, Besteiro S, Sicheritz-Ponten T, Noel CJ, Dacks JB, Foster PG, Simillion C, Van de PY, Miranda-Saavedra D, Barton GJ, Westrop GD, Muller S, Dessi D, Fiori PL, Ren Q, Paulsen I, Zhang H, Bastida-Corcuera FD, Simoes-Barbosa A, Brown MT, Hayes RD, Mukherjee M, Okumura CY, Schneider R, Smith AJ, Vanacova S, Villalvazo M, Haas BJ, Pertea M, Feldblyum TV, Utterback TR, Shu CL, Osoegawa K, de Jong PJ, Hrdy I, Horvathova L, Zubacova Z, Dolezal P, Malik SB, Logsdon JM, Jr, Henze K, Gupta A, Wang CC, Dunne RL, Upcroft JA, Upcroft P, White O, Salzberg SL, Tang P, Chiu CH, Lee YS, Embley TM, Coombs GH, Mottram JC, Tachezy J, Fraser-Liggett CM, Johnson PJ.** 2007. Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. Science **315:**207–212. http://dx.doi.org/10.1126/science.1132894.
13. **Schneider RE, Brown MT, Shiflett AM, Dyall SD, Hayes RD, Xie Y, Loo JA, Johnson PJ.** 2011. The *Trichomonas vaginalis* hydrogenosome proteome is highly reduced relative to mitochondria, yet complex compared with mitosomes. Int J Parasitol **41:**1421–1434. http://dx.doi.org/10.1016/j.ijpara.2011.10.001.
14. **Henze K.** 2007. The proteome of *T. vaginalis* hydrogenosomes, p 163–178. *In* Tachezy J (ed), Hydrogenosomes and mitosomes: mitochondria of anaerobic eukaryotes. Springer, Berlin, Germany.
15. **Rada P, Dolezal P, Jedelsky PL, Bursac D, Perry AJ, Sedinova M, Smiskova K, Novotny M, Beltran NC, Hrdy I, Lithgow T, Tachezy J.** 2011. The core components of organelle biogenesis and membrane transport in the hydrogenosomes of *Trichomonas vaginalis*. PLoS One **6:**e24428. http://dx.doi.org/10.1371/journal.pone.0024428.
16. **Liapounova NA, Hampl V, Gordon PM, Sensen CW, Gedamu L, Dacks JB.** 2006. Reconstructing the mosaic glycolytic pathway of the anaerobic eukaryote Monocercomonoides. Eukaryot Cell **5:**2138–2146. http://dx.doi.org/10.1128/EC.00258-06.
17. **Bapteste E, Moreira D, Philippe H.** 2003. Rampant horizontal gene transfer and phospho-donor change in the evolution of the phosphofructokinase. Gene **318:**185–191. http://dx.doi.org/10.1016/S0378-1119(03)00797-2.
18. **Shirakihara Y, Evans PR.** 1988. Crystal structure of the complex of phosphofructokinase from *Escherichia coli* with its reaction products. J Mol Biol **204:**973–994. http://dx.doi.org/10.1016/0022-2836(88)90056-3.
19. **Poorman RA, Randolph A, Kemp RG, Heinrikson RL.** 1984. Evolution of phosphofructokinase: gene duplication and creation of new effector sites. Nature **309:**467–469. http://dx.doi.org/10.1038/309467a0.
20. **Mony BM, Mehta M, Jarori GK, Sharma S.** 2009. Plant-like phosphofructokinase from *Plasmodium falciparum* belongs to a novel class of ATP-dependent enzymes. Int J Parasitol **39:**1441–1453. http://dx.doi.org/10.1016/j.ijpara.2009.05.011.
21. **Vogtle FN, Wortelkamp S, Zahedi RP, Becker D, Leidhold C, Gevaert K, Kellermann J, Voos W, Sickmann A, Pfanner N, Meisinger C.** 2009. Global analysis of the mitochondrial N-proteome identifies a processing peptidase critical for protein stability. Cell **139:**428–439. http://dx.doi.org/10.1016/j.cell.2009.07.045.
22. **Gakh O, Cavadini P, Isaya G.** 2002. Mitochondrial processing peptidases. Biochim Biophys Acta **1592:**63–77. http://dx.doi.org/10.1016/S0167-4889(02)00265-3.
23. **Chacinska A, Koehler CM, Milenkovic D, Lithgow T, Pfanner N.** 2009. Importing mitochondrial proteins: machineries and mechanisms. Cell **138:**628–644. http://dx.doi.org/10.1016/j.cell.2009.08.005.
24. **Bradley PJ, Lahti CJ, Plumper E, Johnson PJ.** 1997. Targeting and translocation of proteins into the hydrogenosome of the protist Trichomonas: similarities with mitochondrial protein import. EMBO J **16:**3484–3493. http://dx.doi.org/10.1093/emboj/16.12.3484.
25. **Dolezal P, Smid O, Rada P, Zubacova Z, Bursac D, Sutak R, Nebesarova J, Lithgow T, Tachezy J.** 2005. Giardia mitosomes and trichomonad hydrogenosomes share a common mode of protein targeting. Proc Natl Acad Sci U S A **102:**10924–10929. http://dx.doi.org/10.1073/pnas.0500349102.
26. **Zimorski V, Major P, Hoffmann K, Bras XP, Martin WF, Gould SB.** 2013. The N-terminal sequences of four major hydrogenosomal proteins are not essential for import into hydrogenosomes of *Trichomonas vaginalis*. J Eukaryot Microbiol **60:**89–97. http://dx.doi.org/10.1111/jeu.12012.
27. **Mentel M, Zimorski V, Haferkamp P, Martin W, Henze K.** 2008. Protein import into hydrogenosomes of *Trichomonas vaginalis* involves both N-terminal and internal targeting signals: a case study of thioredoxin reductases. Eukaryot Cell **7:**1750–1757. http://dx.doi.org/10.1128/EC.00206-08.
28. **Katoh K, Kuma K, Toh H, Miyata T.** 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res **33:**511–518. http://dx.doi.org/10.1093/nar/gki198.
29. **Hall TA.** 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser **41:**95–98.
30. **Stamatakis A.** 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics **22:**2688–2690. http://dx.doi.org/10.1093/bioinformatics/btl446.
31. **Stamatakis A, Hoover P, Rougemont J.** 2008. A rapid bootstrap algorithm for the RAxML Web servers. Syst Biol **57:**758–771. http://dx.doi.org/10.1080/10635150802429642.
32. **Lartillot N, Rodrigue N, Stubbs D, Richer J.** 2013. PhyloBayes MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. Syst Biol **62:**611–615. http://dx.doi.org/10.1093/sysbio/syt022.
33. **Hrdy I, Hirt RP, Dolezal P, Bardonova L, Foster PG, Tachezy J, Embley TM.** 2004. Trichomonas hydrogenosomes contain the NADH dehydrogenase module of mitochondrial complex I. Nature **432:**618–622. http://dx.doi.org/10.1038/nature03149.
34. **Niedenthal RK, Riles L, Johnston M, Hegemann JH.** 1996. Green fluorescent protein as a marker for gene expression and subcellular localization in budding yeast. Yeast **12:**773–786.
35. **Sutak R, Dolezal P, Fiumera HL, Hrdy I, Dancis A, Delgadillo-Correa MG, Johnson PJ, Müller M, Tachezy J.** 2004. Mitochondrial-type assembly of FeS centers in the hydrogenosomes of the amitochondriate eukaryote *Trichomonas vaginalis*. Proc Natl Acad Sci U S A **101:**10368–10373. http://dx.doi.org/10.1073/pnas.0401319101.
36. **Drmota T, Proost P, Van Ranst M, Weyda F, Kulda J, Tachezy J.** 1996. Iron-ascorbate cleavable malic enzyme from hydrogenosomes of

Trichomonas vaginalis: purification and characterization. Mol Biochem Parasitol **83:**221–234. http://dx.doi.org/10.1016/S0166-6851(96)02777-6.
37. **Chi AS, Deng Z, Albach RA, Kemp RG.** 2001. The two phosphofructokinase gene products of *Entamoeba histolytica*. J Biol Chem **276:**19974–19981. http://dx.doi.org/10.1074/jbc.M011584200.
38. **Mertens E, Van Schaftingen E, Müller M.** 1989. Presence of a fructose-2,6-bisphosphate-insensitive pyrophosphate: fructose-6-phosphate phosphotransferase in the anaerobic protozoa *Tritrichomonas foetus*, *Trichomonas vaginalis* and *Isotricha prostoma*. Mol Biochem Parasitol **37:**183–190. http://dx.doi.org/10.1016/0166-6851(89)90150-3.
39. **Gregg C, Kyryakov P, Titorenko VI.** 2009. Purification of mitochondria from yeast cells. J Vis Exp **30:**1417. http://dx.doi.org/10.3791/1417.
40. **Michels PA, Chevalier N, Opperdoes FR, Rider MH, Rigden DJ.** 1997. The glycosomal ATP-dependent phosphofructokinase of *Trypanosoma brucei* must have evolved from an ancestral pyrophosphate-dependent enzyme. Eur J Biochem **250:**698–704. http://dx.doi.org/10.1111/j.1432-1033.1997.00698.x.
41. **Mustroph A, Sonnewald U, Biemelt S.** 2007. Characterisation of the ATP-dependent phosphofructokinase gene family from *Arabidopsis thaliana*. FEBS Lett **581:**2401–2410. http://dx.doi.org/10.1016/j.febslet.2007.04.060.
42. **Delgadillo MG, Liston DR, Niazi K, Johnson PJ.** 1997. Transient and selectable transformation of the parasitic protist *Trichomonas vaginalis*. Proc Natl Acad Sci U S A **94:**4716–4720. http://dx.doi.org/10.1073/pnas.94.9.4716.
43. **Gerber J, Neumann K, Prohl C, Muhlenhoff U, Lill R.** 2004. The yeast scaffold proteins Isu1p and Isu2p are required inside mitochondria for maturation of cytosolic Fe/S proteins. Mol Cell Biol **24:**4848–4857. http://dx.doi.org/10.1128/MCB.24.11.4848-4857.2004.
44. **Smid O, Matuskova A, Harris SR, Kucera T, Novotny M, Horvathova L, Hrdy I, Kutejova E, Hirt RP, Embley TM, Janata J, Tachezy J.** 2008. Reductive evolution of the mitochondrial processing peptidases of the unicellular parasites *Trichomonas vaginalis* and *Giardia intestinalis*. PLoS Pathog **4:**e1000243. http://dx.doi.org/10.1371/journal.ppat.1000243.
45. **Burstein D, Gould SB, Zimorski V, Kloesges T, Kiosse F, Major P, Martin WF, Pupko T, Dagan T.** 2012. A machine learning approach to identify hydrogenosomal proteins in *Trichomonas vaginalis*. Eukaryot Cell **11:**217–228. http://dx.doi.org/10.1128/EC.05225-11.
46. **Alves AM, Euverink GJ, Santos H, Dijkhuizen L.** 2001. Different physiological roles of ATP- and PP(i)-dependent phosphofructokinase isoenzymes in the methylotrophic actinomycete *Amycolatopsis methanolica*. J Bacteriol **183:**7231–7240. http://dx.doi.org/10.1128/JB.183.24.7231-7240.2001.
47. **Plaxton WC, Tran HT.** 2011. Metabolic adaptations of phosphate-starved plants. Plant Physiol **156:**1006–1015. http://dx.doi.org/10.1104/pp.111.175281.
48. **Treberg JR, MacCormack TJ, Lewis JM, Almeida-Val VM, Val AL, Driedzic WR.** 2007. Intracellular glucose and binding of hexokinase and phosphofructokinase to particulate fractions increase under hypoxia in heart of the Amazonian armored catfish (*Liposarcus pardalis*). Physiol Biochem Zool **80:**542–550. http://dx.doi.org/10.1086/520129.
49. **Graham JW, Williams TC, Morgan M, Fernie AR, Ratcliffe RG, Sweetlove LJ.** 2007. Glycolytic enzymes associate dynamically with mitochondria in response to respiratory demand and support substrate channeling. Plant Cell **19:**3723–3738. http://dx.doi.org/10.1105/tpc.107.053371.
50. **Markos A, Miretsky A, Müller M.** 1993. A glyceraldehyde-3-phosphate dehydrogenase with eubacterial features in the amitochondriate eukaryote, *Trichomonas vaginalis*. J Mol Evol **37:**631–643. http://dx.doi.org/10.1007/BF00182749.
51. **Opperdoes FR, Borst P.** 1977. Localization of nine glycolytic enzymes in a microbody-like organelle in *Trypanosoma brucei*: the glycosome. FEBS Lett **80:**360–364. http://dx.doi.org/10.1016/0014-5793(77)80476-6.
52. **Plaxton WC.** 1996. The organization and regulation of plant glycolysis. Annu Rev Plant Physiol Plant Mol Biol **47:**185–214. http://dx.doi.org/10.1146/annurev.arplant.47.1.185.
53. **Gancedo C, Flores CL, Gancedo JM.** 2014. Evolution of moonlighting proteins: insight from yeasts. Biochem Soc Trans **42:**1715–1719. http://dx.doi.org/10.1042/BST20140199.
54. **Commichau FM, Rothe FM, Herzberg C, Wagner E, Hellwig D, Lehnik-Habrink M, Hammer E, Volker U, Stulke J.** 2009. Novel activities of glycolytic enzymes in *Bacillus subtilis*: interactions with essential proteins involved in mRNA processing. Mol Cell Proteomics **8:**1350–1360. http://dx.doi.org/10.1074/mcp.M800546-MCP200.
55. **Kinnby B, Booth NA, Svensater G.** 2008. Plasminogen binding by oral streptococci from dental plaque and inflammatory lesions. Microbiology **154:**924–931. http://dx.doi.org/10.1099/mic.0.2007/013235-0.
56. **Katakura Y, Sano R, Hashimoto T, Ninomiya K, Shioya S.** 2010. Lactic acid bacteria display on the cell surface cytosolic proteins that recognize yeast mannan. Appl Microbiol Biotechnol **86:**319–326. http://dx.doi.org/10.1007/s00253-009-2295-y.
57. **Lucattini R, Likic VA, Lithgow T.** 2004. Bacterial proteins predisposed for targeting to mitochondria. Mol Biol Evol **21:**652–658. http://dx.doi.org/10.1093/molbev/msh058.
58. **Garg S, Stolting J, Zimorski V, Rada P, Tachezy J, Martin WF, Gould SB.** 2015. Conservation of transit peptide-independent protein import into the mitochondrial and hydrogenosomal matrix. Genome Biol Evol **7:**2716–2726. http://dx.doi.org/10.1093/gbe/evv175.

Publication V

The role of charge in protein targeting evolution

Sriram G. Garg[1] and Sven B. Gould[1]

1 - Institute for Molecular Evolution, Heinrich-Heine Universität Düsseldorf, Universitätsstraße 1, Düsseldorf 40225, Germany

The presented manuscript was peer-reviewed and published in **Trends in Cell Biology** with an impact factor of **11.532** on 11th August 2016

Contribution as **first author**

80%

Conceived significant portions of the hypothesis including mining the literature for relevant citations. Wrote the text and discussion. Designed and illustrated Figures 1, 2 and generated the data represented in Figure 2

CellPress

Opinion

The Role of Charge in Protein Targeting Evolution

Sriram G. Garg[1] and Sven B. Gould[1,*]

Two eukaryotic compartments are of endosymbiotic origin, the mitochondrion and plastid. These organelles need to import hundreds of proteins from the cytosol. The import machineries of both are of independent origin, but function in a similar fashion and recognize N-terminal targeting sequences that also share similarities. Targeting, however, is generally specific, even though plastid targeting evolved in the presence of established mitochondrial targeting. Here we review current advances on protein import into mitochondria and plastids from diverse eukaryotic lineages and highlight the impact of charged amino acids in targeting. Their presence or absence alone can determine localization, and comparisons across diverse eukaryotes, and their different types of mitochondria and plastids, uncover unexplored avenues of protein import research.

Eukaryotic Compartmentalization Engenders Protein Targeting

The most important morphological difference between the prokaryotic and eukaryotic cytosol is compartmentalization. One can distinguish between two different types of eukaryotic compartments: those that originate from **endosymbiosis** (see Glossary), the mitochondrion and the plastid [1,2], and all others, namely the endoplasmic reticulum, the nucleus, the Golgi apparatus, the lysosome, multivesicular bodies, peroxisomes, vacuoles, endosomes, and the vesicles that connect them to each other, which together constitute the compartments of the **endomembrane system** [3,4]. Here we use the term organelle to designate mitochondria and plastids as the two compartments of endosymbiotic origin.

Apart from a few exceptions, such as the proteins destined for the peroxisome [5] or the nucleus [6], the initial step of protein targeting to the different compartments of the endomembrane system is the cotranslational insertion of the nascent polypeptide chain through the **SEC complex** into the lumen of the endoplasmic reticulum (ER) [7]. Such secretory proteins contain an N-terminal **signal peptide**, which is recognized by the signal recognition particle while still undergoing synthesis. Cotranslational translocation across SecY/Sec61 is a mechanism that was inherited from the prokaryotic secretion system [4,7,8].

Prokaryotes, both bacteria and **archaea**, secrete proteins across their plasma membrane. Two main secretion systems can be distinguished: the Sec system, which is homologous to the eukaryotic SecY/Sec61 system of the ER membrane, and the twin-arginine translocon (TAT). Both protein secretion systems recognize cargo through targeting sequences that are characterized by a positively charged amino-terminal region [9]. The TAT targeting sequence, and its two highly conserved and eponymous arginine residues, appear to require no soluble factor to associate first with the membrane and subsequently the TAT itself [10,11]. Furthermore, the proton motive force (PMF) across a membrane alone can be sufficient to initiate secretion of TAT-targeted proteins through electrophoresis [12,13] that is then completed by the TAT translocon, whose oligomerization also partly depends on the PMF [14]. Being independent from any additional form of energy currency other than the PMF – a prerequisite for driving ATP synthesis

Trends

Theory has it that mitochondria require proteins destined for their matrix to carry a charged N-terminal targeting sequence to overcome the energized inner membrane, but cross-species comparisons suggest an alternative and evolutionary ancient mechanism to work in parallel.

Algae and plants are required to distinguish between mitochondrial and plastid proteins, whose targeting motifs share many similarities, and charge appears to be an underestimated power for discrimination.

Phosphorylation of plastid N-terminal targeting sequences adds negative charges and might aid in avoiding false targeting to the mitochondrial matrix that requires positive charges.

A global comparison of targeting sequences, including those of organisms with noncanonical organelles, suggests a predominant selection pressure on mitochondrial and plastid targeting sequences is on charged amino acids.

[1]Institute for Molecular Evolution, University of Düsseldorf, 40225 Düsseldorf, Germany

*Correspondence: gould@hhu.de (S.B. Gould).

 http://dx.doi.org/10.1016/j.tcb.2016.07.001

and hence present anyway – presents an evolutionary advantage and a reason why such targeting and translocation mechanisms are found conserved across all bacterial and eukaryotic phyla.

Eukaryotic signal peptides that target proteins for cotranslational import into the ER carry an overall less positively charged N-terminus than the signal peptides of prokaryotes that target cargo to their plasma membrane [15]. This is because there is a need to avoid crosstalk with the targeting sequences of mitochondrial proteins, for reasons we discuss below. Nuclear-encoded proteins destined for the mitochondrial matrix and plastid stroma also carry N-terminal targeting sequences with characteristic charges, but they evolved in the wake of the endosymbiotic origin of the organelles to which they are targeted.

Endosymbiotic Theory for Organelle Origins

Endosymbiosis had a decisive impact on the evolution of complex life. It gave rise to all macroscopic life (eukaryotes and ultimately animals) and heritable **photosynthesis** in eukaryotes (algae and plants). How so? First, the origin of eukaryotes hinges upon the origin of mitochondria. The transition from prokaryotic to eukaryotic cells commenced with the endosymbiotic integration of one prokaryote into another, namely an alphaproteobacterium into an archaeal host [2,15,16] (see Figure I in Box 1).

Some still debate the level of complexity the host had reached before the alphaproteobacterium came to reside in its cytosol [4,17,18], but it is evident that no other prokaryote – in the 4 billion years of life's evolution – evolved a complexity that vaguely resembles that of a eukaryote, except for one archaeal lineage that acquired the mitochondrion [19,20]. The only eukaryote that appears to lack mitochondria altogether lost them secondarily, which hence changes nothing with respect to how we need to think about the origin of eukaryotes [21]. The universal presence of mitochondria in eukaryotes is testament to the importance of the organelle to eukaryotic biology. Mitochondria provided the energy to evolve complexity [22] and arguably even the seed from which the eukaryotic endomembrane system itself evolved [4]. There is every indication that the emergence of eukaryotes required a mitochondrion.

Second, photosynthesis in eukaryotes stems from the endosymbiotic acquisition and integration of a **cyanobacterium** by a **heterotrophic** host (see Figure I in Box 1). The monophyletic acquisition of the plastid some 1.2 billion years ago by a host of unknown nature generated three major algal lineages: the glaucophytes, rhodophytes and chloroplastida, which include land plants [23–25]. Like mitochondria, the **primary plastids** of all algae and land plants are surrounded by two membranes. Many ecologically relevant algal groups and infamous parasites evolved thereof such as the agent of malaria, ***Plasmodium***, however, house **complex plastids.** These plastids (in *Plasmodium* and related apicomplexan parasites known as the **apicoplast**) are surrounded by more than two membranes and they evolved as a result of secondary endosymbiosis (Box 1). The integration of endosymbionts, primary or otherwise, into the host's biology and their evolutionary transformation into organelles is accompanied by the streamlining of the endosymbiont's genome and the transfer of genetic material to the host nucleus. This process is known as **endosymbiotic gene transfer (EGT)** [26]. The result is that hundreds of proteins required to maintain the biochemistry of both the mitochondrion and plastid are now nuclear-encoded and need to be imported from the cytosol where they are synthesized (Figure 1, Key Figure).

Principles of Protein Targeting to Mitochondria

Virtually all eukaryotes house mitochondria or organelles derived thereof, called **hydrogenosomes** or **mitosomes** (that is **reduced mitochondria**). The homology of their import machineries suggest that the **last eukaryotic common ancestor (LECA)** had a mitochondrion and a

Glossary

Amphiphilic α-helix: α-helix with hydrophilic amino acids on the one side and hydrophobic amino acids on the other.
Apicoplast: the reduced complex plastid of apicomplexan parasites such as *Plasmodium* and *Toxoplasma*. Has retained a minimal genome and is surrounded by four membranes.
Archaea: one of the two domains of prokaryotic life, the other one being Bacteria. Usually inhabit extreme habitats.
Archaeplastidal ancestor: the ancestral protist that engulfed a cyanobacterium from which the three algal lineages (Glaucophyta, Rhodophyta and Chloroplastida) evolved.
Bipartite leader sequence: a targeting sequence made up of two consecutive domains, each with a different targeting purpose
Complex plastid: plastids that originated through secondary endosymbiosis involving two eukaryotic partners, an alga and heterotrophic host; surrounded by three or more membranes, usually four.
Cyanobacterium: a group of gram-negative bacteria with the ability to photosynthesize.
ΔΨ: a potential across a membrane, generated by both the electrochemical gradient and the proton motif force
Electron transport chain (ETC): a series of biochemical oxidation–reduction reactions that results in the transfer of electrons via electron carriers. This electron flow can be coupled to chemiosmotic ATP synthesis by generating a proton gradient across the membrane.
Endomembrane system: elaborate intracellular membrane system unique to eukaryotes; includes e.g. the nucleus, and the endoplasmic reticulum, the lysosome, and all vesicle trafficking processes
Endosymbiosis: a symbiotic relationship where one partner, the symbiont, resides within another, the host.
Endosymbiotic gene transfer (EGT): describes the process of loss of genes from the endosymbiont to the nucleus of the host.
Heterotrophic: a lifestyle used to describe organisms that cannot synthesize their own food and are

Box 1. Eukaryogenesis and the Primary and Secondary Endosymbiotic Origin of Plastids

Both the mitochondrion and plastids are of endosymbiotic origin, but the circumstances under which they evolved and their effects on evolution differ significantly [1]. Eukaryogenesis describes the origin of eukaryotes and is the results of one prokaryote (an ancestor of extant alphaproteobacteria) coming to reside in another (an archaeal cell). The transition of the alphaproteobacterium to the mitochondrion inside the host was key to the origin of eukaryotes themselves [22,88]. With the endosymbiotic origin of the mitochondrion also the endomembrane system evolved in the last eukaryotic common ancestor (LECA), providing the blueprint for all eukaryotic cells [4].

Plastids trace back to the integration of a cyanobacterium by a eukaryotic host of unknown origin [23,24] (Figure I). From this event three algal lineages evolved: the glaucophytes, rhodophytes and chloroplastida, and from the latter, land plants downstream. Plastids evolved in a fully-fledged eukaryotic cell. Primary endosymbiosis, including the origin of the mitochondrion, generated organelles surrounded by two membranes.

Secondary endosymbiosis describes the uptake of one eukaryote (an alga) by another (again a heterotrophic host) (Figure I). This kind of incorporation has generated intricate chimeras with sometimes four genomes and in all cases plastids surrounded by additional membranes, usually four. These additional membranes generate a situation in which all complex plastid proteins are initially recognized by a signal peptide in the cytosol, for the purpose of cotranslational import into the ER, a situation unique to these protists. Four genomes are found in cryptophyte and chlorarachniophyte algae. Next to the nuclear, mitochondrial, and plastidal genome, there is that of the nucleomorph, the remnant nucleus of the algal endosymbiont. Like the nucleus, the nucleomorph encodes plastid proteins, too, but these only cross the inner two membranes homologous to the membranes of canonical, primary plastids. Several ecologically important protist groups are of secondary endosymbiotic origin, including: (i) haptophyte algae, the deposit of calcareous plates of their ancestors is what formed the White Cliffs of Dover; (ii) diatoms, substantial marine primary producers; (iii) dinoflagellates, infamous for causing toxic red tides, and having some of the biggest genomes known, and (iv) apicomplexan parasites, their best known representative being *Plasmodium falciparum*, the agent causing malaria. All endosymbiotic events have in common that they lead to novel compartments inside the cell that require additional levels of protein targeting.

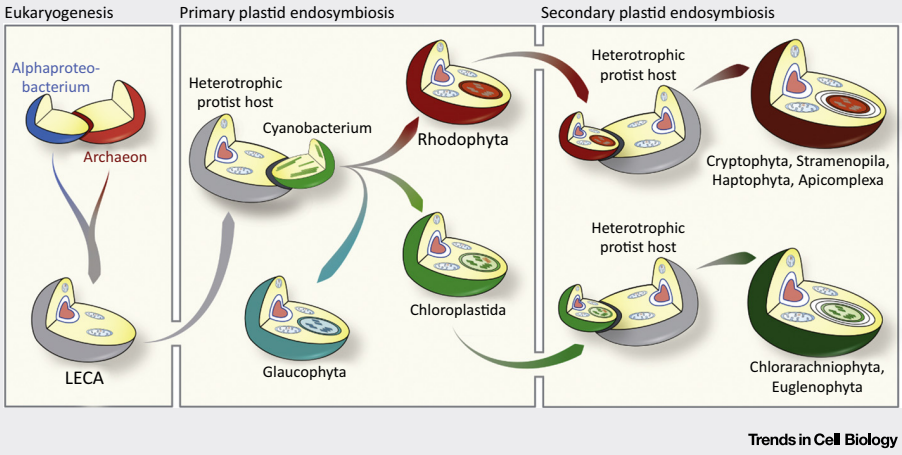


Figure I. Endosymbiotic Events that Shaped Eukaryote Evolution. For details please refer to Box 1 text.

minimal import machinery [27]. It was in LECA that mitochondrial import evolved, and with it N-terminal targeting sequences (NTSs). In the late 1970s it was independently discovered that nuclear-encoded proteins destined for mitochondria or plastids are synthesized on cytosolic ribosomes as precursor proteins, whose N-termini are cleaved upon reaching the matrix and stroma, respectively [28,29]. The presence of a cleavable NTS was long thought to be a prerequisite for mitochondrial and plastid targeting, but only until the analysis of mitochondrial membrane proteins that were found to lack an NTS [30]. The current paradigm has it that matrix (and stromal) proteins require a cleavable NTS, while integral membrane proteins and proteins of the organelle's intermembrane space (i.e., those that do not transverse the matrix or stroma) can be generally imported in the absence of such a motif [31,32], although there are exceptions to the rule and some details are more involved [33].

dependent on the environment and other organisms for organic substrates.

Hydrogenosomes: specialized forms of mitochondria, devoid of a genome and an electron transport chain for oxidative phosphorylation. Energy is instead generated through substrate level phosphorylation. Usually found in anaerobic protists

Intergenic sequences: stretches of DNA in the genome between two genes that do not code for proteins and occasionally can play regulatory roles.

Last eukaryotic common ancestor (LECA): the ancestral eukaryotic lineage from which all eukaryotic supergroups evolved.

Mitosomes: the most reduced forms of mitochondria, that lack a genome and any means of producing energy but remain a site for iron-cluster biogenesis.

Photosynthesis: a process by which some organisms use energy from sunlight to produce carbon-rich compounds from carbon dioxide and water.

***Plasmodium*:** a parasitic lineage of the phylum Apicomplexa. *Plasmodium falciparum* causes human malaria.

Primary plastids: plastids that trace back to primary endosymbiosis involving a prokaryote (the cyanobacterium) and eukaryote (protist host). Surrounded by two membranes.

Reduced mitochondria: all organelles that are evolutionarily of mitochondrial origin, but with different levels of reduced genome capacity and biochemistry.

SEC complex: a multicomponent translocon located in the endoplasmic reticulum membrane; involved in cotranslational translocation of secretory proteins.

Signal peptide: N-terminal targeting motif of a secretory protein. Recognized by the signal recognition particle that mediates interaction with the SEC complex; in prokaryotes at the plasma membrane, in eukaryotes at the ER.

Key Figure

Overview of Organelle and Import Evolution in Relation to Charge

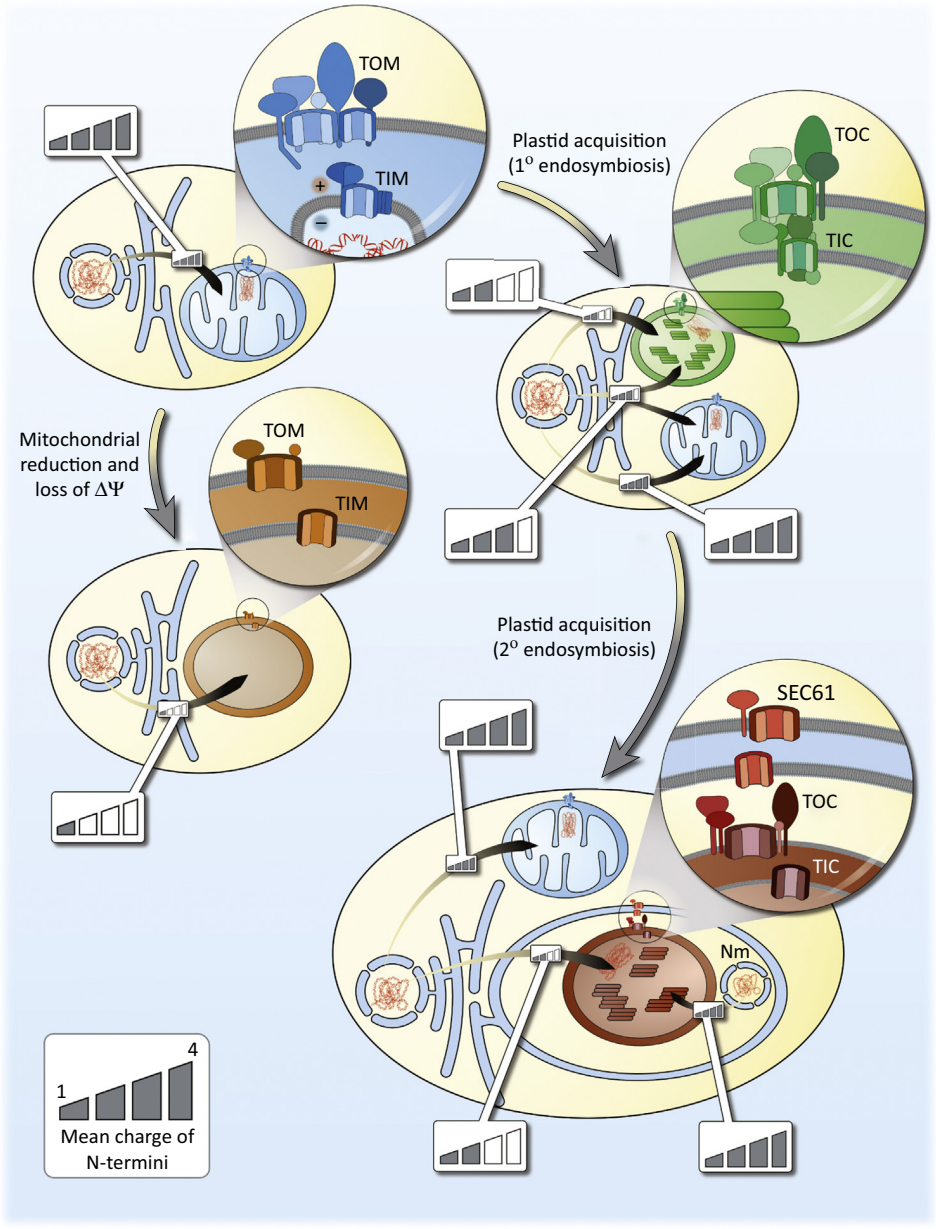


Figure 1. Mitochondria (shown in blue) import proteins across their two membranes into the matrix through the TOM and TIM (translocase of the outer/inner mitochondrial membrane) machinery. $\Delta\Psi$, indicated by the plus and minus symbols, runs across the inner mitochondrial membrane. In some eukaryotic lineages mitochondria experienced evolutionary reduction, leading, for example, to hydrogenosomes (brown), which lost the mitochondrial electron transport chain (ETC) and $\Delta\Psi$. A result of this is also a loss of charged N-terminal targeting sequences (NTS) and a reduced number of proteins that constitute TOM and TIM. Plastid acquisition through primary endosymbiosis requires the cells to discriminate nuclear-encoded mitochondrial from nuclear-encoded plastid proteins, also by the means of positive charges in the NTS.

(Figure legend continued on the bottom of the next page.)

The mitochondrial NTS (mNTS) bears several characteristic features, one of which is a positive charge. The length of mNTSs can vary, typically between 15 and 50 amino acids, but a part of the sequence has usually the potential to fold into **amphiphilic α-helices**, and they are generally depleted of acidic amino acids while being enriched in alanine, leucine, lysine, and in particular arginine [34,35]. The mNTSs are recognized by the receptor platform of the TOM complex (translocase of the outer/inner mitochondrial membrane; Figure 1). This complex includes Tom20, Tom22, and Tom70, which subsequently present the precursors to Tom40, the core translocase of the outer membrane [32]. The import of precursor proteins into isolated mitochondria whose receptor platform was affected through protease treatment suggested an alternative import mechanism, and questioned the importance of mNTS recognition at the outer mitochondrial membrane [36]. Yet, one cannot be sure to what degree all currently known receptors of the TOM complex were affected by the treatment and some caution is hence warranted. The discovery of organisms that harbor reduced mitochondria allowed the comparative analysis of organellar protein import [37–39] and provided some reasoning for the import of precursor proteins in the absence of receptors. These approaches highlighted the extent of reductive evolution, which, surprisingly, not only affected the number of proteins required to build a functional import apparatus, but also the mNTS and its charge.

Reduced mitochondria no longer harbor their own genome and translation machinery [40]; they are required to import all proteins they need to function from the cytosol. While less in number, the proteins that make up the import machineries of reduced mitochondria are homologous to those of typical mitochondria [39,41,42]. Several reports showed that protein translocation across the two membranes of reduced mitochondria is similar to mitochondria and the recognition and import of precursor proteins also required an NTS [43–45]. Later work, however, showed that an NTS is not a mandatory necessity for the successful import of proteins into the matrix of such organelles [46–50]. This is evident from (i) genome data showing that some nuclear-encoded mitosomal proteins already lost their NTS on the level of coding sequence [41] and (ii) experimental work that showed that the removal of the NTS in several cases only had a marginal, if any, effect on correct targeting and translocation [46,50,51]. Furthermore, NTS-independent targeting is conserved across hydrogenosomes and mitochondria, suggesting that the ancestral state of mitochondrial protein targeting was NTS independent [51]. It appears that the emergence and positively-charged nature of the mNTS was promoted by the electrochemical gradient of the inner mitochondrial membrane.

On the Origin of the Mitochondrial NTS and its Positive Charge

NTSs of mitochondrial matrix proteins are positively charged with an overall enrichment of basic amino acids [34,35]. The reason for this amino acid bias is thought to be the electrochemical gradient generated by the mitochondrial **electron transport chain (ETC)** across the inner membrane, which drives ATP synthesis [52]. The ETC generates a membrane potential ($\Delta\Psi$) that acts in concert with the positively charged N-termini to facilitate import of precursor proteins into the mitochondrial matrix across the inner membrane through electrophoresis [53] and simultaneously activates the Tim23 complex [54]. While $\Delta\Psi$ acts favorable on the mNTS on the membrane side facing the matrix, it poses an impediment on the side of the membrane facing

Some nuclear-encoded proteins are targeted to both organelles simultaneously and their NTSs carry a charge on average ranging between those of proteins targeted exclusively to mitochondria or plastids. In plastids of secondary endosymbiotic origin the stroma of the plastids is separated from the cytosol (and mitochondria) though additional membranes that are sometimes continuous with the membranes of the endoplasmic reticulum (ER). This relaxes the selection pressure on their NTS. The gradual arrows within the cells point from the genome in which a protein is encoded towards the organelle that is targeted. Square boxes on these arrows indicate the approximate positive charge of the NTSs, while the blowups show the protein import machineries of the respective organelles. Abbreviations: Nm, nucleomorph; TOC/TIC, translocases of the outer and inner chloroplast membranes.

the intermembrane space. This is circumvented by the negatively charged C-terminal extension of Tom22 that binds the positive charge of the mNTS and facilitates the immediate threading of the precursor into Tim23 [55]. In this way, the thermodynamically unfavorable conditions of the intermembrane space for the positively charged mNTS are avoided.

The role of $\Delta\Psi$ in providing the selection pressure for the maintenance of an overall positive charge in the mNTS is not only shown by experiments carried out in the laboratory, it is also evident in evolutionary comparisons. Hydrogenosomes and mitosomes have not only lost their entire genome, but also the ETC and ATP-synthesis at their inner membrane, concomitant with the loss of $\Delta\Psi$ [19,56]. Hydrogenosomal targeting sequences, as a result, are shorter in length compared to those of canonical mitochondria and have altogether lost their positive charge (Figure 2A) [51]. In the case of mitosomes, the evolutionary reduction of the NTS is even more pronounced. The NTS of many mitosome proteins has been lost entirely [41,57] and, if present, has a low positive charge which is comparable to that of hydrogenosomal proteins. Furthermore, a recent report shows that also yeast mitochondria import a native precursor which lacks a positively charged NTS, but only in the absence of $\Delta\Psi$ [58]. This supports the hypothesis that a positively charged N-terminus was selected for, and is maintained by, the electrochemical gradient of the inner mitochondrial membrane.

The proteins that prokaryotes secrete across their plasma membrane also make use of an NTS that carries positive charges at the very N-terminal region of the signal peptide [8,59]. One can assume that the archaeon that acquired the alphaproteobacterial endosymbiont – today's mitochondrion – was still secreting proteins across its plasma membrane. This likely posed a problem for the emergence of a positively charged NTS for mitochondrial targeting. In the presence of Sec and TAT protein secretion at the host's plasma membrane that depended on $\Delta\Psi$ [9], the archaeal cell would have, at least partly, secreted proteins across the plasma membrane destined for the organelle. If the early mitochondrion imported proteins in the presence of ongoing ATP-synthesis at the host's plasma membrane, then protein import into the organelle was likely NTS-independent, as indicated by recent data [51]. The emergence of the mitochondrion that synthesizes ATP and exports it to the host cytosol, however, resulted in the loss of $\Delta\Psi$ and ATP-synthesis at the archaeal host's plasma membrane, which is the reason why eukaryotic cells synthesize ATP inside their mitochondria and not at their plasma membrane [4]. Loss of plasma membrane $\Delta\Psi$ would have allowed selection for positively charged NTSs in mitochondrial targeting. In comparison to prokaryotes, the N-terminal charge of eukaryotic signal peptides is less pronounced [14]. This might reduce false targeting of secretory proteins to the mitochondrion. At the same time, the mitochondrial NTS is likely not of signal peptide origin, because (i) most of the proteins that are imported into the matrix are not proteins that bacteria secrete, (ii) those secretory proteins that are of bacterial origin are targeted to the ER rather than the mitochondria in eukaryotes [60–62], and (iii) an NTS can readily originate from non-coding **intergenic sequences** of DNA [63]. In summary the data, in particular the conservation of NTS-independent import into mitochondria and hydrogenosomes [51], speak in favor of ancestral mitochondrial targeting being NTS-independent and that a positively charged NTS evolved in an archaeal host that no longer utilized plasma membrane-localized ATP synthesis.

The Effect of Mitochondrial $\Delta\Psi$ on Plastid Targeting

The endosymbiotic acquisition of a cyanobacterium ushered photosynthesis into the eukaryotic lineage [23,24]. As with the alphaproteobacterium that gave rise to the mitochondrion [2,16,64], the cyanobacterial endosymbiont lost many of its genes to the nuclear genome of the host [26], again necessitating a mechanism to import organelle proteins from the cytosol. Plastid protein import evolved in a eukaryotic cell that had already established targeting to the mitochondrion

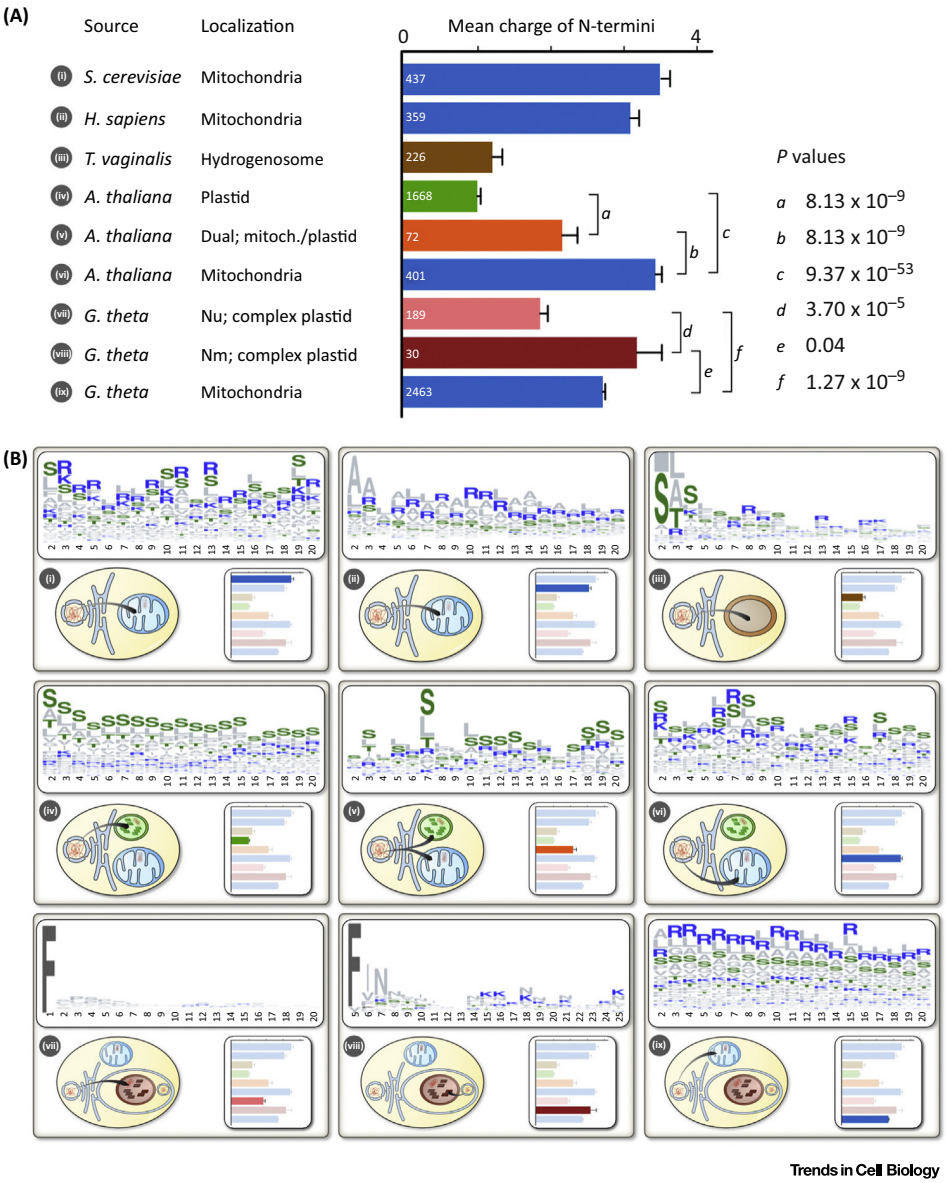


Figure 2. Average Charge of N-terminal Sequences Targeting Mitochondria and Plastids. (A) Bar diagram showing the mean charges of the different NTSs referred to throughout the text, with n indicating the number of protein N-termini screened. Proteins were selected from proteomes of the respective organisms [35,44,85–87] or from publicly available databases. Only those for which TargetP v1.1 predicted a cleavable presequence were analyzed. In each case the first 20 amino acids following the initial methionine were considered, except for the nucleus (Nu)- and nucleomorph (Nm)-encoded proteins, in which case the signal peptide as predicted by SignalP v4.1, preceding the NTS, was first removed. For those comparisons discussed in the text the *P*-values are provided on the right (determined through Wilcoxon signed-rank tests). (B) Depiction of relevant NTSs of the different organisms through amino acid sequence blots, where letter sizes indicate the frequency with which an amino acid is found at the corresponding position. Note, for instance, the high number of serine residues in nuclear-encoded plastid proteins of *Arabidopsis* (IV) and the prominent phenylalanine of nucleus- and nucleomorph-encoded plastid proteins of the cryptophyte alga *Guillardia theta* (VII and VIII, respectively). In addition, a representation of the bar diagram from the top is provided for comparison and schematics of the cells based on Figure 1.

and endomembrane system. This made evolving an organelle targeting mechanism particularly challenging, as the presence of pre-existing import machineries and targeting mechanisms increases the chance of incorrect protein sorting. The bacterial ancestry of the mitochondrion might allow false targeting of some proteins of cyanobacterial origin destined for the plastid, which is most likely the origin of dual-targeted proteins (see below). Of note, minor mistargeting of entire pathways is also thought to drive re-routing of metabolic compartmentalization in eukaryotes in more general [65]. Some organellar biochemistry, however, is less likely to tolerate the import of the wrong proteins, the ETC of both organelles being a prime candidate. Photosynthetic eukaryotes need to sort hundreds of proteins specifically to each organelle, yet, counterintuitively, extant plastid and mitochondrial targeting sequences are strikingly similar [30,66]. The eukaryote that established the first plastid found an elegant solution to plastid targeting, but early ancestors of the green algal lineage, including land plants, were forced to also resort to charge to solve the problem of protein sorting.

The plastid NTS (pNTS) evolved in the **archaeplastidal ancestor**, which also evolved the translocases of the outer and inner chloroplast membranes (TOC and TIC, respectively) [31]. Toc75, the main import channel of the outer membrane, is homologous to members of the bacterial Omp85 family [67] and a more primitive version of Toc75 is found in rhodophytes and glaucophytes [68]. Transit peptides of nuclear-encoded plastid proteins of red and glaucophyte algae are recognized by a conserved motif based on a single bulky aromatic amino acid, in most cases a phenylalanine (Table 1) [69,70]. It was suggested that the ability of the translocon to recognize a bulky aromatic amino acid was inherited from Omp85 itself [71], for which evidence was later provided [72]. The data included evidence for the switch of translocon orientation in the outer membrane, such that the phenylalanine-recognizing domain was now exposed to the cytosol of the host and no longer the intermembrane space of the plastid [73]. This was a simple solution to the problem of distinguishing between nuclear-encoded mitochondrial proteins (that carry a charged NTS) and nuclear-encoded plastid proteins (that carry a phenylalanine). This situation changed, however, for an unknown reason, in the green algal lineage where the phenylalanine-based motif was lost. Concomitant with this loss, is the increase of amino acids

Table 1. Exemplary Targeting Sequences of Various Organisms

Organisms	Gene	N-terminal sequences	Charge	Targeting	Refs
Cyanophora paradoxa **(Glaucophyte)**	PsaL	MAFITAIPAVPSTGAIVPAT	0.0	Plastid	[68]
Cyanidioschyzo merolae **(Rhodophyte)**	RuBisCO	MTDLPFTLDQLRIFQAIVVE	–2.0	Plastid	[69]
Arabidopsis thaliana **(Chlorophyte)**	AtAPX1	MTKNYPTVSEDYKKAVEKCR	2.0	Dual; Mitoch./Plastid	[80]
Oryza sativa **(Chlorophyte)**	OsAPX6	MAVVHRLLRRGLSAASPLPS	3.5	Mitochondria	[80]
	OsAPX5	MAVVHRILRRGLSAASPLPS	3.5	Mitochondria	[80]
	OsAPX7	MAAQRLAALHAAAPSAFSST	1.5	Plastid	[80]
	OsAPX8	MAERIAASLLPAASPSPAPS	1.0	Plastid	[80]
Plasmodium falciparum **(Alveolata)**	Synthetic	SKINNYSLINKYKINKYTHING	4.5	Plastid	[82]
	Synthetic	ITWILLNEVERTARGETPLASTID	–2.0	ER	[82]

Positively charged amino acids are highlighted in blue, negative ones in red and those that can potentially be phosphorylated in green. The bulky aromatic amino acid phenylalanine is underlined.

susceptible to phosphorylation, mostly serine, in the sequences that target the stroma of green plastids (Figure 2) [74]. The phosphorylation of serine residues of the pNTS facilitates the binding of the precursors to a guidance complex composed of a 14-3-3 protein and HSP70 [75,76]. Serine residues are dephosphorylated after binding of the guidance complex to the outer membrane translocase and prior to protein translocation [75,77]. Although serine phosphorylation is not essential for plastid import, it is thought to increase the import rate of precursor proteins [78].

In light of protein sorting, serine phosphorylation might also reduce the chance of false import into the mitochondrial matrix, as it adds negative charge to the mNTS. This idea is supported by the observation that phosphorylation of the mitochondrial NTS can prevent import of mitochondrial proteins [79], but the presence of an outer membrane phosphatase that is shared between the two organelles to dephosphorylate accidentally phosphorylated mitochondrial proteins [77] raises some open questions regarding how mNTS and pNTS phosphorylations are discriminated (see Outstanding Questions). We suggest that the underlying biological reason for pNTS phosphorylation was initially not to increase specificity to plastids, but to keep some plastid precursor proteins out of the plant mitochondria.

Dual Targeted Proteins and Bipartite Leader Sequences for Complex Plastids

Many proteins are simultaneously targeted to mitochondria and plastids. The importance of charge in that regard is apparent upon a simple comparison among the NTSs of mitochondrial and plastid precursor proteins and those targeted to both organelles simultaneously. When we compare their charges, based on a recently published list of dual-targeted *Arabidopsis thaliana* proteins [80], we observe a clear trend regarding the overall net charge of the different NTSs. The 1668 proteins predicted to be solely targeted to the plastid have a mean charge of 1.0, which is significantly lower than the mean charge of 3.4 for the 401 predicted mitochondrial proteins with a cleavable NTS (Figure 2A). The 72 proteins known to be dual targeted have a mean charge of 2.1 (Figure 2A), which is significantly different from both nuclear-encoded plastid and mitochondrial proteins. This is also apparent from logoplots and the overall lower number of arginine residues observed within the first 20 amino acids (Figure 2B). A dual-targeted protein thus carries a residual positive charge that cannot be lost without negatively affecting its ability to be imported into the mitochondrion.

The significance of charge is further apparent from an alternative to dual targeting of a protein, by encoding two isoforms of a protein by separate genes. The *A. thaliana* protein *At*APX1 has experimentally been shown to be targeted to both organelles and the first 20 amino acids carry a median net charge of 2.0 (Table 1), typical for dual-targeted proteins in *Arabidopsis* (Figure 2A). *Oryza sativa*, however, encodes several homologs of *At*APX1. *Os*APX5 and 6 are specifically targeted to mitochondria, while *Os*APX7 and 8 are specifically targeted to plastids [81]. Screening their N-termini demonstrates that the crucial difference is the substitution of two nonpolar hydrophobic amino acids (leucine and isoleucine) with two positively charged amino acids (arginine) in those two copies targeted to the mitochondria (Table 1). A beneficial twist is the ease with which such a change in targeting can occur: a single nucleotide mutation in the wobble base switches between the codons encoding the amino acids serine and arginine. No more was needed than a simple change in charge, achieved via a single point mutation, to reroute those proteins.

The most complex eukaryotic cells are those of secondary endosymbiotic origin (Box 1), in which plastids are separated from mitochondria through additional membranes. The outermost membrane is of ER origin and in some species even continuous with it [82] (Figure 1). The complex plastids of cryptophyte and chlorarachniophyte algae are both surrounded by

four membranes and they still harbor a nucleomorph, the remnant nucleus of the engulfed alga, between the outer and inner membrane pairs. The nucleomorph encodes plastid proteins, too. While targeting to these complex plastids differs in detail, the initial step depends on a canonical signal peptide for the cotranslational targeting across the Sec61/SecY machinery, which precedes the NTS required for translocation across the remaining inner three membranes. Such a targeting sequence is aptly referred to as a **bipartite leader sequence**.

Thus, in the cytosol of these cells, nuclear-encoded plastid proteins are distinguished from mitochondrial proteins through the signal peptide and not the NTS. This topology has relaxed the selection pressure on NTS charge, resulting in a much higher positive charge that is comparable to those of dual-targeted proteins of *Arabidopsis* (Figure 2; VII and V, respectively). The NTS charge of nucleomorph-encoded plastid proteins is even higher (Figure 2; VIII) and exceeds that of nuclear-encoded mitochondrial proteins (Figure 2; IX). The latter might also arise from the high AT-content of the nucleomorph genome and hence a codon bias, but nevertheless demonstrates that the absence of a selection pressure on the NTS (to keep positive charges low and avoid false targeting to mitochondria), permits the presence of positive charges in a NTS targeting plastids. This trend is observed in cryptophytes harboring a reduced red alga as a plastid, as well as in chlorarachniphytes harboring a reduced green alga as a plastid [70]. Positive charges in the NTS of nuclear-encoded proteins of complex plastids now aid in plastid targeting, as demonstrated by using synthetic peptides that resemble English words [83] (Table 1). Moreover, the overall charge rather than a specific position of basic amino acids in the NTS provides the necessary information for accurate targeting to the *Toxoplasma* apicoplast [84]. In summary, dual-targeted proteins and the unique complex plastids of some algae (and apicomplexan parasites) provide evidence for the crucial importance of charge regarding organelle protein sorting.

Concluding Remarks

Mitochondria and plastids are the result of independent endosymbiotic events. Although the import machineries of the two organelles share remarkable functional similarities, plastid protein targeting evolved in the presence of established mitochondrial targeting and targeting to the remaining eukaryotic compartments. The framework of protein targeting to organelles of endosymbiotic origin is resolved, but a broader comparison of extant plastid and mitochondrial diversity reveals the limits of some generalizations. Mitosomal import differs from mitochondrial import, and protein targeting to the plastids of the model system *Arabidopsis* differs from protein targeting to the plastids of all red algae and glaucophytes. Plastid targeting sequences, amidst their numerous similarities to those of proteins targeted to the mitochondrial matrix, contain sufficient information to ensure correct sorting, and the charge of their NTS plays an important role. A simple change in charge is sometimes all that is needed to determine which organelle is targeted by a protein. The loss of charge in targeting sequences of hydrogenosomal and mitosomal proteins, and the shift of charge in proteins targeted to complex plastids, is evidence that a constant selection pressure works on the overall charge of targeting sequences. The role of charge in protein sorting points to numerous new avenues of pursuit in studying organelle protein import.

Outstanding Questions

How do typical mitochondria, such as those of yeast, recognize and import hydrogenosomal proteins that lack positively charged N-terminal targeting sequences?

What is the true purpose of serine phosphorylation in targeting sequences of plastid proteins and how is it distinguished from those phosphorylations occurring on mitochondrial targeting sequences?

How much of the contamination one finds during proteome analyses of organelles is due to the experimental protocol and how much due to mistargeting?

What caused the green algal lineage to lose the phenylalanine that is sufficient for specific targeting to the plastids of red algae and glaucophytes? What advantages (evident from the success of the green algal lineage and land plants) did it bring?

Is the inner mitochondrial membrane partitioned in such a way as to allow import of proteins lacking a positively charged NTS, or are those proteins only imported into mitochondria temporarily lacking $\Delta\Psi$?

Acknowledgments

SBG gratefully acknowledges funding by the German Research Council (1825/4-1 and CRC1208).

References

1. Zimorski, V. *et al.* (2014) Endosymbiotic theory for organelle origins. *Curr. Opin. Microbiol.* 22, 38–48
2. Martin, W.F. *et al.* (2015) Endosymbiotic theories for eukaryote origin. *Phil. Trans. R. Soc. B.* 370, 20140330
3. Schlacht, A. *et al.* (2014) Missing pieces of an ancient puzzle: evolution of the eukaryotic membrane-trafficking system. *Cold Spring Harb. Perspect. Biol.* 6, a016048-a016048

4. Gould, S.B. *et al.* (2016) Bacterial vesicle secretion and the evolutionary origin of the eukaryotic endomembrane system. *Trends Microbiol.* 24, 525–534
5. Smith, J.J. and Aitchison, J.D. (2013) Peroxisomes take shape. *Nat. Rev. Mol. Cell. Biol.* 14, 803–817
6. Marfori, M. *et al.* (2011) Molecular basis for specificity of nuclear import and prediction of nuclear localization. *BBA - Mol. Cell Res.* 1813, 1562–1577
7. Park, E. and Rapoport, T.A. (2012) Mechanisms of Sec61/SecY-mediated protein translocation across membranes. *Annu. Rev. Biophys.* 41, 21–40
8. Cao, T.B. and Saier, M.H., Jr (2003) The general protein secretory pathway: phylogenetic analyses leading to evolutionary conclusions. *BBA - Biomembranes* 1609, 115–125
9. Natale, P. *et al.* (2008) Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane–distinct translocases and mechanisms. *Biochim. Biophys. Acta* 1778, 1735–1756
10. Snyder, A. *et al.* (2006) Role of the *Pseudomonas aeruginosa* PlcH Tat signal peptide in protein secretion, transcription, and cross-species Tat secretion system compatibility. *J. Bacteriol.* 188, 1762–1774
11. Shanmugham, A. *et al.* (2006) Membrane binding of twin arginine preproteins as an early step in translocation. *Biochemistry* 45, 2243–2249
12. Robinson, C. *et al.* (1993) Protein translocation across the thylakoid membrane-a tale of two mechanisms. *FEBS Lett.* 325, 67–69
13. Cline, K. *et al.* (1993) Multiple pathways for protein transport into or across the thylakoid membrane. *EMBO J.* 12, 4105–4114
14. Nielsen, H. *et al.* (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* 10, 1–6
15. Berks, B.C. (2015) The twin-arginine protein translocation pathway. *Annu. Rev. Biochem.* 84, 843–864
16. McInerney, J.O. *et al.* (2014) The hybrid nature of the Eukaryota and a consilient view of life on Earth. *Nat. Rev. Microbiol.* 12, 449–455
17. López-García, P. and Moreira, D. (2015) Open questions on the origin of eukaryotes. *Trends. Ecol. Evol.* 30, 697–708
18. Koonin, E.V. (2015) Origin of eukaryotes from within archaea, archaeal eukaryome and bursts of gene gain: eukaryogenesis just made easier? *Phil. Trans. R. Soc. B.* 370, 20140333
19. Müller, M. *et al.* (2012) Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. *Microbiol. Mol. Biol. Rev.* 76, 444–495
20. Embley, T.M. and Martin, W. (2006) Eukaryotic evolution, changes and challenges. *Nature* 440, 623–630
21. Karnkowska, A. *et al.* A eukaryote without a mitochondrial organelle. *Curr. Biol.* 26, 1274–1284
22. Lane, N. and Martin, W. (2010) The energetics of genome complexity. *Nature* 467, 929–934
23. Archibald, J.M. (2009) The puzzle of plastid evolution. *Curr. Biol.* 19, R81–R88
24. McFadden, G.I. (2014) Origin and evolution of plastids and photosynthesis in eukaryotes. *Cold Spring Harb. Perspect. Biol.* 6, a016105-a016105
25. de Vries, J. *et al.* (2016) Streptophyte terrestrialization in light of plastid evolution. *Trends Plant Sci.* 21, 467–476
26. Timmis, J.N. *et al.* (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat. Rev. Genet.* 5, 123–135
27. Lithgow, T. and Schneider, A. (2010) Evolution of macromolecular import pathways in mitochondria, hydrogenosomes and mitosomes. *Phil. Trans. R. Soc. B.* 365, 799–817
28. Chua, N.H. and Schmidt, G.W. (1978) Post-translational transport into intact chloroplasts of a precursor to the small subunit of ribulose-1,5-bisphosphate carboxylase. *Proc. Natl. Acad. Sci. U.S.A.* 75, 6110–6114
29. Maccecchini, M.L. *et al.* (1979) Import of proteins into mitochondria - precursor forms of the extra-mitochondrially made F1-ATPase subunits in yeast. *Proc. Natl. Acad. Sci. U.S.A.* 76, 343–347
30. Millar, D.G. and Shore, G.C. (1994) Mitochondrial Mas70p signal anchor sequence. Mutations in the transmembrane domain that disrupt dimerization but not targeting or membrane insertion. *J. Biol. Chem.* 269, 12229–12232
31. Schleiff, E.E. and Becker, T.T. (2010) Common ground for protein translocation: access control for mitochondria and chloroplasts. *Nat. Rev. Mol. Cell. Biol.* 12, 48–59
32. Neupert, W. (2015) A perspective on transport of proteins into mitochondria: a myriad of open questions. *J. Mol. Biol.* 427, 1135–1158
33. Chacinska, A. *et al.* (2009) Importing mitochondrial proteins: machineries and mechanisms. *Cell* 138, 628–644
34. Heijne, G.V. *et al.* (1989) Domain structure of mitochondrial and chloroplast targeting peptides. *Eur. J. Biochem.* 180, 535–545
35. Vögtle, F-N. *et al.* (2009) Global analysis of the mitochondrial N-proteome identifies a processing peptidase critical for protein stability. *Cell* 139, 428–439
36. Pfaller, R. *et al.* (1989) Mitochondrial protein import. Bypass of proteinaceous surface receptors can occur with low specificity and efficiency. *J. Biol. Chem.* 264, 34–39
37. Häusler, T. *et al.* (1997) Conservation of mitochondrial targeting sequence function in mitochondrial and hydrogenosomal proteins from the early-branching eukaryotes *Crithidia*, *Trypanosoma and Trichomonas*. *Eur. J. Cell. Biol.* 73, 240–251
38. Lister, R. *et al.* (2005) Protein import into mitochondria: origins and functions today. *Mol. Membr. Biol.* 22, 87–100
39. Doležal, P. *et al.* (2006) Evolution of the molecular machines for protein import into mitochondria. *Science* 313, 314–318
40. Allen, J.F. (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
41. Waller, R.F. *et al.* (2009) Evidence of a reduced and modified mitochondrial protein import apparatus in microsporidian mitosomes. *Eukaryot. Cell* 8, 19–26
42. Hewitt, V. *et al.* (2011) Minor modifications and major adaptations: the evolution of molecular machines driving mitochondrial protein import. *Biochim. Biophys. Acta* 1808, 947–954
43. Bradley, P.J. *et al.* (1997) Targeting and translocation of proteins into the hydrogenosome of the protist Trichomonas: similarities with mitochondrial protein import. *EMBO J.* 16, 3484–3493
44. Carlton, J.M. *et al.* (2007) Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. *Science* 315, 207–212
45. Burri, L. *et al.* (2006) Microsporidian mitosomes retain elements of the general mitochondrial targeting system. *Proc. Natl. Acad. Sci. U.S.A.* 103, 15916–15920
46. Zimorski, V. *et al.* (2013) The N-terminal sequences of four major hydrogenosomal proteins are not essential for import into hydrogenosomes of *Trichomonas vaginalis*. *J. Eukaryot. Microbiol.* 60, 89–97
47. Mentel, M. *et al.* (2008) Protein import into hydrogenosomes of Trichomonas vaginalis involves both N-terminal and internal targeting signals: a case study of thioredoxin reductases. *Eukaryot. Cell* 7, 1750–1757
48. Regoes, A. *et al.* (2005) Protein import, replication, and inheritance of a vestigial mitochondrion. *J. Biol. Chem.* 280, 30557–30563
49. Goldberg, A.V. *et al.* (2008) Localization and functionality of microsporidian iron–sulfur cluster assembly proteins. *Nature* 452, 624–628
50. Hamilton, V. *et al.* (2014) Trypanosome alternative oxidase possesses both an N-terminal and internal mitochondrial targeting signal. *Eukaryot. Cell* 13, 539–547
51. Garg, S. *et al.* (2015) Conservation of transit peptide-independent protein import into the mitochondrial and hydrogenosomal matrix. *Genome Biol. Evol.* 7, 2716–2726
52. Martin, J. *et al.* (1991) Role of an energized inner membrane in mitochondrial protein import. $\Delta\Psi$ drives the movement of presequences. *J. Biol. Chem.* 266, 18051–18057
53. McBride, H.M. *et al.* (1995) Insertion of an uncharged polypeptide into the mitochondrial inner membrane does not require a transbilayer electrochemical potential: effects of positive charges. *BBA - Biomembranes* 1237, 162–168

54. Waegemann, K. *et al.* (2015) Cooperation of TOM and TIM23 complexes during translocation of proteins into mitochondria. *J. Mol. Biol.* 427, 1075–1084
55. Meinecke, M. *et al.* (2006) Tim50 maintains the permeability barrier of the mitochondrial inner membrane. *Science* 312, 1523–1526
56. Müller, M. (1993) The hydrogenosome. *J. Gen. Microbiol.* 139, 2879–2889
57. Heinz, E. and Lithgow, T. (2013) Back to basics: a revealing secondary reduction of the mitochondrial protein import pathway in diverse intracellular parasites. *Biochim. Biophys. Acta* 1833, 295–303
58. Turakhiya, U. *et al.* (2016) Protein import by the mitochondrial presequence translocase in the absence of a membrane potential. *J. Mol. Biol.* 428, 1041–1052
59. Payne, S.H. *et al.* (2012) Unexpected diversity of signal peptides in prokaryotes. *MBio* 3, e00339–e412
60. Hall, J. *et al.* (1990) Eukaryotic and prokaryotic signal peptides direct secretion of a bacterial endoglucanase by mammalian cells. *J. Biol. Chem.* 265, 19996–19999
61. Soole, K.L. *et al.* (1992) Constitutive secretion of a bacterial enzyme by polarized epithelial cells. *J. Cell Sci.* 102, 495–504
62. Allet, B. *et al.* (1997) A bacterial signal peptide directs efficient secretion of eukaryotic proteins in the baculovirus expression system. *Protein Express. Purif.* 9, 61–68
63. Baker, A. and Schatz, G. (1987) Sequences from a prokaryotic genome or the mouse dihydrofolate reductase gene can restore the import of a truncated precursor protein into yeast mitochondria. *Proc. Natl. Acad. Sci. U.S.A.* 84, 3117–3121
64. Rivera, M.C. and Lake, J.A. (2004) The ring of life provides evidence for a genome fusion origin of eukaryotes. *Nature* 431, 152–155
65. Martin, W. (2010) Evolutionary origins of metabolic compartmentalization in eukaryotes. *Phil. Trans. R. Soc. B.* 365, 847–855
66. Murcha, M.W. *et al.* (2014) The plant mitochondrial protein import apparatus - the differences make it interesting. *BBA - Gen. Subjects* 1840, 1233–1245
67. Day, P.M. *et al.* (2014) Evolution and targeting of Omp85 homologs in the chloroplast outer envelope membrane. *Front. Plant Sci.* 5, 535
68. Sommer, M.S. and Schleiff, E. (2014) Evolution of the protein translocons of the chloroplast envelope. In *Endosymbiosis* (Löffelhardt, W., ed.), pp. 81–110, Springer
69. Köhler, D. *et al.* (2015) Identification of protein N-termini in *Cyanophora paradoxa* cyanelles: transit peptide composition and sequence determinants for precursor maturation. *Front. Plant Sci.* 6, 559
70. Patron, N.J. and Waller, R.F. (2007) Transit peptide diversity and divergence: a global analysis of plastid targeting signals. *Bioessays* 29, 1048–1058
71. Steiner, J.M. *et al.* (2005) Homologous protein import machineries in chloroplasts and cyanelles. *Plant J.* 44, 646–652
72. Wunder, T. *et al.* (2007) The invariant phenylalanine of precursor proteins discloses the importance of Omp85 for protein translocation into cyanelles. *BMC Evol. Biol.* 7, 236
73. Sommer, M.S. *et al.* (2011) Chloroplast Omp85 proteins change orientation during evolution. *Proc. Natl. Acad. Sci. U.S.A.* 108, 13841–13846
74. Soll, J. and Schleiff, E. (2004) Protein import into chloroplasts. *Nat. Rev. Mol. Cell. Biol.* 5, 198–208
75. Waegemann, K. and Soll, J. (1996) Phosphorylation of the transit sequence of chloroplast precursor proteins. *J. Biol. Chem.* 271, 6545–6554
76. May, T. and Soll, J. (2000) 14-3-3 proteins form a guidance complex with chloroplast precursor proteins in plants. *Plant Cell* 12, 53–64
77. Law, Y-S. *et al.* (2015) Phosphorylation and dephosphorylation of the presequence of precursor multiple organellar RNA editing factor3 during import into mitochondria from Arabidopsis. *Plant Physiol.* 169, 1344–1355
78. Nakrieko, K.A. *et al.* (2004) Fidelity of targeting to chloroplasts is not affected by removal of the phosphorylation site from the transit peptide. *Eur. J. Biochem.* 271, 509–516
79. Lee, J. *et al.* (2006) Mitochondrial localization of CNP2 is regulated by phosphorylation of the N-terminal targeting signal by PKC: implications of a mitochondrial function for CNP2 in glial and non-glial cells. *Mol. Cell. Neurosci.* 31, 446–462
80. Carrie, C. and Small, I. (2013) A re-evaluation of dual targeting of proteins to mitochondria and chloroplasts. *BBA - Mol. Cell Res.* 1833, 253–259
81. Xu, L. *et al.* (2013) Acquisition, conservation, and loss of dual-targeted proteins in land plants. *Plant Physiol.* 161, 644–662
82. Gould, S.B. *et al.* (2015) Protein import and the origin of red complex plastids. *Curr. Biol.* 25, R515–R521
83. Tonkin, C.J. *et al.* (2008) Evolution of malaria parasite plastid targeting sequences. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4781–4785
84. Tonkin, C.J. *et al.* (2006) N-terminal positively charged amino acids, but not their exact position, are important for apicoplast transit peptide fidelity in *Toxoplasma gondii*. *Mol. Biochem. Parasitol.* 150, 192–200
85. Curtis, B.A. *et al.* (2012) Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs. *Nature* 492, 59–65
86. Vaca Jacome, A.S. *et al.* (2015) N-terminome analysis of the human mitochondrial proteome. *Proteomics* 15, 2519–2524
87. Heazlewood, J.L. *et al.* (2004) Experimental analysis of the Arabidopsis mitochondrial proteome highlights signaling and regulatory components, provides assessment of targeting prediction programs, and indicates plant-specific mitochondrial proteins. *Plant Cell* 16, 241–256
88. Blackstone, N.W. (2013) Why did eukaryotes evolve only once?. Genetic and energetic aspects of conflict and conflict mediation. *Phil. Trans. R. Soc. B.* 368, 20120266

Publication VI

Mitochondria, the cell cycle, and the origin of sex via a syncytial eukaryote common ancestor

Sriram G. Garg[1] and William F. Martin[1]

1 - Institute for Molecular Evolution, Heinrich-Heine Universität Düsseldorf, Universitätsstraße 1, Düsseldorf 40225, Germany

The presented manuscript was peer-reviewed and published in **Genome Biology and Evolution** with impact factor of **4.257** on 26th June 2016

Contribution as **first author**

50%

Conceived significant portions of the hypothesis including mining the literature for relevant citations. Wrote parts of the text and discussion. Designed and illustrated all figures in the manuscript.

Genome Biology and Evolution

SMBE

GBE

Mitochondria, the Cell Cycle, and the Origin of Sex via a Syncytial Eukaryote Common Ancestor

Sriram G. Garg and William F. Martin*

Institute of Molecular Evolution, Heinrich-Heine Universität Düsseldorf, Düsseldorf 40225, Germany

*Corresponding author: E-mail: bill@hhu.de.

Accepted: May 29, 2016

Abstract

Theories for the origin of sex traditionally start with an asexual mitosing cell and add recombination, thereby deriving meiosis from mitosis. Though sex was clearly present in the eukaryote common ancestor, the order of events linking the origin of sex and the origin of mitosis is unknown. Here, we present an evolutionary inference for the origin of sex starting with a bacterial ancestor of mitochondria in the cytosol of its archaeal host. We posit that symbiotic association led to the origin of mitochondria and gene transfer to host's genome, generating a nucleus and a dedicated translational compartment, the eukaryotic cytosol, in which—by virtue of mitochondria—metabolic energy was not limiting. Spontaneous protein aggregation (monomer polymerization) and Adenosine Tri-phosphate (ATP)-dependent macromolecular movement in the cytosol thereby became selectable, giving rise to continuous microtubule-dependent chromosome separation (reduction division). We propose that eukaryotic chromosome division arose in a filamentous, syncytial, multinucleated ancestor, in which nuclei with insufficient chromosome numbers could complement each other through mRNA in the cytosol and generate new chromosome combinations through karyogamy. A syncytial (or coenocytic, a synonym) eukaryote ancestor, or Coeca, would account for the observation that the process of eukaryotic chromosome separation is more conserved than the process of eukaryotic cell division. The first progeny of such a syncytial ancestor were likely equivalent to meiospores, released into the environment by the host's vesicle secretion machinery. The natural ability of archaea (the host) to fuse and recombine brought forth reciprocal recombination among fusing (syngamy and karyogamy) progeny—sex—in an ancestrally meiotic cell cycle, from which the simpler haploid and diploid mitotic cell cycles arose. The origin of eukaryotes was the origin of vertical lineage inheritance, and sex was required to keep vertically evolving lineages viable by rescuing the incipient eukaryotic lineage from Muller's ratchet. The origin of mitochondria was, in this view, the decisive incident that precipitated symbiosis-specific cell biological problems, the solutions to which were the salient features that distinguish eukaryotes from prokaryotes: A nuclear membrane, energetically affordable ATP-dependent protein–protein interactions in the cytosol, and a cell cycle involving reduction division and reciprocal recombination (sex).

Key words: origin, eukaryotes, endomembrane system, meiosis, mitosis, syngamy, karyogamy, coeocytic, reduction division, chromosome segregation, alternation of generations.

Sex Is Essential In Eukaryotes

Few problems have baffled evolutionary biologists more thoroughly than the origin of meiotic sex. Eukaryotes do it, prokaryotes do not. The basic machinery of sexual recombination was present in the last common ancestor of extant eukaryotes, because 1) the genes underpinning sexual recombination are homologous across all eukaryotes studied so far (Ramesh et al. 2005) and 2) all eukaryotic clades either undergo sexual recombination or have the machinery in their genomes to do so (Speijer et al. 2015; Bloomfield 2016). Thus, sex was clearly present in the eukaryote common ancestor and was furthermore manifest in a form homologous to its modern-day incarnations. The classical questions have remained the same for decades: How did it arise, from what, what benefits bore its origin, and what benefits maintained its presence? Many papers and books have been written on the origin of sex (Cleveland 1947; Williams 1975; Maynard Smith 1978; Bell 1982; Bernstein et al. 1984; Uyenoyama and Bengtsson 1989; Hurst and Nurse 1991; Otto and Goldstein 1992; Maynard Smith and Szathmary 1997; Cavalier-Smith 2002; Solari 2002; Wilkins and Holliday 2008; Goodenough and Heitman 2014; Cavalier-Smith 2010; Hörandl and Hadacek 2013). Papers still continue to come in on the topic, a good indication that no one has solved the problems

to everyone's satisfaction (Havird et al. 2015; Radzvilavicius and Blackstone 2015; Speijer et al. 2015).

The issue of why eukaryotes as a lineage never lost sex is most readily attributed to genetic load, or the cumulative effects of sublethal mutations in clonally growing organisms, a population genetic process called Muller's ratchet (Muller 1964; Felsenstein 1974). Without recombination, reproduction is strictly clonal, mutation being inevitable and leading to the steady accumulation of deleterious mutations (Muller 1964). In the absence of recombination, these mutations will ultimately lead to extinction (Muller 1964; Felsenstein 1974; Moran 1996; Crow 2005). Though occasional high ploidy can possibly delay the effects of Muller's ratchet, it cannot alleviate the effects (Kondrashov 1994); ploidy is not a substitute for sex.

In this article, we will be arguing that recombination is essential for long-term lineage survival of both prokaryotes and eukaryotes. We will also argue that recombination rescues organisms from extinction at the hands of Muller's ratchet. Because sex is the only means of recombination known in eukaryotes, it seems likely to us that the avoidance of Muller's ratchet is the reason that eukaryotes have preserved meiotic sex throughout their history, which spans some 1.7 Gyr (Parfrey et al. 2011). However, for readers who doubt the power of Muller's ratchet, we interject that is indeed possible that selective pressures other than the escape of Muller's ratchet are responsible for eukaryotes having retained meiotic recombination. Yet for the purposes of this article, it is immaterial whether Muller's ratchet or some other selective force is responsible for the retention of sex (meiotic recombination) throughout all of eukaryotic history up to the present. For such skeptics, we emphasize: It is an observation from biology (not a prediction from theoretical population genetics) that recombination and the proteins required have been strictly conserved during eukaryote evolution (Ramesh et al. 2005; Speijer et al. 2015). From that we can readily and robustly infer that recombination is essential to long-term eukaryote survival. We are also fully aware that various eukaryotes appear to have lost the ability to undergo sex in some terminal branches (Maynard Smith 1986; Welch and Meselson 2000; Rougier and Werb 2001; Halary et al. 2011; Hand and Koltunow 2014; Speijer et al. 2015). But we reaffirm: If homologous recombination was not essential over the long term in eukaryotes, it would have been lost long ago and in many independent lineages. We posit that sex was conserved throughout eukaryote evolution by purifying selection as a means to escape Muller's ratchet. Only the conserved and clearly essential nature of sex (meiotic recombination) is vital to our inference for its origin, not the exact reason for why sex has been conserved. As Maynard Smith (1986) put it: "... it is clear that for one reason or another it is very difficult to give up sex once you have it." Our article is not about the "... hard to give up ..." part, it is about the "... once you have it ..." part, which is more challenging, because it falls into the prokaryote–eukaryote transition.

It is also very important to note this: Eukaryotes and prokaryotes use conserved mechanisms and homologous enzymes to perform DNA recombination (Camerini-Otero and Hsieh 1995; Ramesh et al. 2005), given the presence of two different DNA molecules within the cell. But crucially, the way(s) in which DNA substrates for recombination enter the cell and come into contact for recombination differ fundamentally across the prokaryote–eukaryote divide, as explained in the following.

In prokaryotes, the mechanisms that bring DNA into the cell for recombination are the mechanisms of lateral gene transfer (LGT): Transformation, conjugation, transduction, and gene transfer agents (Jones and Sneath 1970; Doolittle 1999; Martin 1999; Ochman et al. 2000; Lang et al. 2012). These mechanisms operate unidirectionally, from donor to recipient. Except for some archaeal lineages that undergo cell fusion and recombination in the fused state (Naor and Gophna 2014) these mechanisms do not obey taxonomic boundaries, species or otherwise, and over time they generate the pangenomes typical of prokaryotic taxa (Rasko et al. 2008). What are pangenomes? Pangenomes are readily illustrated as follows: Although 61 different humans (or individuals of any eukaryotic species) possess essentially the same genes (some copy number variation notwithstanding), 61 strains of *Escherichia coli*, each harboring about 4,500 genes, possess in total about 18,000 genes (the pangenome), with only 1,000 genes being present in all strains (the core genome) (Lukjancenko et al. 2010). Thus, homologous recombination in prokaryotes entails the introduction of foreign DNA into the cell through the mechanisms that we typically associate with LGT (transformation, conjugation, transduction, and/or gene transfer agents), and in these cases recombination is never reciprocal. In fusing archaeal species, there is no clear evidence that recombination is homologous (Papke et al. 2004; Naor and Gophna 2014). In prokaryotes, recombination is not reciprocal, is always unidirectional from donor to recipient, and operates with LGT machinery: Conjugation, transformation, transduction, or gene transfer agents.

In eukaryotes, homologous recombination occurs during meiosis, is always reciprocal and occurs between individuals of the same species. The DNA substrates for eukaryotic recombination come into contact through gamete fusion (syngamy) followed either immediately or after a dikaryon or multinucleated stage by nuclear fusion (karyogamy). Because eukaryotes arose from prokaryotes (Williams et al. 2013), eukaryotes must have lost the prokaryotic LGT mechanisms present in their ancestors, while retaining the enzymatic machinery that performs homologous recombination (Ramesh et al. 2005; Speijer et al. 2015; Bloomfield 2016). Without sex (reciprocal recombination), eukaryotes would have succumbed to Muller's ratchet long ago. Sex was an invention of the eukaryote common ancestor that has neither

been replaced nor fundamentally improved upon in the approximately 1.7 Gyr since eukaryotes arose.

The Nature of the Cell That Evolved Sex

Traditionally, approaches to the origin of sex start with a mitotic cell (typically a hypothetical "primitive" eukaryote) and introduce factors and effects that lead to meiosis (Cleveland 1947; Williams 1975; Maynard Smith 1978; Bell 1982; Bernstein et al. 1984; Uyenoyama and Bengtsson 1989; Hurst and Nurse 1991; Otto and Goldstein 1992; Maynard Smith and Szathmary 1997; Cavalier-Smith 2002, 2010; Solari 2002; Wilkins and Holliday 2008; Hörandl and Hadacek 2013; Goodenough and Heitman 2014). That approach sounds reasonable enough at first encounter, but upon closer inspection, some fairly severe problems quickly become apparent. First, if the ancestral eukaryote (the cell that evolved sex) was mitotic, it was an asexual mitotic cell, obviously. In our view, an asexual mitotic cell is a very problematic intermediate, because it raises the question of how it escaped Muller's ratchet both during the time 1) before it evolved sex and 2) while it was evolving mitosis. Yet more pressingly from our perspective, if meiosis arose from mitosis, how and from what did mitosis arise?

Recent progress in understanding eukaryote origins has changed the nature of the problem concerning the origin of sex in some salient respects. Current data on eukaryote origin have it that the host for the origin of mitochondria was an archaeon, not a eukaryote (Cox et al. 2008; Lane and Martin 2010; Williams et al. 2013; Spang et al. 2015). That would in turn suggest that the evolutionary inventions that distinguish eukaryotes from prokaryotes (including sex) arose in a prokaryotic (archaeal) host cell that possessed a mitochondrial (bacterial) symbiont, providing good reasons to doubt that the cell that acquired the mitochondrion was even mitotic at the time that the mitochondrion became established. There are a growing number of reports that implicate a role for mitochondria at the origin of sex (Lane and Martin 2010; Hadjivasiliou et al. 2013; Radzvilavicius and Blackstone 2015; Lane 2015). The basic idea that mitochondria came before sex (Radzvilavicius and Blackstone 2015) is worth exploring.

In fact, one can probably even exclude the possibility that a mitosing cell arose in the absence of mitochondria. How so? A short calculation is insightful. All eukaryotes separate their chromosomes with the help of microtubules. A tubulin dimer has 110 kDa (Oakley 2000), corresponding to about 1,000 amino acids, each of the peptide bonds requiring four ATP for polymer formation (Stouthamer 1978), or 4,000 ATP per dimer. A microtubule filament has about 13 dimers for one 360° turn in a 25-nm filament, the turn covering about 10 nm length, such that 10 nm of microtubule requires about 50,000 ATP for its synthesis (Nogales 2000). If a eukaryote cell is 10 μm long, and the microtubule has to go end to end, that corresponds to 50 million ATP to make one microtubule. If there is one microtubule per centromere (as in yeast), the cell can move one chromosome at a cost of 50 million ATP, the cost of microtubule depolymerization being 1/1,000th that of making the tubulin. If there are ten microtubules per centromere, as in many eukaryotes, we need 0.5 billion ATP to move a chromosome. If we have ten chromosomes per cell, we are at 5 billion ATP to move the chromosomes, but only if every single microtubule hits/attaches to a centromere, which does not happen—maybe 1–10% of the microtubules formed during mitotic cell division actually hit centromeres. That puts us at about 50–500 billion ATP to divide ten chromosomes in a 10-μm cell (or about 5–50 billion ATP to divide one chromosome)—in the modern, highly refined and regulated process. At the onset of microtubule-dependent chromosome segregation (MDCS), when the process was still primitive and improving via purifying selection, the cost of chromosome segregation was probably much higher. How much is 50–500 billion ATP? For comparison, *E. coli* needs a total of about 10–20 billion ATP per cell division (Neidhardt et al. 1990) to synthesize the daughter cell, to physically divide and to keep both cells alive during the process. Similarly, the amount of ATP that ancestral mitochondrial endosymbionts could make available to their host, the nascent eukaryote, simply by not synthesizing 5% unneeded proteins (such as for cell wall and the like) also comes in at about 50 billion ATP, but 50 billion per day (Lane 2014). Such calculations serve to highlight the amount of ATP required by eukaryotic cell biological processes and how mitochondria could contribute to these energetic needs.

Thus, the ancestral eukaryotic cell, the one that learned to divide its chromosomes using microtubules, expended as much ATP to merely segregate one prefabricated chromosome as normal prokaryotes expend to generate an entire daughter cell. This strongly suggests that the cell that learned to segregate chromosomes with microtubules had mitochondrial power (Lane and Martin 2010)—an inference that is consistent with, but independent of, data on the antiquity of mitochondria (McInerney et al. 2014; Ku et al. 2015) and the archaeal ancestry of the host (Cox et al. 2008; Lane and Martin 2010; Williams et al. 2013; Raymann et al. 2015; Spang et al. 2015). Although it should be mentioned that there are criticisms of the idea that mitochondrial power was important at eukaryote origin (Booth and Doolittle 2015; Lynch and Marinov 2015), it should also be mentioned that those criticisms have their own criticisms (Lane and Martin 2015, 2016). It should also be mentioned that mitochondria were not only a source of innovation, but they also caused problems (Blackstone 2013): Having a foreign cell in one's cytosol is a great perturbation in the day-to-day life of any viable prokaryotic cell. Yet by virtually any measure, it is increasingly clear that mitochondria played an important role at eukaryote origin; indeed, even more genes in eukaryotes stem from the ancestral mitochondrion than stem from the

archaeal host (Esser et al. 2004; Ku et al. 2015). Though views on eukaryote origin have changed radically in recent years (Martin and Müller 1998; Cox et al. 2008; Lane and Martin 2010; Katz 2012; Williams et al. 2013; McInerney et al. 2014; Raymann et al. 2015; Spang et al. 2015; Ku et al. 2015), views on the origin of sex have not at all kept pace with that development.

Sex: Embedded in the Cell Cycle and Dependent Upon Energy

As outlined in figure 1, meiosis and mitosis are just part of a more general process at the heart of eukaryotic cell growth and survival: The eukaryotic cell cycle. The cell cycle is, in turn, itself embedded in an even more general process: Carbon and energy metabolism, which run all processes of the cell to begin with. Without carbon and energy, no cell can survive, and no evolution can take place. Carbon and energy govern the immediate survival of the individual, hence its ability to evolve. Energy means ATP synthesis and is the first limiting factor for evolution. Without ATP synthesis, life and evolution come to an immediate halt. Population genetic effects operate within generations, bioenergetic effects operate within minutes.

The typical eukaryotic cell cycle comprises two major stages: The interphase and the mitotic phase or M-phase (fig. 1*A*) (Mitchison 1971). Interphase is further separated into the synthesis (S-) phase, during which the genome is replicated, and the gap (G-) phases, G1 and G2 (Norbury and Nurse 1992). During G1, the cell is metabolically active and prepares for genome replication, sensing the favorability of the environment; DNA replication during S phase follows. During G2, the cell carefully checks the integrity of its genetic material and prepares for mitosis (Norbury and Nurse 1992). M-phase or mitosis entails chromosome segregation followed by cell division or cytokinesis marking the end of the cell cycle. Many variations on this theme exist, for example, the presence or absence of the nuclear envelope in closed and open mitosis (Raikov 1994) or other variants such as cell senescence or specialization where the cells enter into a stage where they cease dividing (Blagosklonny 2011).

Meiosis, also called reduction division (because ploidy is reduced), can also be seen as a part of the cell cycle that is manifested in conjunction with sexual reproduction (Solari 2002). In this article, we use the term "sex" to designate a process in which the nucleus-bound genomes of two parents are brought together in a common cytoplasm (syngamy), whereupon the nuclei eventually fuse (karyogamy) to produce a cell with a double set of chromosomes, which will eventually undergo meiotic recombination and reduction, giving rise to progeny (meiospores) with a single (haploid) set of chromosomes that contain reassorted portions of the parental genomes (Bernstein et al. 1984) starting from a diploid cell. Meiosis typically results in four haploid cells (meiospores or gametes, depending upon how the organism undergoes

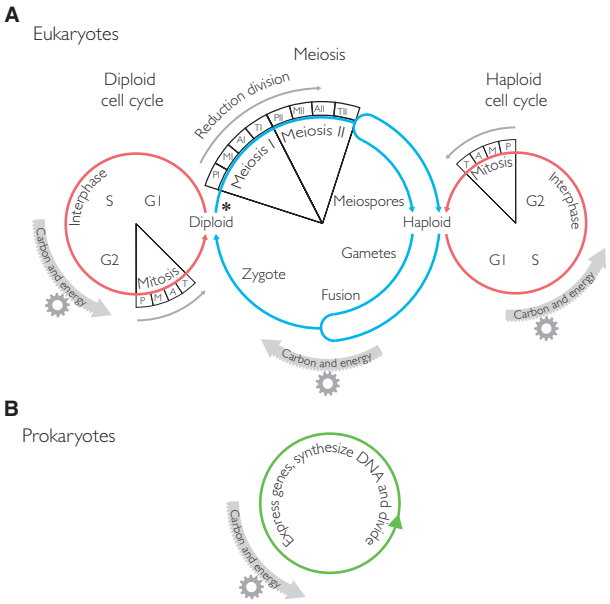


Fig. 1.—(***A***) Cell cycles and life cycles in eukaryotes. Meiosis (reduction division) connects haploid and diploid cell cycles. Each cell cycle is divided into distinct phases G (*gap*), S (*synthesis*) and the meiotic or mitotic phase, within which P (*prophase*), M (*metaphase*), A (*anaphase*), and T (*telophase*) phases are distinguished (see text). Carbon and energy are required to drive all cellular processes, including these cycles, upon which population genetic effects can subsequently operate. The asterisk indicates where, roughly, we would place the starting point in the origin of the process, that is, the symbiotic merger of host and symbiont. (***B***) Cell cycles and life cycles in prokaryotes, schematic. We recognize that there are spore-forming types, stalk-forming types, heterocyst-forming types, and other exceptions among prokaryotes. But in the main, we find it fair to generalize that the prokaryotic life cycle, if compared with the eukaryotic state, is a matter of continuously simultaneous cell and chromosome division.

alternation of generations) whose chromosomes have been reassorted and recombined relative to the mother cell, whereas mitosis yields two haploid or diploid cells, again depending upon alternation of generations, that have identical genomes.

Meiosis begins with duplicated sister chromatids paired with their homologous counterparts. Double-strand breaks (DSBs) initiate recombination. After crossovers are resolved, the sister chromatids are independently assorted. Following cytokinesis, a second round of chromosome segregation without DNA replication ensues, sister chromatids are distributed to the daughter cells, typically haploid gametes. Fusion of two such meiotically generated gametes, which may have a prolonged mitotic life cycle of their own (fig. 1*A*), eventually restores diploidy (Wilkins and Holliday 2008). Importantly, the transition from one cell cycle phase to another is a highly controlled process involving cyclin-dependent kinases (CDKs), which are temporally regulated by their respective cyclins (Morgan 1997). These act as checkpoints that ensure

faithful chromosome replication and segregation are followed by cell division (Hartwell and Weinert 1989), initiating mitosis at its specific phase during the cell cycle.

Prokaryotes clearly have recombination (Camerini-Otero and Hsieh 1995; Papke et al. 2004; Naor and Gophna 2014). But meiosis, mitosis, and a eukaryotic type cell cycle are lacking in prokaryotes altogether (fig. 1*B*). Some might counter that prokaryotes have mitosis or that archaea have a cell cycle (Lindås and Bernander 2013) but these are misnomers: Prokaryotes do not present anything resembling bona fide mitosis (chromosome condensation and microtubule-dependent chromosome separation to nuclear poles or cell poles) that resides at the heart the eukaryotic cell division process, nor do prokaryotes have anything that could be viewed as faintly homologous to the eukaryotic cell cycle, of which mitosis is a part (fig. 1). Archaeal chromosomes typically have multiple origins of replication, requiring a bit more care to ensure proper chromosomal partitioning (Lindås and Bernander 2013), yes. But a eukaryote-like cell cycle? Hardly. The eukaryotic cell cycle has no homolog among prokaryotes.

Reinspecting Old Premises

Compared with literature on the origin of sex or the origin of eukaryotes, literature concerning the evolution of the cell cycle is fairly scarce, with Nasmyth (1995), Novak et al. (1998) and Cross et al. (2011) being notable exceptions, though they do not specifically address cell cycle origin. Literature covering all three topics in one place is scarcer still, Cavalier-Smith's essays (2002, 2010) being exceptions. Yet, like de Duve (2007) did in his day, Cavalier-Smith (2002, 2010, 2014) still rejects the idea that archaea participated in any way in the origin of eukaryotic lineage, steadfastly maintaining that both eukaryotes and archaea arose from actinobacteria. That makes it virtually impossible to integrate his views into any kind of modern synthesis, because phylogenetic analyses indicate that the host for the origin of mitochondria was an archaeon (Cox et al. 2008; Williams et al. 2013; McInerney et al. 2014; Raymann et al. 2015; Spang et al. 2015) and that the ancestor of the mitochondrion was an alphaproteobacterium, with no evidence for other partners at eukaryote origin (Ku et al. 2015). Genomes harbor evidence neither for an actinobacterial origin of eukaryotes (Ku et al. 2015) nor for an actinobacterial origin of archaea (Nelson-Sathi et al. 2015). Hence we acknowledge Cavalier-Smith's contributions to the topic, but address no specifics of his hypotheses regarding actinobacterial origins of archaea, phagotrophy, eukaryotes, or sex (Cavalier-smith 1975, 2014). Not surprisingly, the phylogeny of the proteins involved trace the cell cycle to the eukaryote common ancestor (Krylov et al. 2003).

If meiosis evolved from mitosis, as traditional theories for the origin of sex posit, it arose in some hypothetical lineage of asexual, mitosing eukaryotes that, like all lineages, had to escape Muller's ratchet. Therefore, it utilized either 1) the well-characterized prokaryotic mechanisms of getting DNA into the cell for recombination (the typical prokaryotic LGT mechanisms), or 2) some mechanism of recombination that was compatible with mitosis but did not involve meiosis. In either case, the lineage was necessarily recombining (to avoid Muller's ratchet, we contend), leaving neither selective pressure to evolve anything as complicated as meiosis and sex, nor benefit from it once it arose. This line of thought actually renders the origin of meiosis from mitosis altogether unlikely.

Ordering Events at the Prokaryote–Eukaryote Transition

The eukaryote ancestor had mitochondria, a nucleus, a cell cycle and sex. In what order did these traits arise? Traditional theories holding that meiosis evolved from mitosis, also entail the assumption that mitochondria had nothing to do with the origin of eukaryote complexity. Yet from the energetic standpoint, mitochondria had everything to do with the origin of eukaryote complexity (Lane and Martin 2010), and several recent publications even report how genetic effects emanating from mitochondria could have impacted the origin of sex (Lane 2009; Hörandl and Hadacek 2013; Lane 2014; Havird et al. 2015; Radzvilavicius and Blackstone 2015; Speijer 2015). Cleveland (1947) clearly considered meiosis in the context of the life cycle. Our approach is similar. In eukaryotic microbes, whose common ancestor possessed mitochondria (Embley and Martin 2006), the life cycle is an iteration of the cell cycle (fig. 1). In ordering the sequences of events surrounding the origin of six key characters (mitochondria, the nucleus, meiosis, mitosis, the cell cycle, and sex), we start with minimal premises: A mitochondrial endosymbiont in an archaeal host that lacked the other five traits. For further justification of why we embark from such a simple cell biological starting point, see Gould et al. (2016) with regard to the origin of the endomembrane system and Sousa et al. (2016) with regard to the archaeal nature of the host.

Starting with Carbon and Energy

We start with an endosymbiosis, an archaeon (hereafter called the host) that acquired a bacterial endosymbiont, the common ancestor of mitochondria and hydrogenosomes (hereafter called the symbiont). Various prokaryotes harbor prokaryotic endosymbionts (Wujek 1979; von Dohlen et al. 2001; McCutcheon et al. 2009; Husnik et al. 2013; Kobialka et al. 2016), in cases investigated so far, the symbiotic interactions are metabolic and the host is not phagocytotic. At eukaryote origin, metabolic interactions (Martin and Müller 1998; Searcy 2003; Müller et al. 2012; Degli Esposti 2014) likely facilitated interactions between the mitochondrial symbiont and its host (fig. 2*A*), for which we posit an archaeal

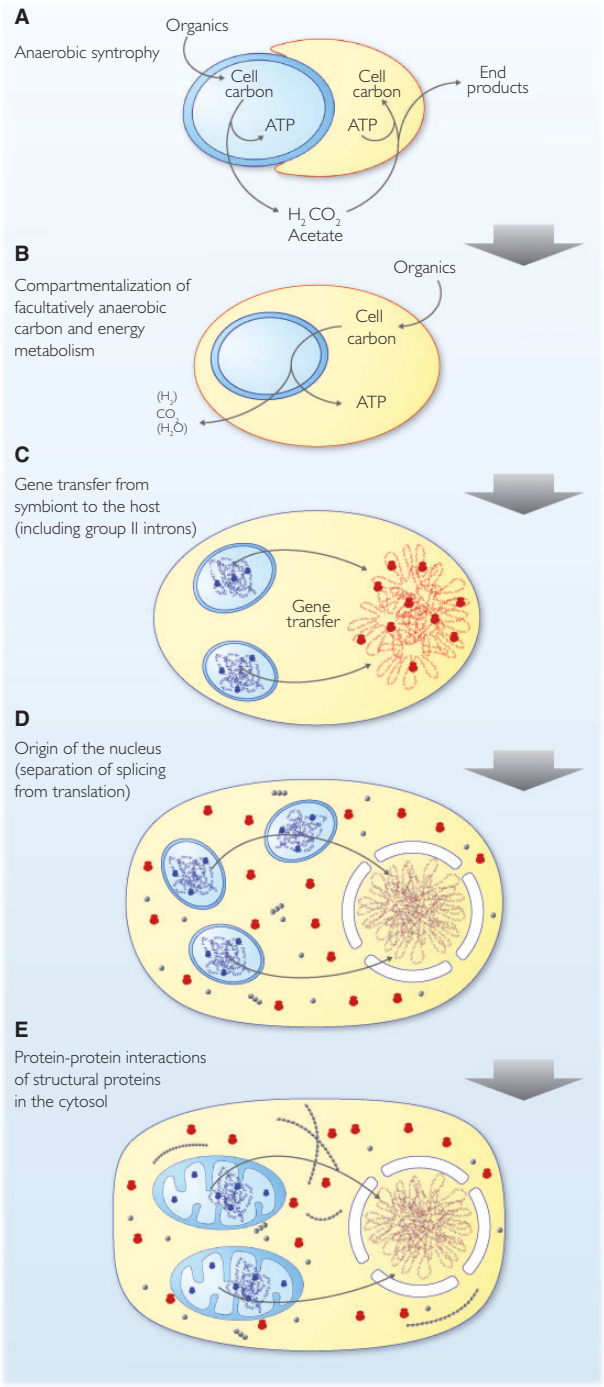


Fig. 2.—Steps en route from endosymbiont acquisition to an ancestral cell cycle. Blue, red and gray represent bacterial, archaeal and eukaryotic components, respectively (see text).

cellular organization. The symbiosis must be stable, neither partner digesting the other, and with a suitable metabolic flux fueling both partners, for example, anaerobic syntrophy with a hydrogen-dependent host (Sousa et al. 2016).

Stable symbiosis requires carbon and energy for both partners, which anaerobic syntrophy can provide. Eukaryotes conserve energy in the cytosol and in internalized bioenergetic organelles—mitochondria. The transfer of genes from the mitochondrial endosymbiont, a facultative anaerobe (Müller et al. 2012; Degli Esposti 2014) to the chromosomes of the host prior to the origin the mitochondrial protein import apparatus can account both for 1) the bacterial origin of the eukaryotic glycolytic pathway and 2) its cytosolic localization (Martin and Müller 1998). With importers for organic compounds in the host's plasma membrane and a glycolytic pathway in the cytosol, the host compartment (the cytosol) has a source of net ATP synthesis (glycolysis) that is independent of chemiosmotic coupling at the plasma membrane, which is lost. To provide net ATP yield in the cytosol, glycolysis must proceed to the pyruvate-generating step. Transfer of the symbiont's glycolytic pathway to the host's cytosol and its carbon substrate importers to the host's plasma membrane does not require inventions, it merely requires the transfer of genes from symbiont to host and the expression of bacterial genes in archaeal chromosomes (Martin and Müller 1998), evidence for which in modern archaeal genomes abounds (Nelson-Sathi et al. 2015).

The endosymbiont's transition into an ATP-exporting organelle requires two things: 1) Ability to import pyruvate—which traverses membranes readily (Bakker and Van Dam 1974)—and oxidize it, yielding approximately five ATP per glucose (Müller et al. 2012) anaerobically or approximately 30 ATP per glucose aerobically (Rich and Maréchal 2010); and 2) ability to export ATP to the cytosol through the mitochondrial ADP/ATP carrier (AAC) (Whatley and Whatley 1979; Radzvilavicius and Blackstone 2015). Radzvilavicius and Blackstone (2015) have suggested that the AAC might even have been invented in, and originally expressed by, the symbiont's genome, requiring no new protein import machinery, only the preexisting protein insertion machinery of the symbiont (fig. 2*B* and *C*), an intriguing idea.

From the energetic standpoint, the AAC consummates the symbiont-to-organelle transition (Whatley and Whatley 1979; Radzvilavicius and Blackstone 2015), although the invention of the mitochondrial protein import machinery—Translocase of the outer membrane (TOM) and Translocase of the Inner Membrane (TIM) complexes (Doležal et al. 2006)—and targeting signals on cytosolic precursor proteins (Garg et al. 2015) allows the symbiont to relinquish genes to the host's chromosomes. Not all genes are relinquished however, those central to the electron transport chain in the inner mitochondrial membrane remain in the organelle, for reasons of redox balance (Allen 2015).

The mitochondrion alters the basic bioenergetic architecture of the cell and the amount of protein that the host compartment (cytosol) can afford to express (Lane and Martin 2010; Lane 2014). The conversion of an alphaproteobacterial endosymbiont with a heterogeneous genome (Martin 1999;

Ku et al. 2015) into an ATP generating compartment furthermore coincided with the complete loss of chemiosmotic energy conservation (ATP synthesis) at the host's plasma membrane (Gould et al. 2016). Eukaryotes have archaeal ribosomes in the cytosol, but the enzymes of eukaryotic carbon and energy metabolism stem from bacteria and trace to the eukaryote common ancestor (Blackstone 2013; Ku et al. 2015). As we see it, mitochondria change not only the bioenergetic state of the host compartment (Lane and Martin 2010) but also the physical content of the host's cytosol through the addition of membrane vesicles consisting of bacterial lipids: Outer membrane vesicles produced by the mitochondrial symbiont (Gould et al. 2016). Both contributions of mitochondria, we contend, carried dramatic consequences for eukaryote evolution and account for the observation that only the cells that became genuinely complex have mitochondria or had them in their past.

For balance, we note here that some readers might not agree with the foregoing proposition. For example, population geneticists contend that mitochondria had nothing whatsoever to do with eukaryote origin (Lynch and Marinov 2015). The philosophically inclined might argue that eukaryote complexity, if it is not an illusion altogether, is not due to mitochondria at all, but to luck (Booth and Doolittle 2015). It is not our intent here to try to convince critics. There is certainly a role for population genetics at eukaryote origin, namely the small population size, increased drift, and reduced power of purifying selection (Nei 1987) obviously inherent to the single origin of both mitochondria and of eukaryotes as a group. But we think that the main hurdle at eukaryote origin is the origin of mitochondria, which (like plastids) is the result of endosymbiosis, not allele frequency changes. All organisms have population genetics, only eukaryotes have mitochondria. Is there a connection between mitochondria and complexity? We think so. Since its inception (Mereschkowsky 1905), endosymbiotic theory has always been a target of disparaging critique. Thus, we acknowledge criticisms divesting mitochondria (and endosymbiosis more generally) of evolutionary significance, and move on.

Gene Transfer, Introns, the Nucleus and Ploidy

All cells that have sex have a nucleus. How does the nucleus fit into eukaryote origin? The mitochondrial endosymbiont is more than a source of energy, it is a source of genes, lots of genes, and large scale chromosomal mutations. This process of gene transfer, from the bacterial symbiont to chromosomes of the archaeal host, is called endosymbiotic gene transfer, or EGT (Martin et al. 1993; Timmis et al. 2004), it is unidirectional, it still operates today involving insertion of whole organelle genomes hundreds of kilobases in length into nuclear DNA (Huang et al. 2005). The mechanism of DNA integration is nonhomologous end joining (Hazkani-Covo and Covo 2008), the mechanism of DNA release to the host is organelle lysis (Huang et al. 2004), the process of EGT has operated throughout eukaryotic history and is observable as an ongoing process, even during human evolution, with the most recent mitochondrion-to-nucleus transfers dating to the Tschernobyl incident (Hazkani-Covo et al. 2010). Eukaryotes are usually described as descendants of archaea (Cox et al. 2008; Williams et al. 2013), but if we look at the whole genome, bacterial genes vastly outnumber archaeal genes in eukaryotes (Esser et al. 2004; Thiergart et al. 2012), and genes that trace to the mitochondrion vastly outnumber those that trace to the host (Ku et al. 2015). At the outset, the host has no nucleus, and as long as cell division is not impaired, the symbiosis of prokaryotes is stable, as long as the environment supports growth.

The stability of the symbiosis changes however, probably as a consequence of EGT, a mutational mechanism that is specific to the eukaryotic lineage: Gene transfer from symbiont to host carries some fateful hitchhikers—self splicing group II introns. Group II introns are important, and their transition into spliceosomal introns could have precipitated the origin of the nucleus (Martin and Koonin 2006). How so? Group II introns occur in prokaryotic genomes (Lambowitz and Zimmerly 2011), they are mobile, they can spread to many copies per genome (Lambowitz and Zimmerly 2004) and they remove themselves through a self-splicing mechanism that involves the intron-encoded maturase (Matsuura et al. 1997). Their splicing mechanism is similar to that in spliceosomal intron removal (Lynch and Richardson 2002), for which reason they have long been viewed as the precursors of both 1) spliceosomal introns and 2) their cognate snRNAs in the spliceosome (Sharp 1985).

The crux of the intron hypothesis for nuclear origin (Martin and Koonin 2006) is that group II introns, which are mobile elements in prokaryotes (Lambowitz and Zimmerly 2011), entered the eukaryotic lineage through gene transfer from the mitochondrial endosymbiont to the archaeal host. In the host's chromosomes they spread to many sites and underwent the transition to spliceosomal introns, as evidenced by the observation that many introns are located at conserved positions across eukaryotic supergroups (Rogozin et al. 2003) and by the presence of spliceosomes in the last eukaryote common ancestor (Collins and Penny 2005). The transition from group II introns to spliceosomal introns evokes a curious situation: Spliceosomal splicing is slow, on the order of minutes per intron (Audibert et al. 2002), whereas translation in ribosomes is fast, on the order of 10 peptide bonds per second (Sørensen et al. 1989). As the transition to spliceosomal introns set in, the host's cytosol was still a prokaryotic compartment with cotranscriptional translation. With the origin of bona fide spliceosomes and spliceosomal splicing, nascent transcripts were being translated (ribosomes are fast) before they can be spliced (spliceosomes are slow). Translation of introns leads to defective gene expression at many loci

simultaneously (though one essential locus would suffice), a lethal condition for the host unless immediately remedied.

The solution to this condition was, we posit, physical separation of the slow process of splicing from the fast process of translation so that the former could go to completion before the latter set in. Separation in cells usually involves membranes, and that is the central tenet of the intron hypothesis: The initial pressure that led to selection for nucleus–cytosol compartmentation (the origin of the nuclear membrane) was the requirement for physical exclusion of active ribosomes from nascent transcripts, to restore gene expression and intron-containing genes (fig. 2*D*). The primordial nuclear membrane allowed the slow process of splicing to go to completion around the chromosomes, thereby initially allowing distal diffusion, later specific export of processed mRNAs to the cytosol for translation, furthermore precipitating the origin of nonsense-mediated decay, a eukaryote-specific machinery that recognizes and inactivates intron-containing mRNAs in the cytosol (Martin and Koonin 2006) (fig. 2*D*).

The reader might protest that we have specified neither a mechanism nor a source for the vesicles that give rise to the nuclear membrane in the host's archaeal cytosol. That is the topic of a separate paper (Gould et al. 2016), in which we outline how outer membrane vesicles produced by the mitochondrial endosymbiont in an archaeal host are likely both the physical source and the evolutionary origin of the eukaryotic endomembrane system.

A primitive nuclear membrane rescues gene expression, and DNA replication can continue to proceed as long as the cytosol supplies dNTP precursors. But there is no mechanism for chromosome segregation in place. Chromosomes replicate without division, polyploidy, extreme polyploidy in all likelihood ensues, and the symbiosis seems to be headed straight toward a dead end. But mitochondria can make a difference.

Protein–Protein Interactions in an Energy-Laden Cytosol

All cells that undergo sex divide their chromosomes with microtubules. Prokaryotes possess genes for tubulin precursors (Erickson 2007), but they do not make microtubules. Why not? Introns give rise to a cell that requires a nuclear membrane to express genes. That configuration is fine from the standpoint of stable gene expression to maintain carbon and energy metabolism. But sequestration of the host's chromosomes within a nuclear compartment has two consequences of exceptional significance. First, though the nuclear membrane rescues gene expression, the chromosomes are no longer attached to the plasma membrane of the cell and segregation of the chromosomes (now contained within the nucleus) is no longer coupled to cell division. This is a problem of severe sorts. Our symbiotic consortium can satisfy its carbon and energy needs by virtue of compartmentalized carbon and energy metabolism between the cytosol and the mitochondrion. It can express intron-containing genes by virtue of a nuclear membrane, but it cannot segregate its chromosomes in the standard prokaryotic manner to produce progeny. Either a solution to the problem of chromosome partitioning is found or extinction is the alternative. The solution to chromosome partitioning stems, we propose, from the second consequence of nucleus–cytosol compartmentation.

The second consequence is that the nuclear membrane generates a fundamentally new kind of cell compartment in the biological realm of that day: A cytosol that is free of active chromatin. The eukaryotic cytosol is not only a compartment of protein–protein interactions (Martin and Koonin 2006), it can afford, energetically, to express the proteins that might interact (Lane and Martin 2010). The eukaryotic cytosol is unique in that it is a dedicated translation compartment where protein–protein interactions can take place at a magnitude never before possible in any prokaryotic cell (fig. 2*E*). Energy is crucial for that, because protein synthesis consumes about 75% of a cell's energy budget (Harold 1986). It is also true that for the world of proteinprotein interactions that emerged in the eukaryotic cytosol to materialize, orders of magnitude more ribosomes than typical of a prokaryotic cytosol need to be synthesized. This requires amplification of rDNA genes, which eukaryotes realize by various means, including the increase of rDNA genes to thousands of chromosomal copies (McGrath and Katz 2004), and the specific amplification of rDNA genes through rolling circle plasmids and other extrachromosomal elements in various eukaryotic lineages (Hourcade and Dressler 1973; McGrath and Katz 2004; Kobayashi 2011). Providing the cytosol with abundant protein requires not only abundant ATP but also very large numbers of ribosomes.

In addition to having a chromatin-free cytosol, hence a dedicated translation compartment, the stem eukaryote has, by virtue of mitochondria (Lane and Martin 2010), effectively unlimited ATP for protein synthesis. So despite being unable to divide in a well-coordinated manner, it can synthesize proteins (and ribosomes) in amounts unattainable by any prokaryotic cell, because of mitochondrial ATP synthesis. This enables the symbiotic consortium (the nascent eukaryote) to explore protein expression in a manner that no prokaryote could. Relative to prokaryotes, the existence of mitochondria in the nascent stem eukaryote enables energetically unpenalized protein overexpression. Mitochondria do not force an evolutionary transition, but they enable it. The stem eukaryote can overexpress virtually every protein, so countless protein expression experiments are possible.

Given the watchful eye of natural selection, which expression experiments might be successful? That is, which proteins might become expressed at high amounts? Massive overexpression of metabolic enzymes is not a viable option, as it will inevitably impair carbon and energy flux. In contrast, overexpression of enzymatically inert structural proteins such as tubulin, actin, and other structural proteins typical of the eukaryotic cytosol, but in their ancestral prokaryotic forms

(FtsZ, MreB, Ta0583, CetZ, archaeal Cdv's, the precursors of ESCRT complex proteins, etc.), will not alter metabolism. Expression of structural proteins will simply sink carbon and nitrogen into proteins that 1) can accumulate without interfering with carbon flux, and 2) that spontaneously assemble into higher order structures (Jékely 2014) while actually requiring ATP hydrolysis for their disaggregation. Cytoskeletal proteins aggregate spontaneously and consume ATP to depolymerize or disaggregate (Fleury-Aubusson 2003; Gould et al. 2011).

This is particularly interesting because it suggests that the origin of filamentous or otherwise aggregated cytoskeletal components was not a slow, stepwise evolutionary process, but rather that it was a spontaneous consequence of dramatically increased ATP availability (for protein synthesis), requiring a small additional supply for disassembly into monomers. A general underlying theme of eukaryotic cytoskeletal proteins is that—if synthesized in sufficient amounts—they spontaneously assemble into larger, ordered structures and require ATP hydrolysis for their disassembly or depolymerization. Cytoskeletal proteins could thus be seen as relicts of ancient overexpression experiments in which selection was acting to bring forth polymers that could undergo reversible self-assembly. Such experiments might also still be going on today, as this would explain why the intermediate filament proteins of various protist lineages seem not to share common ancestry with the intermediate filament proteins of animals and fungi (Gould et al. 2011), having arisen independently instead.

In short, at this stage in the prokaryote to eukaryote transition, the cytosol expresses proteins that make structures and move things through ATP and GTP, as opposed to converting substrates. ATP has to be in very abundant supply for that, otherwise the proteins could not be synthesized. At such a stage, the spectrum of eukaryotic-specific cytological novelties could have taken root in terms of becoming heritable and fixed. But there are still some unsolved problems with the process of heredity: Chromosome segregation.

Chromosome Division

Overexpressing cytoskeletal and other (novel) structural proteins is now an energetically affordable option for the mitochondrion bearing cell. Prokaryotic tubulin precursors are very similar in form and function to their eukaryotic counterparts (Erickson 2007), hence little in the way of protein sequence modification is required for tubulin monomers to assume new function, but rather the limitation is synthesizing them in large amounts. Synthesizing large amounts of tubulin monomers results in microtubules, and unconstrained (unregulated) polymerization of microtubules leads to a cytosol teeming with spontaneously polymerizing and GTP-dependent (ultimately ATP-dependent) depolymerizing microtubules: A world of molecular movement bearing the possibility of MDCS (fig. 3*A*). Obviously, chromosome movement requires attachment sites on the chromosomes: Primitive centromeric regions. Attachment sites are not a completely novel eukaryotic invention, because prokaryotic chromosomes attach to the plasma membrane, allowing them to be separated at cell division (Toro and Shapiro 2010). Indeed, the archaeal protein that serves as an attachment site for large plasmid segregation in *Sulfolobus*, ParB, is similar at the structural level to CenpA, which is a pivotal protein of eukaryotic microtubule chromosome attachment and chromosome segregation (Schumacher et al. 2015). So the basic machinery for attaching to DNA that was not bound to the plasma membrane (plasmids) was apparently in place in the host, and prokaryotic protein attachment sites for ParB-dependent segregation are present in prokaryotes (Mierzejewska and Jagura-Burdzy 2012), such that the initial process of physically segregating DNA with microtubules possibly hinged more upon merely being able to synthesize enough tubulin to get the job done than it did on evolving an orchestrated chromosome choreography. Synthesizing large amounts of protein required mitochondria.

Primitive MDCS likely occurred in the persistent presence of a nucleus (a closed mitosis state), because of the continued need to separate splicing from translation. The host's chromosomes could interact with the inner leaf of the nuclear membrane through pre-existing chromosome-membrane attachment mechanisms (Toro and Shapiro 2010), whereas the microtubules could interact with primitive nuclear pore complexes (Zuccolo et al. 2007) that permitted diffusion of spliced mRNA from the nucleus to the cytosol, but excluded the diffusion of active ribosomes from the nucleus. Alternatively, microtubules might simply have formed within the nucleoplasm, attaching to chromosomes directly and pushing them and the nucleus apart in an ATP-dependent manner, as it occurs in *Vaucheria* (Takahashi et al. 2003), *Bryopsis* (McNaughton and Goff 1990), or diatoms (Pickett-Heaps et al. 1982; Cande and McDonald 1985). At the outset, some combination of both is not unlikely.

MDCS provides a means for, and leads to, unregulated division of continuously replicating chromosomes. But given the tools and the energy, chromosome segregation does not know when to stop (!) because there is neither a cell cycle nor a coordinated mitotic division process. Thus, the ability to move chromosomes apart, while initially en route to becoming a virtue, suddenly becomes a horrible vice: Microtubules continuously separate chromosomes, down to a state where no more segregation is possible (perhaps one chromosome or plasmid per nucleus).

Before going further, a short note about forces in chromosome segregation is in order. When we say MDCS, the reader might think that we mean the pulling apart of chromosomes via attachment of microtubuli to microtubule organizing centers at poles of the cell, or similar. Pulling is not what we have in mind. The simpler process of pushing chromosomes apart is what we have in mind. At this point in our inference, the cell

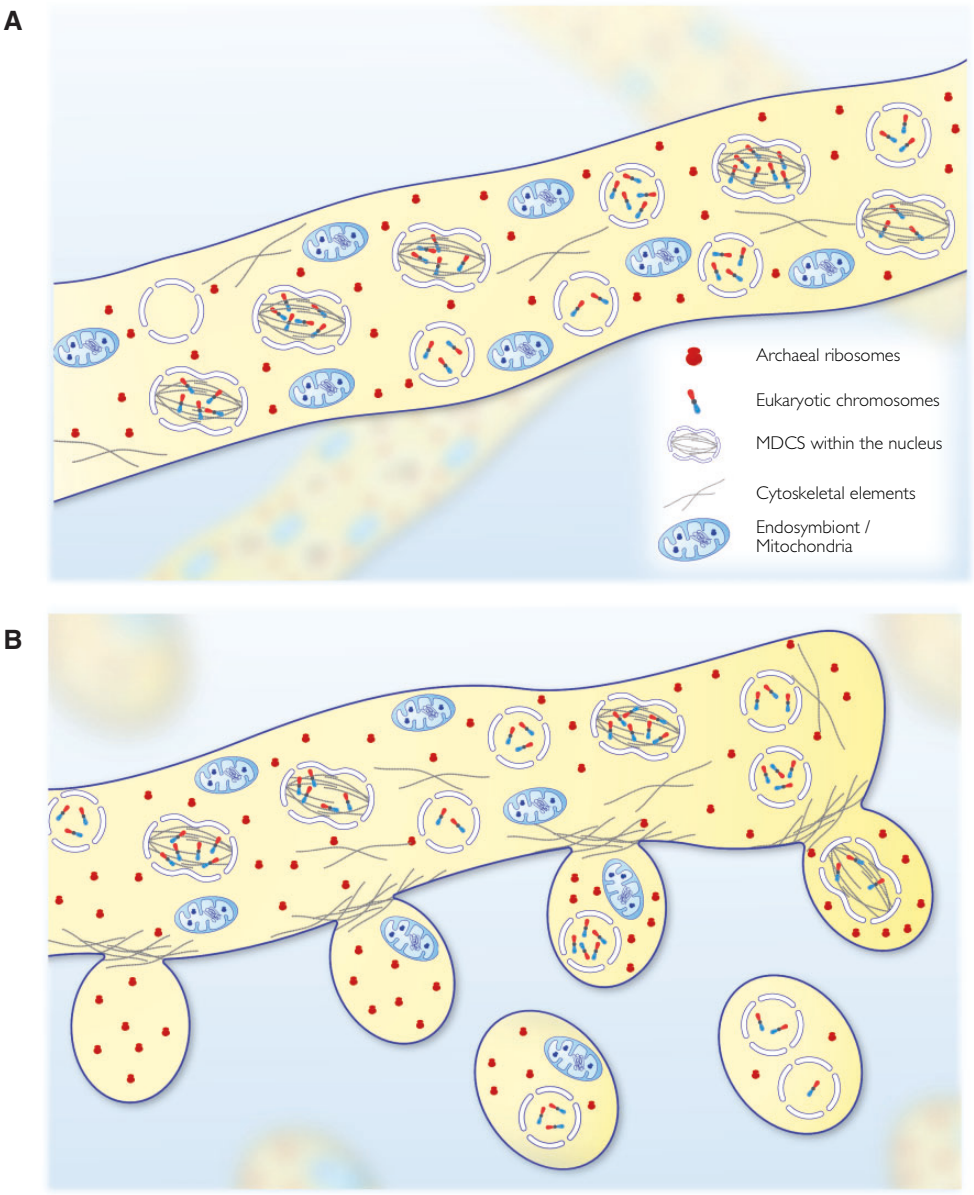


Fig. 3.—(*A*) A coenocytic eukaryote common ancestor(Coeca) with multiple independently dividing nuclei. (*B*) Accidental budding off of spores using the archaeal vesicle secretion mechanisms. Only spores containing (at least) a complete genome enclosed within a nucleus and at least one (compatible) mitochondrion are viable. All other possibilities result in inviable spores, providing strong selection among spores for viable gene, chromosome and mitonuclear combinations.

has a nucleus but neither mitosis nor a cell cycle, so chromosome segregation and nuclear division are the issues, not coordinating nuclear division with cell division. In several lineages of eukaryotes that maintain their nuclear membrane intact at chromosome division, the chromosomes and nuclei are pushed apart by microtubules. This occurs in coenocytic eukaryotes such as *Vaucheria* (Takahashi et al. 2003) or *Bryopsis* (McNaughton and Goff 1990), where the central rod-like microtubule structure that pushes the chromosomes and nuclei apart is called the interzonal spindle. It occurs in trichomonads, where the central rod-like microtubule structure that pushes the flagellar apparatuses (and ultimately the chromosomes) apart is designated either as the central spindle (Raikov 1994) or as the paradesmosis (Bricheux et al. 2007). Pushing also occurs in *Schizosachharomyces pombe*, where the spindle pole bodies located within the nucleus are pushed apart by microtubules (Castagnetti et al. 2015). In diatoms, the shape and behavior of the spindle were shown to be highly

suggestive of a pushing mechanism to separate chromosomes (Pickett-Heaps et al. 1982); purified, isolated diatom spindles were later directly shown to exert a pushing force in vitro upon addition of ATP (Cande and McDonald 1985). Thus, when we say chromosome segregation, we have an ATP-dependent pushing mechanism for chromosome segregation in mind. An overview of variation in mitotic types among eukaryotes is provided by Raikov (1994).

High ploidy (see Gene Transfer, Introns, the Nucleus and Ploidy section), which was inevitable before MDCS came into play, points to a possible reason why linear chromosomes, which the eukaryote ancestor certainly came to possess at some point, would be preferable or better suited to survival than circular chromosomes. If ploidy became high, linear chromosomes, which do occur in prokaryotes (Bentley et al. 2002), would be much easier to separate than circular chromosomes, which generate concatamers upon replication. Linear chromosomes would not require disentanglement of multiply replicated circles, and hence would appear advantageous for a primitive MDCS process.

This kind of chromosome division—a primitive microtubule-dependent division that is independent of plasma membrane movement and cell expansion and that segregates (linear) chromosomes out of heavily polyploid nuclei—is, we suggest, the evolutionary origin of reduction division, the cardinal event in meiosis. But because the segregation process does not know when to stop, the chromosome sets that emerge as products of this kind of uncontrolled reduction will strongly tend to lack sufficient chromosomes (or genes, or both) for stable heredity. At this stage, chromosomes have to be in nuclei for gene expression (because of spliceosomal splicing) and they can replicate without the need for invention, using preexisting prokaryotic replication machinery. Thanks to mitochondrial metabolic power, they can be pushed apart by microtubules in the presence of nuclei (corresponding to closed mitosis), but they are not pushed apart in a coordinated manner to start.

Cell Division? Things Actually Work Better without It

If (note the "if") there is cell division going on concomitant with this kind of primitive and crude chromosome segregation, then many, most or all of the progeny from this "emergency solution" or "evolutionary loophole" to chromosome separation will not capable of continued reproduction for lack of chromosome sets that would permit self-sufficient and self-sustaining replicating progeny. Daughter cells might inherit enough active cytosolic protein from lost genes to keep them viable for days or months, but not enough genes to keep a lineage going.

In addition, continuous gene transfer from symbiont to host (Timmis et al. 2004; Hazkani-Covo et al. 2010) generates archaeal host chromosomes that are, both within and across individuals, heterogeneous with respect to bacterial chromosome insertions (Lane 2009,2014). This generates disrupted genes, DSBs, and chromosomes that are rapidly evolving in terms of gene content. The future does not look bright for this population of energetically overachieving but genetically underdeveloped cells. Short of a miracle, is this inference going to go anywhere? Probably not, were it not for two observations that come into play, each of which can potentially contribute to solving some very hard problems surrounding the origin of the eukaryotic lifestyle:

1. If these cells divide, a very curious property of the archaeal host cell could rescue progeny: Archaeal cells can fuse. That archaeal cells fuse has been reported for the crenarchaeote *Sulfolobus* (Schleper et al. 1995), for several *Thermococcus* species (Kuwabara et al. 2005), and the euryarchaeote *Haloferax* (Naor and Gophna 2014). It is thus a property found within both the crenarchaeal and the euryarchaeal groups, hence attributable to our host without need for invention. The ability to fuse is a preexisting property of the host, and if not lost during earlier phases of the transition, now fulfills a lifesaving function: It creates new combinations of chromosomes, chromosomes that can be very different in number and nature (the products of uncontrolled reduction). Fusion could, in principle, lead to restoration of viable gene and chromosome numbers at this stage, but a lot of fusion would have to be going on: Fusion rates would have to be roughly the same as division rates in order to keep the system going. It is possible, but it is a long shot. The ability to fuse is probably more important when it comes to closing the life cycle (syngamy), dealt with in a later section. That modern eukaryotes can undergo fusion (plasmogamy), probably to generate recombination, is documented for some amoebae (Tekle et al. 2014).

2. If these cells do not divide, but just grow in length, the basic machinery of host cell division being somehow impaired but symbiont division—requiring dynamins (Purkanti and Thattai 2015) and ftsZ (Beech et al. 2000)—remaining intact, the result is a filamentous cell having nucleus-surrounded chromosomes that are segregated within the cytosol alongside autonomously dividing mitochondria. This cellular habit, the syncytium or coenocytic state, with many dividing nuclei and organelles occurs in several eukaryotic groups, being perhaps best-known among fungi (Roper et al. 2011), green algae (Verbruggen et al. 2009), and algae with red secondary plastids like *Vaucheria* (Gavrilova and Rundanova 1999).

The decisive advantage of a syncytial habit over host cell fusion at this stage is 2-fold: 1) In a syncytium, many different nuclei with deficient chromosome sets can "complement each other" simultaneously through mRNA in the cytosol, keeping the coenocyte alive; and 2) in a syncytium, nuclei can fuse (karyogamy: every eukaryote with sex does it) and divide. Chromosomes can undergo replication, recombination, segregation, and reduction, while remaining heavily buffered from selection because defects are rescued through

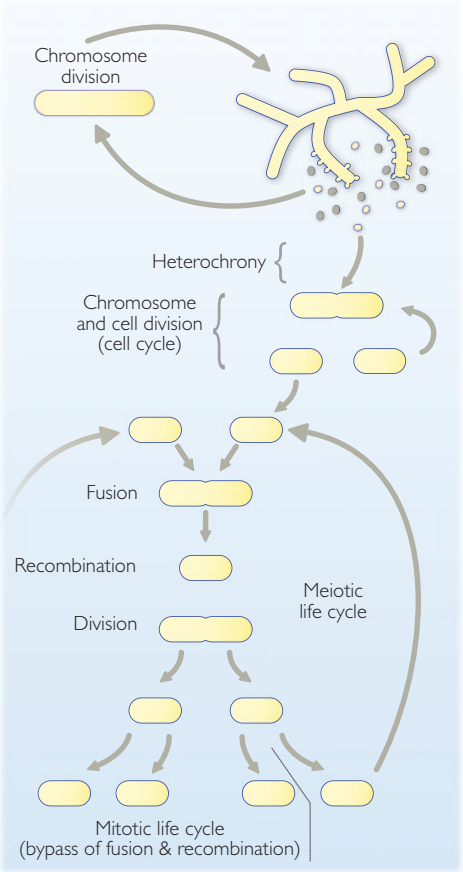


Fig. 4.—Possible life cycles of a coenocytic eukaryotic common ancestor. Viable meiospores (yellow) that bud off the syncytium have two possibilities. They can either 1) germinate to a new syncytium and continue with a coenocytic life cycle with multiple nuclear divisions or 2) undergo cell division (spore secretion) immediately after nuclear division, a case of heterochrony (spore secretion before filament formation). If this results in viable progeny, they can undergo fusion, recombination and division, which are characteristic of meiotic lifecycles. Mitotic life cycles can be easily derived from bypassing the fusion (syngamy, karyogamy) and recombination phases of a meiotic life cycle (see also fig. 1).

complementing mRNAs in the cytosol. A complete set of essential genes (or many sets thereof) can be expressed in the syncytium, but from chromosomes dispersed across many different individual nuclei, some (many?) containing perhaps only one chromosome. The concept of a syncytial eukaryote common ancestor (fig. 4*A*) has many virtues. Coexisting defective nuclei of the kind that we have in mind have been observed within contemporary syncytia among the charophytes (Hasitschka-Jenschke 1960).

A Syncytial Eukaryote Common Ancestor

Some readers might gasp at the seemingly radical notion of a syncytial eukaryote common ancestor. But the closer one inspects the idea, the more robust it appears. The reasons are as follows.

First, in a syncytium, there is no pressure to solve the evolutionary problems of inventing the very complex and very novel (relative to prokaryotes) eukaryotic solutions to coordinated chromosome division and coordinated cell division all at once. The syncytium not only allows defective or incomplete chromosome sets in individual nuclei to persist through mRNA-mediated complementation in the cytosol, but also allows primitive nuclear divisions and nuclear fusions (karyogamy) to generate continuously new and potentially useful chromosome combinations. In principle, a syncytium could become very long, possibly branched (many prokaryotes including cyanobacteria and actinomycetes can branch), and thus generate ample opportunity for selection through the physical separation of chromosome combinations. Because the cell wall is the host's, cell fusion as in archaea is a possibility that could, in principle, generate further combinations of fit nuclei at growing tips.

Second, the syncytial state is better suited to sorting out cytonuclear interactions than uninucleate cells. The genetic interactions between mitochondria and the nucleus have come under intense interest of late, not only because they are important for modern biology but also because they were likely important very early in eukaryote evolution, also at the origin of sex (Lane 2005,2009,2014; Havird et al. 2015; Radzvilavicius and Blackstone 2015; Speijer 2015; Speijer et al. 2015) and at the origin of anisogamy (Allen 1996; de Paula et al. 2013). The syncytial state allows mitochondria and nuclei to mix and interact in myriad combinations, without requiring that viable offspring (packaged as single cells) be produced. The combination of mitochondria and nuclei into diaspores provides a very strong selective mechanism with which to select for compatible mitonuclear interactions, but from a multinucleated reserve that was genetically buffered against deleterious effects of single nuclei or mitochondria harboring incompatibilities. Syncytial buffering plays an important role at this major evolutionary transition.

Third, a syncytium provides time, nutrients, energy, and a spatially differentiated landscape (territories) for nuclei to undergo selection for refining the simple process of coordinated chromosome division. Better nuclear division means more nuclei, as progeny, within the syncytium, so a clear selective advantage for nuclei capable of increasingly refined nuclear division is apparent. From our inference, it is evident that repeated rounds of DNA replication prior to unrefined division seem more likely at the onset than a fully regimented replicate-fall-in-line-up-and-divide mode of chromosome segregation. With RecA-type homologous recombination going on between DNA strands, the overall habit of this kind of reductive chromosome division would have more in common with meiosis than mitosis (a segment of the cell cycle, which for lacking cytokinesis does not have much utility in a syncytium anyway). We do not endeavor here to offer an explanation for

why eukaryotic chromosomes came to undergo condensation, pairing, and alignment prior to nuclear division (closed state, without dissolution of the nuclear membrane). However, we do suggest that the process of ordered chromosome and nuclear division arose in a syncytial state, independent of cell division processes. This has the advantage of lowering the barriers of evolutionary invention that the first eukaryote had to surmount simultaneously. That is, a syncytial intermediate breaks down the almost intractably complex process of mitotic division into simpler yet still selectable component parts: Coordinated division of nucleus + chromosome being a simpler problem to solve than coordinated division of nucleus + chromosome + mitochondrion + cell. Our suggestion that coordinated nuclear and chromosome division evolved in a syncytial common ancestor would also directly account for the observation that all eukaryotes share the same basic conserved pattern of microtubule dependence in chromosome segregation (Koshland 1994) whereas their processes of cell division within and across supergroups are varied, for example, longitudinal fission in *Euglena*, phragmoblast formation in higher plants, gamete formation in *Acetabulariua*, or budding in yeast.

Fourth, the coenocyte offers possible transitional state solutions to the problem of cell division: Budding (spores). This is sketched in figure 3*B*. Archaea and bacteria produce membrane vesicles (Deatherage and Cookson 2012; Schwechheimer and Kuehn 2015; Gould et al. 2016). In archaea, the vesicles are pinched off from the cell surface with proteins of the Cdv (for cell division) family, which are archaeal precursors of the eukaryotic ESCRT III (for Vps2 and Vps4) proteins. These proteins are involved in making vesicles that protrude outwards from the cytosol, not inwards (as in endocytic processes). Reasonably assuming that our archaeal host had Cdv proteins, these could generate vesicles at the plasma membrane (Ellen et al. 2010). If no nuclei or mitochondria become contained in such a vesicle, fine, no problem, but no progeny. If nuclei with incomplete chromosome sets or lacking mitochondria become contained, also fine, also no progeny. But if vesicles come to contain both a mitochondrion and a nucleus (possibly more than one to start) with chromosome sets that are sufficiently complete to permit the formation of a new coenocyte, and sufficient proteins to initiate growth, the vesicle is a diaspore. This provides a very effective and powerful system of selection for combinations of nuclei and mitochondria that can found a new syncytium.

But, the diaspores are clonal. That brings us back to the dreaded dead end street of Muller's ratchet. Unless, that is, the diaspores can fuse, like archaeal cells can, so as to permit new combinations of chromosomes and genes. If that happens, then what started out as a hopeless symbiotic consortium has basically completed a meiotic cell cycle (fig. 1) with a syncytial "diploid" stage. The diaspores are homologous to meiospores, their fusion is homologous to syngamy, and karyogamy can take place either immediately or in the syncytial state.

Fifth, and finally in this section, a coenocytic eukaryote common ancestor would go a long way to explaining why all of the traits that are common to eukaryotes were assembled in the eukaryote common ancestor, without intermediate forms in the prokaryote to eukaryote transition: There were attempts at the spawning of intermediates through budding, but only 1) those diaspores that came to possess complete chromosome sets and 2) those that were able to fuse with other diaspores of different chromosome parentage were viable on the long term. Together, these considerations, though not tested by modeling but clearly modelable using the stochastic corrector framework (Grey et al. 1995), provide cause to pursue the idea of a coenocytic eukaryote common ancestor, or Coeca. Nonetheless, if we look across eukaryotic supergroups, there is little alternative to the view that the last eukaryote ancestor was unicellular and that it possessed sex and a cell cycle, which bring us to the next section.

Coupling Chromosome Division to Cell Division

It is notable that chromosome division and cell division are not tightly coupled in many modern eukaryotes (Parfrey et al. 2008). During early eukaryogenesis, before a bona fide cell cycle had evolved, chromosome and cell division might not have been coupled at all. The syncytial intermediate inferred so far could probably divide nuclei and chromosomes (karyokinesis) with some proficiency and produce spores corresponding to meiospores, some of which would be viable (fig. 3*B*). It must have possessed a basic machinery for the scission act of cell division, otherwise it would have been unable to cleave off spores. In the simplest scenario, spores would do what their parental coenocyte did: Grow into another coenocyte, getting better through selection at ordered chromosome and nuclear division (fig. 4). If however, heterochrony—phenotypic expression at the wrong stage of development, here a change in the time spent between sporulation and being a syncytium—sets in, such that the spore cleavage process took place very early in filament growth, our eukaryote might have had a chance to attain a primitive form of that which its prokaryotic ancestors took for granted: Binary cell division, but this time with nuclei and mitochondria. In this respect, the vesicle (spore) formation function of Cdvs in archaea and their ESCRT homologs in eukaryotes, which are involved in cell division (Ellen et al. 2010; Morita et al. 2010), remained conserved.

Bypassing the syncytial stage (spore formation from spores, possibly akin to budding is yeasts) would generate cells whose fitness would dramatically improve by mutations that led to a temporally coordinated regulation of chromosome and cell division. In principle, a fairly straightforward process of selection could have brought forth the basics of a cell cycle (next section) as the solution to that problem. The products of these

divisions would however, be clonal, returning our attention to Muller's ratchet and the need for recombination to avoid extinction. Fusion of cells would more or less correspond to gamete fusion (syngamy) in the sexual cycle of modern eukaryotes. Yet recombination does not take place until karyogamy has occurred, and the reader will note that in figure 4 we have not indicated nuclei. The reason is that there are many possibilities regarding the timing of karyogamy (immediate, dikaryon phase, multinucleated phase), recombination and reduction, in addition to the issue of whether the mitotic cells are haploid or diploid (compare figs. 1 and 4), we just leave it open. We note however that the requirement for recombination (sex) in our inference—recalling that the starting point for this essay was the origin of sex—clearly traces alternation of generations into the eukaryote common ancestor.

Also of note, the proteins that serve to condense chromosomes during the cell cycle and align homologous chromosomes during homologous recombination, members of the SMC family, for structural maintenance of chromatin (Jeppsson et al. 2014), are extremely important in the prokaryote to eukaryote transition. SMC homologs are present in prokaryotes (Soppa 2001; Soppa et al. 2002), so as a gene family they are not a eukaryote-specific invention, but their gene family diversification and their cell cycle-specific expression (Jeppsson et al. 2014) clearly are eukaryote-specific attributes. We suggest that, as in the case of tubulin and some other cytoskeletal proteins, energetically unpenalized overexpression of preexisting prokaryotic genes for structural proteins in the eukaryote ancestor led to very useful and highly conserved processes: Chromatin condensation during the cell cycle and homologous chromatid pairing during meiotic recombination. The advent of cohesins (members of the SMC family) and the origin of homologous pairing was clearly important in eukaryote evolution, even the key event in the origin of meiosis under the synapsis homolog model (Wilkins and Holliday 2008). However, like other models that start with a mitosing cell that lacks meiosis, the synapsis homolog model takes the origin of mitosis as a given, which in our inference is an explanandum, hence the two models address very different things. Our proposal lacks mitosing cells incapable of recombination.

The cells at the bottom of figure 4 could represent the last eukaryote common ancestor, as such they would have had all the many traits that the last eukaryote common ancestor had, including the energy metabolic repertoire of a facultative anaerobe (Müller et al. 2012), a nucleus plus complex endomembrane system (Gould et al. 2016), and nowhere mentioned so far nor drawn in any figure, a eukaryotic flagellum. Clearly, a flagellum would improve the fitness of all single-celled stages. We have nowhere referred to phagocytosis, because despite occasional staunch claims in the literature, it is by no means clear that the eukaryotic ancestor was phagocytotic (Gould et al. 2016; Sousa et al. 2016). Although the ancestral eukaryotic habit sketched in figure 4 looks very much like a chytridiomycete, the spores of which have a flagellum (James et al. 2006), and would have a similar physiology (facultative anaerobes), any resemblance is not by design. Some fungi are syncytial and fungi are not phagocytotic.

Steps En Route to a Cell Cycle

With a heritable means to segregate chromosomes, natural selection can improve the basic invention: Cells with better, more refined and more accurate chromosome segregation proliferate according to their fitness. Though the nucleus permits continuous gene expression, it decouples the process of metabolite accumulation from cell and chromosome division—which in prokaryotes are tightly linked. Without means to coordinate metabolism with division, no single-celled eukaryotes will arise. Cyclins, also an attribute of the eukaryote common ancestor, apparently solved this problem by establishing a hierarchy of cytosol-based sensing and decision making, so as to sense 1) when metabolites had been accumulated (G1 checkpoint), 2) when chromosomes had been replicated (S phase checkpoint), and 3) when cell division could be initiated (Norbury and Nurse 1992; Morgan 1997).

The cyclin/CDK system reflects the origin of nucleus–cytosol compartmentalization. In the prokaryotic cytosol, DNA-binding proteins like DnaA bind directly to the chromosome where they are sensed by the replication machinery (Mott and Berger 2007); when DNA is cytosolic, the concentration of DnaA is linked to metabolism. In the ancestral eukaryotic cell, however, this communication is disrupted by the presence of the nuclear membrane, precipitating the need for new sensing mechanisms and possibly marking the advent of nuclear DNA-binding proteins with cytosolic-binding partners. Cyclins are indeed homologous to archaeal TFIIB (Gibson et al. 1994) and CDKs are ser/thr kinases, which are common among prokaryotes (Pereira et al. 2011; Kennelly 2014). For the cyclin/CDK system, no fundamental inventions were required, but interactions across novel compartments.

A primitive cyclin/CDK system could have established communication between the chromosomes and the cell division machinery, which had become disrupted by the origin of the nuclear membrane. In yeast it is possible to drive both mitosis and meiotic cell cycles by a single-engineered cyclin–CDK complex (Gutiérrez-Escribano and Nurse 2015) suggesting that a simpler ancestral network consisting of a few essential components could, in principle, underlie the origin of a primitive eukaryotic cell cycle regulation network.

A consequence of the cell cycle is that eukaryotes condense chromosomes and shut down once per cell division. Bacteria can shut down gene expression globally by shutting down ribosomes, for example, in toxin-mediated plasmid responses (Van Melderen and Bast 2009; Bertram and Schuster 2014). Global gene expression shutdown in eukaryotes entails chromatin-modification (Ptashne 2005). Chromatin-based shut down of gene expression is a eukaryotic invention, its

evolutionary onset likely accompanied cell cycle origin (Maurer-Alcala and Katz 2015). Of course, a well-regulated cell cycle need not arise, but if it does not, no progeny will ensue. Genetic variation favoring the fixation of basic regulatory mechanisms governed by cyclins could generate the basic fabric of a cell cycle.

A basic cell cycle affords the first eukaryotes a plethora of fundamentally new possibilities. With mitochondrial power and ATP-dependent cytosolic structural proteins that can move and do things in the cytosol, they can undergo extended phases of gene expression without the burden of continuously dividing chromosomes (as in prokaryotes). This decoupling of chromosome replication from gene expression, and the regular shutdown of gene expression once per cell division is a hallmark of eukaryote biology. It enables long phases of gene expression from chromosomes that are not dividing, but are specifically dedicated to the gene expression process. The cell cycle shuts gene expression down, initiates chromosome and cell division, and then allows gene expression anew. Processes of cell development became possible that unfolded from the shut-down-and-reboot nature of chromatin condensation at every cell division. Mitochondrial power allowed eukaryotes to explore new protein function and protein overexpression in the cytosol, where eukaryote complexity takes root.

With the basic sequence of cell division, gene expression, cell fusion, and recombination in place, a meiotic cell cycle (recombination at every division) becomes dispensable. Occasional recombination suffices to escape Muller's ratchet (Muller 1964; Felsenstein 1974). From this more complex meiotic starting point, mitotic shortcuts in either the haploid or the diploid state are readily attained by shortening meiosis (fig. 1), as are variants extending the number of mitotic divisions between meiosis (fig. 4). Most cell divisions are once again clonal, as in prokaryotes, but with new chromosome segregation mechanisms. With the conserved core of meiosis and mitosis in place, ploidy phase variation was possible. Eukaryotes, especially protists, exhibit baroque diversity among ploidy cycles (Parfrey et al. 2008; Parfrey and Katz 2010), but they do not relinquish their genes for sex (Ramesh et al. 2005).

Some Consequences of a Primordial Coenocytic Model

Karyogamy and karyokinesis within the syncytial intermediate have a curious attribute: Together, they homogenize populations of individually unviable chromosome sets otherwise headed to extinction. This is of interest in several ways.

First, it fits very well with the checkpoints in the cell cycle that carefully monitor, hence insure, proper chromosome replication: Do not enter into chromosome and cell division until the chromosomes are fully replicated. This would have been an important milestone en route to achieving a regulated cell cycle of the type underpinning eukaryote cell division today. However, regulated chromosome division at the coenocytic state could have evolved without the need for simultaneously coupling regulated chromosome division to regulated cell division, because the content of spores resulting from scission (our suggested precursor to cell division) was initially random, viable contents being selected.

Second, at the syncytial stage, eukaryotes had solved their carbon and energy problems with the help of glycolysis in the cytosol and terminal oxidation plus ATP export in mitochondria. With their core metabolic problems solved in a virtually unimprovable manner (no known eukaryote has ever replaced or supercharged its mitochondria) members of the emergent eukaryotic lineage could no longer genuinely benefit from LGT with prokaryotes. They did not need new terminal oxidases in the inner mitochondrial membrane or NADH oxidizing enzymes in the cytosol. They needed maintenance and improvements in the regulation of their operational yet still clumsy cell and chromosome division. That is to say, what they needed for lineage survival they could not get from prokaryotes, only from other eukaryotic chromosomes, namely variants on the themes surrounding the formation and regulation of novel structures and processes in the cytosol that emerge from suddenly affordable ATP-costly protein overexpression in that compartment and ATP-dependent protein aggregation states and protein interaction states therein. The genes and proteins underpinning eukaryotic specific traits (nucleus, endomembrane system, and the like) arose in the eukaryotic ancestor, eukaryotic cells were thus the only existing source of newly emergent genes that characterized the eukaryotic lineage. The transition to sex marked the departure from the LGT mechanisms that impart recombination among prokaryotes, and simultaneously marked the advent of reciprocal recombination among kin as the mechanism to escape Müller's ratchet in the eukaryotic lineage. The origin of eukaryotes was the origin of vertical lineage inheritance (Ku et al. 2015), and sex was required to keep vertically evolving lineages viable.

Third, recurrent genome fusions generate multiple gene copies and large gene families, much in the same way that whole-genome duplications do among eukaryotes today. On the one hand, this creates massive paralogy among ancestral eukaryotic genes. In the presence of mitochondria, which permit the new gene copies to be expressed as protein, it also creates large gene families for the genes specific to the eukaryotic lineage, thereby allowing experimentation, functional specialization and fixation of members within eukaryote-specific gene families involved in membrane traffic, cell structure, and signaling. Clearly, our proposal predicts the existence of massive paralogy in the eukaryote common ancestor.

Fourth and finally, the syncytial and meiospore fusion aspects help to explain why eukaryote origins left no intermediate forms. Only when the whole suite of genes required for survival as a nucleus-bearing cell was assembled, were the

products of cell division (meiospore) genuinely self-sufficient to an extent that would allow diversification into descendant lineages.

Conclusion

Our proposal has it that eukaryotic sex arose as a necessary replacement for prokaryotic gene transfer mechanisms in a cell that had evolved a nucleus (as a consequence of mitochondrial group II introns) and that was consequently able to undergo the transition to MDCS as a fortuitous byproduct of mitochondrial energetics. The mitochondrion transformed its host, generating a cytosol 1) where protein could be synthesized in amounts no prokaryote had ever experienced, 2) where proteins could interact without interfering either with chromatin or with cell division, and 3) where spontaneous aggregation of structural proteins to higher order structures and ATP-dependent disaggregation became possible. Selection was acting upon the outcome of spontaneous protein interactions—spontaneous chemical processes were thus decisive at this major evolutionary transition.

Once it exists, the benefits of sex are manifold. The how and why of how sex came into existence are more challenging. The DNA damage hypothesis (Hörandl and Hadacek 2013) posits that meiosis evolved as a repair response to oxidative damage. Our model would not exclude a role for recombination in repair, since the enzymes involved in meiotic recombination are homologous to prokaryotic repair machineries (Camerini-Otero and Hsieh 1995; Ramesh et al. 2005), and because mitochondria (the source of reactive oxygen species in eukaryotes) are present in the cell that evolved sex. Homolog synapsis involving cohesins, members of the SMC protein family (Jeppsson et al. 2014), has been suggested as the key innovation that allowed for the evolution of meiosis from mitosis (Wilkins and Holliday 2008). In our proposal, the chromosomes are initially heterogeneous. Hence until a homogeneous, ancestral eukaryote genome with defined chromosome numbers having haploid and diploid states during an evolved cell cycle and life cycle had emerged, homologous pairing would not have been obviously beneficial. Congruent with that view, many prokaryotes possess clear homologs of the SMC proteins that promote eukaryote chromosome condensation and that lead to pairing of sister chromatids (Soppa 2001; Soppa et al. 2002), yet prokaryotes neither condense their chromatin at cell division nor do they possess sister chromatids.

A number of major evolutionary transitions in eukaryote evolution involve endosymbiosis: The origin of mitochondria, the origin of plastids, and the origin of major algal groups through secondary endosymbiosis. Each of those endosymbiotic transitions also left a major impact on the genome in the form of gene transfers from organelles to the nucleus (Martin and Müller 1998; Martin et al. 2002; Ku et al. 2015). But endosymbiosis is not readily accommodated either by mathematics or by the gradualist paradigm of population genetics, which is why mitochondria play no role whatsoever in population genetic approaches to understanding the prokaryote–eukaryote transition (Lynch and Conery 2003; Lynch 2006; Lynch and Marinov 2015). Indeed, population genetic investigations tend to recognize no difference at all between prokaryotes and eukaryotes, both of which appear to map out along an uninterrupted continuum by the measure of population genetic parameters (Lynch and Conery 2003). That is not a criticism of population genetics, it is a statement about the evolutionary divide separating eukaryotes from prokaryotes. Viewed solely through the looking glass of population genetics and allele frequencies, one would not even be able to tell the difference between a lion and a palm tree, because on the long term, both have sex and diploid genetics. Clearly, population genetics does not tell us everything that we need to know about eukaryote evolution, if we are interested in the physiological and cell biological differences that distinguish eukaryotes from prokaryotes.

At the prokaryote to eukaryote transition, many major changes took place: Mitochondria, endomembrane system, nucleus, meiosis, mitosis, and the origin of sex. We have endeavored here to account for those differences in one essay. In doing so we necessarily crossed a border from modern endosymbiotic theory, starting from an archaeaon as the host, to population genetics (sex). We were able to sketch, albeit in broad strokes, a general outline that bridged the evolutionary gap between prokaryotes and eukaryotes and that resulted in a mitochondriate, nucleated, sexually recombining and mitosing cell. Whether population genetic approaches would be able to predict the single origin of plastids or mitochondria during evolution remains an open question. Whether the eukaryotic cell cycle or endomembrane system could be modeled as a population genetic parameter is also an open question. When it comes to modeling the processes underlying the cell biological differences that distinguish eukaryotes from prokaryotes and taking them apart into simpler, conceptually tractable components, endosymbiotic theory works fairly well. Yet endosymbiotic theory is founded in comparative biology and comparative physiology, not in mathematics and statistics, hence as a field it does not interface well with population genetics. Maybe future progress will improve that circumstance.

Sex, the cell cycle, chromosome division, and alternation of generations are processes. Inferring the evolutionary origin of processes is arguably more difficult than inferring the origin of structures. Understanding the origin of organelles and structures that distinguish eukaryotes from prokaryotes, for example, the endomembrane system, is important, and a strong case can be made that the origin of mitochondria had the decisive role in endomembrane origin (Gould et al. 2016). A better understanding of eukaryote origin requires understanding the evolution of eukaryotic processes. The syncytial stage has many virtues as an evolutionary intermediate. In terms of

structuring the problem of eukaryote process origins, it also helps to break down the complex eukaryotic cell cycle and life cycle (fig. 1) into simpler component processes (fig. 4), and in doing so draws attention to the largely (but not completely: Cavalier-Smith 2010) neglected issue of the origin of alternation of generation in eukaryotes. Students of biology have had to learn alternation of generations for over a century, yet without an evolutionary context that could account for the origin of the sexual processes (karyogamy and reduction division) that connect the alternating generations. We have made an effort here to put sex and the alternation of generations into the evolutionary context of modern endosymbiotic theory. Our evolutionary intermediates obtain new gene variants through karyogamy in a syncytium and through archaeal-type spore fusion. The end result of the inference is a population of free-living, unicellular, sexually recombining and mitosing cells that have a cell cycle, archaeal ribosomes in the cytosol, and mitochondria.

In recent years, views on the origin of eukaryotes have changed in that 1) the mitochondrion is now recognized to be ancestrally present in the eukaryote common ancestor and in that 2) the host is now considered to have been an archaeon (Martin and Müller 1998; Martin and Koonin 2006; Cox et al. 2008; Lane and Martin 2010; Williams et al. 2013; McInerney et al. 2014; Raymann et al. 2015; Spang et al. 2015; Sousa et al. 2016; Gould et al. 2016). Views on the origin of sex have not changed at the same pace, though there is increasing interest in the possible role(s) of mitochondria in promoting the establishment of meiotic recombination (Lane 2009,2015; Speijer et al. 2015; Havird et al. 2015; Radzvilavicius and Blackstone 2015). Here, we have considered the origin of sex on the basis of those newer premises.

In contrast to earlier theories, we do not assume that the cell in which sex arose had already evolved a mitotic cell cycle. Rather, we start from a symbiotic association of two prokaryotes, which led to the origin of mitochondria, and consider the evolutionary sequence of events. As the most notable departures from previous theories on the origin of eukaryotes, mitosis or sex, 1) we posit that mitotic division is evolutionarily derived from meiotic division; 2) we place both processes in their natural context of eukaryotic cell cycle origin; 3) we propose a coenocytic eukaryote common ancestor, Coeca, which allowed nuclei harboring defective in chromosome sets to complement each other through mRNA in the cytosol; 4) we suggest that the first form of eukaryotic cell division was budding of meiospores from the coenocyte, a process that selected viable combinations of chromosomes and mitochondria; and 5) we suggest that the ability of meiospores to fuse, a property of archaeal cells, allowed them to undergo homologous meiotic recombination, or sex. It is an observation that sex has been retained in all eukaryotic lineages for over 1.7 Gyr, we suggest that the reason for its retention is the same as the reason for its fixation at eukaryote origins: It saves eukaryotes from extinction through Muller's ratchet. Our inference orders the origin of major evolutionary innovations in the eukaryotic lineage as follows: Mitochondria, followed by the nucleus and endomembrane system, MDCS, reduction division in a syncytial eukaryote common ancestor, meiospore production, division and fusion leading to a meiotic life cycle and cell cycle (sex), and finally mitosis through bypassing of recombination and reduction division during the life cycle. The syncytial nature of the eukaryote common ancestor is a particularly interesting, and possibly useful, element of our proposal. It is interesting because it decouples the processes of chromosome segregation, ploidy cycling, karyogamy, progeny generation, cell division, and the cell cycle from each other. It is useful because it allows one to consider the evolution of each of those actions as a tendentially independent biological process, which fits well with the diversity, or lack thereof, observed for each of these processes among contemporary eukaryotic groups.

Acknowledgments

The authors gratefully acknowledge funding from the European Research Council (ERC 666053 to W.F.M.). They thank James O. McInerney, Laura A. Katz, Nick Lane, Eors Szathmary, Sven B. Gould, Neil Blackstone, John F. Allen, John Logsdon, and especially Klaus V. Kowallik for constructive discussions. The authors gratefully acknowledge funding from the European Research Council (ERC 666053 to W.F.M.).

Literature Cited

Allen JF. 1996. Separate sexes and the mitochondrial theory of ageing. J Theor Biol. 180:135–140.
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.
Audibert A, Weil D, Dautry F. 2002. In vivo kinetics of mRNA splicing and transport in mammalian cells. Mol Cell Biol. 22:6706–6718.
Bakker EP, Van Dam K. 1974. The movement of monocarboxylic acids across phospholipid membranes: evidence for an exchange diffusion between pyruvate and other monocarboxylate ions. Biochim Biophys Acta–Biomembranes. 339:285–289.
Beech PL., et al. 2000. Mitochondrial FtsZ in a chromophyte alga. Science 287:1276–1279.
Bell G. 1982. The masterpiece of nature: the evolution and genetics of sexuality. Berkeley: Univ. of California Press.
Bentley SD., et al. 2002. Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). Nature 417:141–147.
Bernstein H, Byerly HC, Hopf FA, Michod RE. 1984. Origin of sex. J Theor Biol. 110:323–351.
Bertram R, Schuster CF. 2014. Post-transcriptional regulation of gene expression in bacterial pathogens by toxin-antitoxin systems. Front Cell Infect Microbiol. 4:6.
Blackstone NW. 2013. Why did eukaryotes evolve only once? Genetic and energetic aspects of conflict and conflict mediation. Phil Trans R Soc B. 368:20120266.
Blagosklonny MV. 2011. Cell cycle arrest is not senescence. Aging 4:94–101.
Bloomfield G. 2016. Atypical ploidy cycles, Spo11, and the evolution of meiosis. Sem Cell Dev Biol. 54:158–164.

Booth A, Doolittle WF. 2015. Eukaryogenesis, how special really? Proc Natl Acad Sci U S A. 112:10278–10285.
Bricheux G, Coffe G, Brugerolle G. 2007. Identification of a new protein in the centrosome-like "atractophore" of *Trichomonas vaginalis*. J Microbiol Immunol Infect. 153:133–140.
Esser C., et al. 2004. A genome phylogeny for mitochondria among alpha-proteobacteria and a predominantly eubacterial ancestry of yeast nuclear genes. Mol Biol Evol. 21:1643–1660.
Camerini-Otero RD, Hsieh P. 1995. Homologous recombination proteins in prokaryotes and eukaryotes. Annu Rev Genet. 29:509–552.
Cande WZ, McDonald KL. 1985. *In vitro* reactivation of anaphase spindle elongation using isolated diatom spindles. Nature 316:168–170.
Castagnetti S, Božič B, Svetina S. 2015. Mechanical and molecular basis for the symmetrical division of the fission yeast nuclear envelope. Phys Chem Chem Phys. 17:15629–15636.
Cavalier-smith T. 1975. The origin of nuclei and of eukaryotic cells. Nature 256:463–468.
Cavalier-Smith T. 2002. Origins of the machinery of recombination and sex. Heredity 88:125–141.
Cavalier-Smith T. 2010. Origin of the cell nucleus, mitosis and sex: roles of intracellular coevolution. Biol Direct. 5:7.
Cavalier-smith T. 2014. The neomuran revolution and phagotrophic origin of eukaryotes and cilia in the light of intracellular coevolution and a revised tree of life. Cold Spring Harb Perspect Biol. 6:a016006.
Cleveland LR. 1947. The origin and evolution of meiosis. Science 105:287–289.
Collins L, Penny D. 2005. Complex spliceosomal organization ancestral to extant eukaryotes. Mol Biol Evol. 22:1053–1066.
Cox CJ, Foster PG, Hirt RP, Harris SR, Embley TM. 2008. The archaebacterial origin of eukaryotes. Proc Natl Acad Sci U S A. 105:20356–20361.
Cross FR, Buchler NE, Skotheim JM. 2011. Evolution of networks and sequences in eukaryotic cell cycle control. Phil Trans R Soc B. 366:3532–3544.
Crow JF. 2005. Timeline–Hermann Joseph Muller, evolutionist. Nat Rev Genet. 6:941–945.
de Duve C. 2007. The origin of eukaryotes: a reappraisal. Nat Rev Genet. 8:395–403.
de Paula WBM, Lucas CH, Agip ANA, Vizcay-Barrena G, Allen JF. 2013. Energy, ageing, fidelity and sex: oocyte mitochondrial DNA as a protected genetic template. Phil Trans R Soc B. 368:20120263.
Deatherage BL, Cookson BT. 2012. Membrane vesicle release in bacteria, eukaryotes, and archaea: a conserved yet underappreciated aspect of microbial life. Infect Immun. 80:1948–1957.
Degli Esposti M. 2014. Bioenergetic evolution in proteobacteria and mitochondria. Genome Biol Evol. 6:3238–3251.
Doležal P, Likic V, Tachezy J, Lithgow T. 2006. Evolution of the molecular machines for protein import into mitochondria. Science 313:314–318.
Doolittle WF. 1999. Phylogenetic classification and the universal tree. Science 284:2124–2128.
Ellen AF, Zolghadr B, Driessen AMJ, Albers S-V. 2010. Shaping the archaeal cell envelope. Archaea 2010:1–13.
Embley TM, Martin W. 2006. Eukaryotic evolution, changes and challenges. Nature 440:623–630.
Erickson HP. 1995. FtsZ, a prokaryotic homolog of tubulin? Cell 80:367–370.
Erickson HP. 2007. Evolution of the cytoskeleton. BioEssays 29:668–677.
Felsenstein J. 1974. The evolutionary advantage of recombination. Genetics 78:737–756.
Fleury-Aubusson A. 2003. "Novel cytoskeletal proteins in protists": introductory remarks. J Eukaryot Microbiol. 50:3–8.
Garg S., et al. 2015. Conservation of transit peptide-independent protein import into the mitochondrial and hydrogenosomal matrix. Genome Biol Evol. 7:2716–2726.
Gavrilova OV, Rundanova EE. 1999. Cell architecture in the morphogenesis of coenocytic alga *Vaucheria sessilis*. I. The morphology of germination and the behaviour of nuclei. Protistology 1:5–9.
Gibson TJ, Thompson JD, Blocker A, Kouzarides T. 1994. Evidence for a protein domain superfamily shared by the cyclins, TFIIB and RB/p107. Nucleic Acids Res. 22:946–952.
Goodenough U, Heitman J. 2014. Origins of eukaryotic sexual reproduction. Cold Spring Harb Perspect Biol. 6:a016154.
Gould SB, Garg SG, Martin WF. 2016. Bacterial vesicle secretion and the evolutionary origin of the eukaryotic endomembrane system. Trends Microbiol. 24:525–534.
Gould SB, et al. 2011. Ciliate pellicular proteome identifies novel protein families with characteristic repeat motifs that are common to Alveolates. Mol Biol Evol. 28:1319–1331.
Grey D, Hutson V, Szathmáry E. 1995. A re-examination of the stochastic corrector model. Proc R Soc B. 262:29–35.
Gutiérrez-Escribano P, Nurse P. 2015. A single cyclin-CDK complex is sufficient for both mitotic and meiotic progression in fission yeast. Nat Commun. 6:6871.
Hadjivasiliou Z, Lane N, Seymour RM, Pomiankowski A. 2013. Dynamics of mitochondrial inheritance in the evolution of binary mating types and two sexes. Proc R Soc B. 280:20131920.
Halary S., et al. 2011. Conserved meiotic machinery in *Glomus spp.*, a putatively ancient asexual fungal lineage. Genome Biol Evol. 3:950–958.
Hand ML, Koltunow AMG. 2014. The genetic control of apomixis: asexual seed formation. Genetics 197:441–450.
Harold FM. 1986. The vital force. New York: WH Freeman.
Hartwell LH, Weinert TA. 1989. Checkpoints: controls that ensure the order of cell cycle events. Science 246:629–634.
Hasitschka-Jenschke G. 1960. Beitrag zur Karyologie von Characeen. Österreichische Botanische Zeitschrift. 107:228–240.
Havird JC, Hall MD, Dowling DK. 2015. The evolution of sex: a new hypothesis based on mitochondrial mutational erosion. Bioessays 37:951–958.
Hazkani-Covo E, Covo S. 2008. Numt-mediated double-strand break repair mitigates deletions during primate genome evolution. PLoS Genet. 4:e1000237.
Hazkani-Covo E, Zeller RM, Martin W. 2010. Molecular poltergeists: mitochondrial DNA copies (numts) in sequenced nuclear genomes. PLoS Genet. 6:e1000834.
Hörandl E, Hadacek F. 2013. The oxidative damage initiation hypothesis for meiosis. Plant Reprod. 26:351–367.
Hourcade D, Dressler D. 1973. The amplification of ribosomal RNA genes involves a rolling circle intermediate. Proc Natl Acad Sci U S A. 70:2926–2930.
Huang CY, Ayliffe MA, Timmis JN. 2004. Simple and complex nuclear loci created by newly transferred chloroplast DNA in tobacco. Proc Natl Acad Sci U S A. 101:9710–9715.
Huang CY, Grünheit N, Ahmadinejad N, Timmis JN, Martin W. 2005. Mutational decay and age of chloroplast and mitochondrial genomes transferred recently to angiosperm nuclear chromosomes. Plant Physiol. 138:1723–1733.
Hurst LD, Nurse P. 1991. A note on the evolution of meiosis. J Theor Biol. 150:561–563.
Husnik F., et al. 2013. Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. Cell 153:1567–1578.
James TY., et al. 2006. A molecular phylogeny of the flagellated fungi (Chytridiomycota) and description of a new phylum (Blastocladiomycota). Mycologia 98:860–871.
Jékely G. 2014. Origin and evolution of the self-organizing cytoskeleton in the network of eukaryotic organelles. Cold Spring Harb Perspect Biol. 6:a016030–a016030.

Jeppsson K, Kanno T, Shirahige K, Sjögren C 2014. The maintenance of chromosome structure: positioning and functioning of SMC complexes. Nat Rev Mol Cell Biol. 15:601–614.
Jones D, Sneath PH. 1970. Genetic transfer and bacterial taxonomy. Bacteriol Rev. 34:40–81.
Katz LA. 2012. Origin and diversification of eukaryotes. Annu Rev Microbiol. 66:411–427.
Kennelly PJ. 2014. Protein Ser/Thr/Tyr phosphorylation in the Archaea. J Biol Chem. 289:9480–9487.
Kobayashi T. 2011. Regulation of ribosomal RNA gene copy number and its role in modulating genome integrity and evolutionary adaptability in yeast. Cell Mol Life Sci. 68:1395–1403.
Kobialka M, Michalik A, Walczak M, Junkiert L, Szklarzewicz T. 2016. Sulcia symbiont of the leafhopper *Macrosteles laevis* (Ribaut, 1927) (Insecta, Hemiptera, Cicadellidae: Deltocephalinae) harbors *Arsenophonus* bacteria. Protoplasma 253:903–912.
Kondrashov AS. 1994. The asexual ploidy cycle and the origin of sex. Nature 370:213–216.
Koonin EV. 2015. Origin of eukaryotes from within archaea, archaeal eukaryome and bursts of gene gain: eukaryogenesis just made easier? Phil Trans R Soc B. 370:20140333.
Koshland D. 1994. Mitosis: back to the basics. Cell 77:951–954.
Krylov DM, Nasmyth K, Koonin EV. 2003. Evolution of eukaryotic cell cycle regulation: stepwise addition of regulatory kinases and late advent of the CDKs. Curr Biol. 13:173–177.
Ku C., et al. 2015. Endosymbiotic origin and differential loss of eukaryotic genes. Nature 524:427–432.
Kuwabara T., et al. 2005. *Thermococcus coalescens* sp. nov., a cell-fusing hyperthermophilic archaeon from Suiyo Seamount. Int J Syst Evol Microbiol. 55:2507–2514.
Lambowitz AM, Zirnmerly S. 2004. Mobile group II introns. Annu Rev Genet. 38:1–35.
Lambowitz AM, Zimmerly S. 2011. Group II introns: mobile ribozymes that invade DNA. Cold Spring Harb Perspect Biol. 3:a003616.
Lane N. 2005. Power, sex, suicide: mitochondria and the meaning of life. Oxford: Oxford University Press.
Lane N. 2009. Life ascending. London: Profile Books.
Lane N. 2014. Bioenergetic constraints on the evolution of complex life. Cold Spring Harb Perspect Biol. 6:a015982.
Lane N. 2015 The vital question. Why is life the way it is? London: Profile Books.
Lane N, Martin W. 2010. The energetics of genome complexity. Nature 467:929–934.
Lane N, Martin WF. 2015. Eukaryotes really are special, and mitochondria are why. Proc Natl Acad Sci U S A. 112:E4823.
Lane N, Martin WF. 2016. Mitochondria, complexity, and evolutionary deficit spending. Proc Natl Acad Sci U S A. 113:E666.
Lang AS, Zhaxybayeva O, Beatty JT. 2012. Gene transfer agents: phage-like elements of genetic exchange. Nat Rev Microbiol. 10:472–482.
Lindås AC, Bernander R. 2013. The cell cycle of archaea. Nat Rev Microbiol. 11:627–638.
Lukjancenko O, Wassenaar TM, Ussery DW. 2010. Comparison of 61 sequenced *Escherichia coli* genomes. Microb Ecol. 60:708–720.
Lynch M. 2006. The origins of eukaryotic gene structure. Mol Biol Evol. 23:450–468.
Lynch M, Conery JS. 2003. The origins of genome complexity. Science 302:1401–1404.
Lynch M, Marinov GK. 2015. The bioenergetic costs of a gene. Proc Natl Acad Sci U S A. 112:15690–15695.
Lynch M, Richardson AO. 2002. The evolution of spliceosomal introns. Curr Opin Genet Dev. 12:701–710.
Martin W. 1999. Mosaic bacterial chromosomes: a challenge en route to a tree of genomes. BioEssays 21:99–104.
Martin W, Brinkmann H, Savonna C, Cerff R. 1993. Evidence for a chimeric nature of nuclear genomes: eubacterial origin of eukaryotic glyceraldehyde-3-phosphate dehydrogenase genes. Proc Natl Acad Sci U S A. 90:8692–8696.
Martin W, Koonin EV. 2006. Introns and the origin of nucleus–cytosol compartmentalization. Nature 440:41–45.
Martin W, Müller M. 1998. The hydrogen hypothesis for the first eukaryote. Nature 392:37–41.
Martin W., et al. 2002. Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc Natl Acad Sci U S A. 99:12246–12251.
Matsuura M, et al. 1997. A bacterial group II intron encoding reverse transcriptase, maturase, DNA endonuclease activities: biochemical demonstration of maturase activity and insertion of new genetic information within the intron. Genes Dev. 11:2910–2924.
Maurer-Alcala XX, Katz LA. 2015. An epigenetic toolkit allows for diverse genome architectures in eukaryotes. Curr Opin Genet Dev. 35:93–99.
Maynard Smith J. 1978. The evolution of sex. Cambridge: Cambridge Univ. Press.
Maynard Smith J. 1986. Evolution: contemplating life without sex. Nature 324:300–301.
Maynard Smith J., Szathmary E. 1997. The major transitions in evolution. Oxford: Oxford Univ. Press.
McCutcheon JP, McDonald BR, Moran NA. 2009. Convergent evolution of metabolic roles in bacterial co-symbionts of insects. Proc Natl Acad Sci U S A. 106:15394–15399.
McGrath CL, Katz LA. 2004. Genome diversity in microbial eukaryotes. Trends Ecol Evol. 19:32–38.
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. 12:449–455.
McNaughton EE, Goff LJ. 1990. The role of microtubules in establishing nuclear spatial patterns in multinucleate green algae. Protoplasma 157:19–37.
Mereschkowsky C. 1905. Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. Biol Centralbl. 25:593–604. [English translation in Martin W, Kowallik KV (1999) Eur. J. Phycol. 34:287-295].
Mierzejewska J, Jagura-Burdzy G. 2012. Prokaryotic ParA-ParB-parS system links bacterial chromosome segregation with the cell cycle. Plasmid 67:1–14.
Mitchison JM. 1971. The biology of the cell cycle. London: Cambridge Univ. Press.
Moran NA. 1996. Accelerated evolution and Muller's ratchet in endosymbiotic bacteria. Proc Natl Acad Sci U S A. 93:2873–2878.
Morgan DO. 1997. Cyclin-dependent kinases: engines, clocks, and microprocessors. Annu Rev Cell Dev Biol. 13:261–291.
Morita E., et al. 2010. Human ESCRT-III and VPS4 proteins are required for centrosome and spindle maintenance. Proc Natl Acad Sci U S A. 107:12889–12894.
Mott ML, Berger JM. 2007. DNA replication initiation: mechanisms and regulation in bacteria. Nat Rev Microbiol. 5:343–354.
Muller HJ. 1964. The relation of recombination to mutational advance. Mutat Res. 106:2–9.
Müller M., et al. 2012. Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. Microbiol Mol Biol Rev. 76:444–495.
Naor A, Gophna U. 2014. Cell fusion and hybrids in Archaea. Bioengineered 4:126–129.
Nasmyth K. 1995. Evolution of the cell-cycle. Phil Trans R Soc B. 349: 271–281.
Nei M. 1987. Molecular evolutionary genetics. New York: Columbia University Press
Neidhardt FC, Ingraham JL, Schaechter M. 1990. Physiology of the bacterial cell. Sunderland (MA): Sinauer Associates Incorporated

Nelson-Sathi S., et al. 2015. Origins of major archaeal clades correspond to gene acquisitions from bacteria. Nature 517:77–80.
Nogales E. 2000. Structural insights into microtubule function. Annu Rev Biochem. 69:277–302.
Norbury C, Nurse P. 1992. Animal cell cycles and their control. Annu Rev Biochem. 61:441–468.
Novak B, Csikasz-Nagy A, Gyorffy B, Nasmyth K, Tyson JJ. 1998. Model scenarios for evolution of the eukaryotic cell cycle. Phil Trans R Soc B. 353:2063–2076.
Oakley BR. 2000. An abundance of tubulins. Trends Cell Biol. 10:537–542.
Ochman H, Lawrence JG, Groisman EA. 2000. Lateral gene transfer and the nature of bacterial innovation. Nature 405:299–304.
Otto SP, Goldstein DB. 1992. Deleterious mutations and the origin of the meiotic ploidy cycle. Genetics 131:745–751.
Papke RT, Koenig JE, Rodriguez-Valera F, Doolittle WF. 2004. Frequent recombination in a saltern population of *Halorubrum*. Science 306:1928–1929.
Parfrey LW, Katz LA. 2010. Dynamic genomes of eukaryotes and the maintenance of genomic integrity. Microbe 5:156–163.
Parfrey LW, Lahr DJG, Katz LA. 2008. The dynamic nature of eukaryotic genomes. Mol Biol Evol. 25:787–794.
Parfrey LW, Lahr DJG, Knoll AH, Katz LA. 2011. Estimating the timing of early eukaryotic diversification with multigene molecular clocks. Proc Natl Acad Sci U S A. 108:13624–13629.
Pereira SFF, Goss L, Dworkin J. 2011. Eukaryote-like serine/threonine kinases and phosphatases in bacteria. Microbiol Mol Biol Rev. 75:192–212.
Pickett-Heaps JD, Tippit DH, Porter KR. 1982. Rethinking mitosis. Cell 29:729–744.
Ptashne M. 2005. Regulation of transcription: from lambda to eukaryotes. Trends Biochem Sci. 30:275–279.
Purkanti R, Thattai M. 2015. Ancient dynamin segments capture early stages of host–mitochondrial integration. Proc Natl Acad Sci U S A. 112:2800–2805.
Radzvilavicius AL, Blackstone NW. 2015. Conflict and cooperation in eukaryogenesis: implications for the timing of endosymbiosis and the evolution of sex. J R Soc Interface 12:20150584.
Raikov IB. 1994. The diversity of forms of mitosis in protozoa–a comparative review. Eur J Protistol. 30:253–269.
Ramesh MA, Malik S-B, Logsdon JM Jr. 2005. A phylogenomic inventory of meiotic genes. Curr Biol. 15:185–191.
Rasko DA., et al. 2008. The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. J Bacteriol. 190:6881–6893.
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. 112:6670–6675.
Rich PR, Maréchal A. 2010. The mitochondrial respiratory chain. Essays Biochem. 47:1–23.
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. Curr Biol. 13:1512–1517.
Roper M, Ellison C, Taylor JW, Glass NL. 2011. Nuclear and genome dynamics in multinucleate ascomycete fungi. Curr Biol. 21: R786–R793.
Rougier N, Werb Z. 2001. Minireview: parthenogenesis in mammals. Mol Reprod Dev. 59:468–474.
Schleper C, Holz I, Janekovic D, Murphy J, Zillig W. 1995. A multicopy plasmid of the extremely thermophilic archaeon Sulfolobus effects its transfer to recipients by mating. J Bacteriol. 177:4417–4426.
Schumacher MA., et al. 2015. Structures of archaeal DNA segregation machinery reveal bacterial and eukaryotic linkages. Science 349:1120–1124.
Schwechheimer C, Kuehn MJ. 2015. Outer-membrane vesicles from gram-negative bacteria: biogenesis and functions. Nat Rev Microbiol. 13:605–619.
Searcy DG. 2003. Metabolic integration during the evolutionary origin of mitochondria. Cell Res. 13:229–238.
Sharp PA. 1985. On the origin of RNA splicing and introns. Cell 42: 397–400.
Solari AJA. 2002. Primitive forms of meiosis: the possible evolution of meiosis. Biocell 26:1–13.
Soppa J. 2001. Prokaryotic structural maintenance of chromosomes (SMC) proteins: distribution, phylogeny, and comparison with MukBs and additional prokaryotic and eukaryotic coiled-coil proteins. Gene 278:253–264.
Soppa J, et al. 2002. Discovery of two novel families of proteins that are proposed to interact with prokaryotic SMC proteins, and characterization of the Bacillus subtilis family members ScpA and ScpB. Mol Microbiol. 45:59–71.
Sørensen MA, Kurland CG, Pedersen S. 1989. Codon usage determines translation rate in *Escherichia coli*. J Mol Biol. 207:365–377.
Sousa FL, Neukirchen S, Allen JF, Lane N, Martin WF. 2016. Lokiarchaeon is hydrogen dependent. Nature Microbiol. 1:16034.
Spang A., et al. 2015. Complex archaea that bridge the gap between prokaryotes and eukaryotes. Nature 521:173–179.
Speijer D. 2015. Birth of the eukaryotes by a set of reactive innovations: new insights force us to relinquish gradual models. Bioessays 37:1268–1276.
Speijer D, Lukes J, Elias M. 2015. Sex is a ubiquitous, ancient, and inherent attribute of eukaryotic life. Proc Natl Acad Sci U S A. 112:8827–8834.
Stouthamer AH. 1978. Energy-yielding pathways. In: Gunsalus IC, editor. The bacteria vol VI: bacterial diversity. New York: Academic Press. p. 389–462.
Takahashi F, Yamaguchi K, Hishinuma T, Kataoka H. 2003. Mitosis and mitotic wave propagation in the coenocytic alga, *Vaucheria terrestris* sensu Goetz. J Plant Res. 116:381–387.
Tekle YI, Anderson OR, Leckya AF. 2014. Evidence of parasexual activity in "asexual amoebae" *Cochliopodium* spp. (Amoebozoa): extensive cellular and nuclear fusion. Protist 165:676–687.
Thiergart T, Landan G, Schenk M, Dagan T, Martin WF. 2012. An evolutionary network of genes present in the eukaryote common ancestor polls genomes on eukaryotic and mitochondrial origin. Genome Biol Evol. 4:466–485.
Timmis JN, Ayliffe MA, Huang CY, Martin W. 2004. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Rev Genet. 5:123–135.
Toro E, Shapiro L. 2010. Bacterial chromosome organization and segregation. Cold Spring Harb Perspect Biol. 2:a000349.
Uyenoyama MK, Bengtsson BO. 1989. On the origin of meiotic reproduction: a genetic modifier model. Genetics 123:873–885.
Van Melderen L., De Bast MS 2009. Bacterial toxin–antitoxin systems: more than selfish entities? PLoS Genet. 5:e1000437.
Verbruggen H., et al. 2009. A multi-locus time-calibrated phylogeny of the siphonous green algae. Mol Phylogenet Evol. 50:642–653.
von Dohlen CD, Kohler S, Alsop ST, McManus WR. 2001. Mealybug beta-proteobacterial endosymbionts contain gamma-proteobacterial symbionts. Nature 412:433–436.
Welch DBM, Meselson M. 2000. Evidence for the evolution of bdelloid rotifers without sexual reproduction or genetic exchange. Science 288:1211–1215.
Whatley JM, John P, Whatley FR. 1979. From extracellular to intracellular: the establishment of mitochondria and chloroplasts. Proc R Soc B. 204:165–187.
Wilkins AS, Holliday R. 2008. The evolution of meiosis from mitosis. Genetics 181:3–12.
Williams GC. 1975. Sex and evolution. Princeton, NJ: Princeton Univ. Press.

Williams TA, Foster PG, Cox CJ, Embley TM. 2013. An archaeal origin of eukaryotes supports only two primary domains of life. Nature 504:231–236.
Wujek DE. 1979. Intracellular bacteria in the blue-green alga *Pleurocapsa minor*. Trans Am Microsc Soc. 98:143–145.
Zuccolo M., et al. 2007. The human Nup107–160 nuclear pore subcomplex contributes to proper kinetochore functions. EMBO J. 26:1853–1864.

Associate editor: Laura Katz

Summary of results

The origin of eukaryotes from prokaryotes spans the greatest evolutionary divide in the living world. It has been long recognized that the process of eukaryogenesis did not leave any intermediates for biologists to examine. Examination of extant eukaryotes and eukaryotic gene families paint a picture of a complex Last Eukaryotic Common Ancestor (LECA) (Koumandou *et al.*, 2013). Despite several efforts a unifying theory for eukaryogenesis is still missing and attempts to synthesize one are rare. Evolutionary traits of an organism or lineage can be considered as solutions to problems put forth by natural selection, and hence can be used to work backwards up the evolutionary chain of events. In this thesis the process of eukaryogenesis is broken into three major transitions that led to the evolution of eukaryotic specific traits such as the endomembrane system, organellar protein import, and the cell cycle and sex. These transitions are evaluated from the standpoint of endosymbiotic theory that posits the origin of eukaryotes from an endosymbiotic event between an archaeal host and a bacterial symbiont (Martin and Müller, 1998).

Endosymbiotic theories for eukaryote origin

Since its inception endosymbiotic theory has always been riddled with controversies and debates, with many outright criticisms. With impetus from Lynn Margulis and others the endosymbiotic theory was revived. The role of endosymbiotic interactions in the evolution of the mitochondria and plastids now has a permanent place in biology. Many scientists used endosymbiotic theory to explain the origin of the eukaryotic nucleus, an endomembrane system, the flagella and other aspects of eukaryotic biology. In Martin *et al.*, 2015, models that have invoked endosymbiotic theory for the origin of eukaryotes or otherwise starting from Constantin Mereschkowsky in 1905 to the most recent adaptations are reviewed highlighting the strengths and drawbacks of all these theories with respect to eukaryote origins in light of new phylogenetic evidence which postulate an archaeal host for the endosymbiosis. The ubiquitous presence of the mitochondria in eukaryotes is an observation that needs to be explained by any theory attempting to explain eukaryote origin. Endosymbiosis of a cyanobacterium also gave rise to the archaeplastidal lineage of photosynthetic eukaryotes. The role of anaerobes at eukaryotic origin is consistent with the recent inference that the host that acquired the cyanobacterial endosymbiont was a facultative anaerobe.

Bacterial vesicle secretion and the evolutionary origin of the eukaryotic endomembrane system

The origin of the endomembrane system has always been one of the more challenging questions in eukaryogenesis given the fact that the endomembrane system defines the eukaryotes. Many models have been proposed to explain the origin of the endomembrane system and most invariably start with the invagination of the plasma membrane, while one of the more recent ones, the inside-out model for eukaryote origins starts with *"extruded membrane-bound blebs"* , but nevertheless has the ER lumen homologous with the extracellular environment (Baum and Baum, 2013). The distinctively similar biology of the ER lumen and the Gram-negative periplasm has not been addressed in previous models. In Gould *et al.*, 2016, the secretion of outer membrane vesicles (OMVs) containing periplasm by the endosymbiont

within the cytosol of the host is proposed to have naturally given rise to the components endomembrane system. The model puts into context existing evidence to provide a novel proposal for the series of events that led to the origin of the endomembrane system and vesicular flux.

A central pillar of the proposal is that bacterial vesicles fill the cytosol of the archaeal host, providing an alternative target for the archaeal secretory machinery. Bacterial vesicles also transform the host membrane lipid composition from archaeal ether-linked lipids to their bacterial ester-linked counterparts. Archaeal ESCRTIII proteins which are involved in cell division and capable of bending membranes away from the cytosol, re-model the cytosolic vesicles to form a nucleus in response to invasion by group II introns. The lumen of these vesicles continue to perform their original biology of glycosylation, calcium storage, protein secretion and oxidative protein folding but in the host's cytosol. In eukaryotes this persists to the present day. The observation that mitochondria secrete vesicles (Soubannier *et al.*, 2012; Mohanty and McBride, 2013; Sugiura *et al.*, 2014), which contribute to the formation of the peroxisome and autophagosomes, lends more credence to the idea that the endosymbiont was and is capable of secreting vesicles that form subcellular compartments. The model accounts both for the origin of the endomembrane system and its functions.

Conservation of transit peptide-independent protein import into the mitochondrial and hydrogenosomal matrix

Since their discovery, N-terminal targeting sequences (NTSs) have been a paradigm for protein targeting to mitochondria and related organelles. While there are proteins that have been shown to be targeted to mitochondria independent of an NTS, they are either membrane proteins with internal or C-terminal targeting sequences. Analysis of hydrogenosomal proteins of *Trichomonas vaginalis* showed that some of them can be imported into hydrogenosomes independent of an NTS (Mentel *et al.*, 2008; Zimorski *et al.*, 2013). Furthermore, the targeting information could not be pinpointed to any particular region or domain within the proteins (Zimorski *et al.*, 2013). In the present manuscript it was tested if this propensity of NTS-independent targeting was also extended to mitochondrial import. It could be determined that not only is NTS-independent import conserved in *Saccharomyces cerevisiae*, but also mitochondrial proteins can be targeted to hydrogenosomes independent of their NTSs. This reveals an inherent propensity of ancestral mitochondrial targeting to be NTS-independent. The evolution and persistence of NTSs for mitochondrial import and their dispensability in hydrogenosomal import sheds light on the selection pressures involved in the evolution and maintenance of positively charged NTSs. 'A correlation is found between the presence of the electrochemical gradient ($\Delta\Psi$) of the mitochondria which is essential for protein import and the presence of positively charged NTSs. In the absence of $\Delta\Psi$, as in hydrogenosomes and mitosomes, the selection pressure for maintaining an NTS is lifted and the NTS begins to loose its charge and ultimatelz it is no longer required, consequently reverting organellar protein import to an ancestral state of being NTS-independent.

N-Terminal presequence-independent import of phosphofructokinase into hydrogenosomes of *Trichomonas vaginalis*

The conservation of NTS-independent targeting to organelles was further explored in Rada *et al.*, 2015 using the ATP-dependent phosphofructokinase (ATP-PFK), an enzyme of the glycolytic pathway, in both bacteria and eukaryotes. Not only were isoforms of *T.vaginalis* ATP-PFK imported into the hydrogenosomes independent of their NTSs, *Tv*ATP-PFK was also imported into *S.cerevisiae* mitochondria. As in the case with proteins from Garg *et al.*, 2015 *Tv*ATP-PFK showed an intrinsic propensity to get targeted to the *S.cerevisiae* mitochondria. A closer look at *Sc*ATP-PFK showed that it contains a negatively charged N-terminal domain that is absent in *Tv*ATP-PFK. Deletion of this N-terminus allowed the *S.cervesiae* ATP-PFK to be associated with mitochondria. The presence of the negatively charged NTS of *Sc*ATP-PFK had no effect on hydrogenosome association in *T.vaginalis* and removal of the negatively charged N-terminal domain led to complete import into hydrogenosomes. In addition, it was also observed that *Escherichia coli* ATP-PFK could be imported into hydrogenosomes. The presence of a dedicated NTS allows for specific binding and recognition of proteins by the import apparatus at lower concentrations without interfering with the functional regions of the proteins. The loss of a specific sequence determining specificity would increase the concentration of protein required for binding and recognition. This explains the varying degrees of organellar localization when using different promoters. The addition of the negatively charged N-terminal domain was selected and maintained for the purpose of exclusion of proteins from the mitochondria. The role of negatively charged N-terminal domains has not been hitherto investigated when it comes to mitochondrial protein targeting, which may be important for shaping and re-compartmentalizing the metabolism of eukaryotes following endosymbiosis (Martin, 2010).

The role of charge in protein targeting evolution

The role of mitochondrial biology in the selection of positively or negatively charged N-termini for facilitating or impeding import into the organelle prompted investigation of organisms with plastids. The presence of the mitochondria and a fully evolved mitochondrial import machinery posed a unique challenge to the evolution of plastid protein import. In Garg and Gould, 2016 peptide characteristics of N-termini of plastid targeted proteins and proteins targeted to the mitochondria were analyzed, upon on which it was observed that the plastid NTSs (pNTSs) on average has a significantly lower positive charge than mitochondrial NTSs (mNTSs). It was previously known that the serine residues enriched in pNTSs are prone to be phosphorylated, although the necessity of this phosphorylation for specificity of import has been questioned. It could be reasoned that in the case of the newly emerging class of dual-targeted proteins that target both the mitochondria and plastids, phosphorylation of a positively charged NTS could direct targeting to the plastid. In support of this view, it is observed that the NTSs of dual targeted proteins often have a higher positive charge than plastid targeted proteins, while simultaneously being enriched in phosphorylatable serines. Presence of mitochondria in the same cytosol as the plastid selected for less positively charged NTSs in proteins targeted to plastids. This selection was, however, lifted in the case of secondary plastids where nucleomorph-encoded proteins that are targeted to the plastids now have a positively charged N-terminus. It has previously been shown that in some cases, the magnitude of positive charge is sufficient to impart specificity to import, which explains the lack of a consensus sequences regarding protein targeting to mitochondria and plastids concurrently, providing clues to the evolution of organellar protein import.

Mitochondria, the cell cycle, and the origin of sex via a syncytial eukaryote common ancestor

While many theories exists for the origin of sex in eukaryotes very few address the origin of the cell cycle (Nasmyth, 1995; Novak *et al.*, 1998; Krylov *et al.*, 2003). In Garg and Martin, 2016 an inference for the origin of sex in a mitochondrion possessing archaeal cell is presented which is essence posits the following. Symbiotic association led to the origin of mitochondria and gene transfer to the host genome. This precipitated the origin of introns and the origin of the nucleus, giving rise to a dedicated translational compartment, the eukaryotic cytosol, in which — by virtue of mitochondria — metabolic energy was not limiting. This permitted, spontaneous ATP-dependent processes of protein aggregation and assembly in the cytosol, processes that could become selectable and that ultimately gave rise to continuous microtubule dependent chromosome segregation (MDCS) and reductive division. The chromosome heterogeneity that resulted from random endosymbiotic gene transfer is rescued in a filamentous, multinucleated syncytial cell. A syncytial ancestor accounts for the massive paralogy of genes observed in eukaryotes, while accounting for the conserved nature of chromosome division versus the variation in cell-division and life cycles in eukaryotes. Recombination and repair alongside MDCS within the syncytial ancestor provided the initial framework for the evolution of the cell cycle which would have resembled meiosis from which haploid and diploid mitosis could easily be derived. A universal presence of multinucleated organisms in the eukaryotic supergroups, which remains to be tested, would provide some clues to the conserved nature of a syncytial common ancestor of eukaryotes.

Concluding remarks

In the context of eukaryote origins, the term "major transitions" is used to encompass all the eukaryotic traits that are absent in prokaryotes. It is true that many eukaryotic proteins share homology with their prokaryotic counterparts, in fact it is the cornerstone of the endosymbiotic theory for eukaryote origin. However, biology life is often realized as the emergent property of the components building it — in this case proteins and enzymes. It is the unique collection of proteins and the equally unique functional network they formed that allowed eukaryotes to evolve.

In the most basic view of eukaryote origin, one has a starting point (the prokaryotes), an end point (extant eukaryotes), and a black box in between. It is often easy to attribute something fanciful to this black box as a means to get from prokaryotes to eukaryotes but, Occam's razor implores us to hold on tight to the observations and to use a minimum of imagination. There are many theories and models that strive to explain the origin of eukaryotes but either they are based on non-parsimonious assumptions or fail to explain all aspects of eukaryogenesis (reviewed in Martin *et al.*, 2015).

This thesis attempts to push the boundaries of the endosymbiotic theory of eukaryote origin as postulated by the hydrogen hypothesis (Martin and Müller, 1998) to explain some of the notable major transitions in eukaryote origins. outer membrane vesicle (OMV) secretion by the mitochondrial endosymbiont provides the necessary framework for the selection and evolution of the endomembrane system and sub-cellular complexity (Gould *et al.*, 2016). Endosymbiotic gene transfer (EGT) results in the origin of introns and the evolution of the nucleus. EGT also necessitates the evolution of protein import mechanisms into the mitochondria, which uses the electrochemical gradient of the organelle to impart specificity to the targeted proteins (Garg *et al.*, 2015; Rada *et al.*, 2015; Garg and Gould, 2016). The genetic heterogeneity that would result from random EGT drives the evolution of recombination and repair in a syncytial common ancestor of eukaryotes (Garg and Martin, 2016). The strength this inference is that it provides a mutually consistent framework for the evolution of eukaryotic traits.

Endosymbiosis drastically changed the evolutionary landscape and the selection pressures which compelled the evolution of new biological processes. Nevertheless, evolution can only improve upon pre-existing building blocks, which in the case of eukaryote origins are the building blocks contributed by the archaeal host and bacterial endosymbiont. Three major transitions at eukaryote origin were addressed in this thesis namely (i) origin of the eukaryotic endomembrane system, (ii) protein import to the mitochondria and (iii) cell cycle and sex.

Examples of some other major transitions include true multicellularity which is a unique eukaryotic trait. While prokaryotes may form colonies and work in consort they do not reproduce as a single unit and although clues suggest that multicellularity may have evolved multiple times (Knoll, 2011); the capacity to evolve multicellularity is rooted in eukaryote origin. The same can be said about apoptosis or programmed

cell death, where the mitochondria plays a central role in regulating and orchestrating apoptosis (Bhola and Letai, 2016). The mitochondria interact with other sub-compartments in the eukaryotic cell for functions that extend beyond simple lipid and ion exchange (Murley and Nunnari, 2016; Gatta and Levine, 2016). There are even implications of mitochondria in cancer and countless other pathological diseases (Evans and Neuman, 2016). In all eukaryotic processes studied to date the mitochondria play a central role. In light of its pivotal role in eukaryotic biology the contribution of the mitochondria in shaping the last common ancestor of eukaryotes and their role in the major transitions at eukaryote origin merit continued attention.

References

Aebi M (2013). N-linked protein glycosylation in the ER. *BBA - Mol. Cell Res.* **1833**: 2430–2437.

Alberts B (2008). *Molecular Biology of the Cell*. Garland Science, New York.

Allen JF (1996). Separate sexes and the mitochondrial theory of ageing. *J. Theor. Biol.* **180**: 135–140.

Altmann R (1890). *Die Elementarorganismen und ihre Beziehungen zu den Zellen*. Verlag von Veit & Comp., Leipzig.

Archibald J (2014). *One Plus One Equals One*. Oxford University Press, Oxford.

Arnold T, Poynor M, Nussberger S, Lupas AN, and Linke D (2007). Gene duplication of the eight-stranded β-barrel Ompx produces a functional pore: a scenario for the evolution of transmembrane β-barrels. *J. Mol. Biol.* **366**: 1174–1184.

Barton NH and Charlesworth B (1998). Why sex and recombination? *Science* **281**: 1986–1990.

Baum DA and Baum B (2013). An inside-out origin for the eukaryotic cell. *BMC Biol* **12**: 76–76.

Becker T, Böttinger L, and Pfanner N (2012). Mitochondrial protein import: from transport pathways to an integrated network. *Trends Biochemi. Sci.* **37**: 85–91.

Beijerinck MW (1942). Concerning a contagium vivum fluidum as a cause of the spot-disease of tobacco leaves. Translated from "Über ein Contagium vivum fluidum als Ursache der Fleckenkrankheit der Tabaksblätter" (1898). *Verhandelingen der Koninklijke akademie van Wetenschappen te Amsterdam* (in German). **65**: 1–22.

Bell G (1982). *The Masterpiece of Nature: the Evolution and Genetics of Sexuality*. University of California Press, Berkeley.

Bell G (1988). *Sex and Death in Protozoa*. Cambridge University Press, Cambridge.

Bendich AJ and Drlica K (2000). Prokaryotic and eukaryotic chromosomes: what's the difference. *BioEssays* **22**: 481–486.

Bentley SD, Chater KF, Cerdeno-Tarraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D, Bateman A, Brown S, Chandra G, Chen CW, Collins M, Cronin A, Fraser A, Goble A, Hidalgo J, Hornsby T, Howarth S, Huang CH, Kieser T, Larke L, Murphy L, Oliver K, O'Neil S, Rabbinowitsch E, Rajandream MA, Rutherford K, Rutter S, Seeger K, Saunders D, Sharp S, Squares R, Squares S, Taylor K, Warren T, Wietzorrek A, Woodward J, Barrell BG, Parkhill J, and Hopwood DA (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor A3(2)*. *Nature* **417**: 141–147.

Bernstein C and Bernstein H (1991). *Aging, Sex, and DNA Repair*. Academic Press, Michigan.

Bernstein H and Bernstein C (2010). Evolutionary origin of recombination during meiosis. *BioScience* **60**: 498–505.

Bhola PD and Letai A (2016). Mitochondria — judges and executioners of cell death sentences. *Mol. Cell* **61**: 695–704.

Biondi EG, Reisinger SJ, Skerker JM, Arif M, Perchuk BS, Ryan KR, and Laub MT (2006). Regulation of the bacterial cell cycle by an integrated genetic circuit. *Nature* **444**: 899–904.

Blom T, Somerharju P, and Ikonen E (2011). Synthesis and biosynthetic trafficking of membrane lipids. *Cold Spring Harb. Perspect. Biol.* **3**: a004713–a004713.

Boettcher B and Barral Y (2014). The cell biology of open and closed mitosis. *Nucleus* **4**: 160–165.

Bölter B and Soll J (2016). Once upon a time – chloroplast protein import research from infancy to future challenges. *Mol. Plant* **9**: 798–812.

Borst P (1972). Mitochondrial nucleic-acids. *Annu. Rev. Biochem.* **41**: 333–&.

Bottomley W and Bohnert HJ (1982). The biosynthesis of chloroplast proteins. In B Parthier and D Boulter (Eds.) *Nucleic Acids and Proteins in Plants II*, pp. 531–596. Springer Berlin Heidelberg.

Brandizzi F and Barlowe C (2013). Organization of the ER–Golgi interface for membrane traffic control. *Nat. Rev. Mol. Cell. Biol.* **14**: 382–392.

Brown MT, Goldstone HMH, Bastida-Corcuera F, Delgadillo-Correa MG, McArthur AG, and Johnson PJ (2007). A functionally divergent hydrogenosomal peptidase with protomitochondrial ancestry. *Mol. Microbiol.* **64**: 1154–1163.

Brown L, Wolf JM, Prados-Rosales R, and Casadevall A (2015). Through the wall: extracellular vesicles in Gram-positive bacteria, mycobacteria and fungi. *Nat. Rev. Microbiol.* **13**: 620–630.

Bui E, Bradley PJ, and Johnson PJ (1996). A common evolutionary origin for mitochondria and hydrogenosomes. *Proc. Natl. Acad. Sci. U.S.A.* **93**: 9651–9656.

Burri L, Williams BAP, Bursac D, Lithgow T, and Keeling PJ (2006). Microsporidian mitosomes retain elements of the general mitochondrial targeting system. *Proc. Natl. Acad. Sci. U.S.A.* **103**: 15916–15920.

Camerini-Otero RD and Hsieh P (1995). Homologous recombination proteins in prokaryotes and eukaryotes. *Annu. Rev. Genet.* **29**: 509–552.

Carballido-Lopez R (2006). The bacterial actin-like cytoskeleton. *Microbiol. Mol. Biol. Rev.* **70**: 888–909.

Cavalier-Smith T (1987). The origin of eukaryote and archaebacterial cells. *Ann. N.Y. Acad. Sci.* **503**: 17–54.

Cavalier-Smith T (2002). Origins of the machinery of recombination and sex. *Heredity* **88**: 125–141.

Cavalier-Smith T (2010). Origin of the cell nucleus, mitosis and sex: roles of intracellular coevolution. *Biol. Direct* **5**: 7.

Cavicchioli R (2011). Archaea — timeline of the third domain. *Nat. Rev. Microbiol.* **9**: 51–61.

Chacinska A, Koehler CM, Milenkovic D, Lithgow T, and Pfanner N (2009). Importing mitochondrial proteins: machineries and mechanisms. *Cell* **138**: 628–644.

Chatton E (1925). *Pansporella perplexa*. Réflexions sur la biologie et la phylogenie des protozoaires. *Ann. Sci. Nat. Zool* **VII**: 1–84.

Chua NH and Schmidt GW (1978). Post-translational transport into intact chloroplasts of a precursor to the small subunit of ribulose-1,5-bisphosphate carboxylase. *Proc. Natl. Acad. Sci. U.S.A.* **75**: 6110–6114.

Clements A, Bursac D, Gatsos X, Perry AJ, Civciristov S, Celik N, Likic VA, Poggio S, Jacobs-Wagner C, Strugnell RA, and Lithgow T (2009). The reducible complexity of a mitochondrial molecular machine. *Proc. Natl. Acad. Sci. U.S.A.* **106**: 15791–15795.

Cleveland LR (1947). The origin and evolution of meiosis. *Science* **105**: 287–289.

Cohan FM and Aracena S (2012). Prokaryotic sex: Eukaryote-like qualities of recombination in an archaean lineage. *Curr. Biol.* **22**: R601–R602.

Cooper S (2006). *Bacterial Growth and Division*. Wiley-VCH Verlag GmbH & Co., Weinheim.

Cox CJ, Foster PG, Hirt RP, Harris SR, and Embley TM (2008). The archaebacterial origin of eukaryotes. *Proc. Natl. Acad. Sci. U.S.A.* **105**: 20356–20361.

Crow JF (2005). Hermann Joseph Muller, evolutionist. *Nat. Rev. Genet.* **6**: 941–945.

Dacks JB, Field MC, Buick R, Eme L, Gribaldo S, Roger AJ, Brochier-Armanet C, and Devos DP (2016). The changing view of eukaryogenesis - fossils, cells, lineages and how they all come together. *J. Cell Sci.* **129**: 3695–3703.

De Duve C (1969). Evolution of the peroxisome. *Ann. N.Y. Acad. Sci.* **168**: 369–381.

De Hertogh B, Lantin AC, Baret PV, and Goffeau A (2004). The archaeal P-type ATPases. *J. Bioenerg. Biomembr.* **36**: 135–142.

Dey G, Thattai M, and Baum B (2016). On the archaeal origins of eukaryotes and the challenges of inferring phenotype from genotype. *Trends Cell Biol.* **26**: 476–485.

Di Maio R, Barrett PJ, Hoffman EK, Barrett CW, Zharikov A, Borah A, Hu X, McCoy J, Chu CT, Burton EA, Hastings TG, and Greenamyre JT (2016). α-Synuclein binds to TOM20 and inhibits mitochondrial protein import in Parkinson's disease. *Sci. Transl. Med.* **8**: 342ra78–342ra78.

Doherty GJ and McMahon HT (2009). Mechanisms of endocytosis. *Annu. Rev. Biochem.* **78**: 857–902.

von Dohlen CD, Kohler S, Alsop ST, and McManus WR (2001). Mealybug beta-proteobacterial endosymbionts contain gamma-proteobacterial symbionts. *Nature* **412**: 433–436.

Doležal P, Likic V, Tachezy J, and Lithgow T (2006). Evolution of the molecular machines for protein import into mitochondria. *Science* **313**: 314–318.

Doolittle WF (1980). Revolutionary concepts in evolutionary cell biology. *Trends Biochemi. Sci.* **5**: 146–149.

van Dooren GG, Yeoh LM, Striepen B, and McFadden GI (2016). The import of proteins into the mitochondrion of *Toxoplasma gondii*. *J. Biol. Chem.* **291**: 19335–19350.

Dudek J, Rehling P, and van der Laan M (2013). Mitochondrial protein import: common principles and physiological networks. *Biochim. Biophys. Acta* **1833**: 274–285.

Duggin IG, Aylett CHS, Walsh JC, Michie KA, Wang Q, Turnbull L, Dawson EM, Harry EJ, Whitchurch CB, Amos LA, and Löwe J (2015). CetZ tubulin-like proteins control archaeal cell shape. *Nature* **519**: 362–365.

Dyer N (2009). Tubulin and its prokaryotic homologue FtsZ: a structural and functional comparison. *Sci. Prog.* **92**: 113–137.

Egan AJ and Vollmer W (2013). The physiology of bacterial cell division. *Ann. N.Y. Acad. Sci.* **1277**: 8–28.

Ellis RJ (1981). Chloroplast proteins: synthesis, transport, and assembly. *Ann. Rev. Plant Physiol.* **32**: 111–137.

Embley TM, Van der Giezen M, Horner DS, Dyal PL, and Foster P (2003). Mitochondria and hydrogenosomes are two forms of the same fundamental organelle. *Phil. Trans. R. Soc. Lond. B.* **358**: 191–201.

Embley TM and Hirt RP (1998). Early branching eukaryotes? *Curr. Opin. Genet. Dev.* **8**: 624–629.

Embley TM and Martin W (2006). Eukaryotic evolution, changes and challenges. *Nature* **440**: 623–630.

Evans A and Neuman N (2016). The mighty mitochondria. *Mol. Cell* **61**: 641.

Fader CM and Colombo MI (2009). Autophagy and multivesicular bodies: two closely related partners. *Cell Death Differ.* **16**: 70–78.

Feigel A, Englander A, and Engel A (2009). Sex is always well worth its two-fold cost. *PLoS ONE* **4**: e6012.

Felsenstein J (1974). The evolutionary advantage of recombination. *Genetics* **78**: 737–756.

Field MC, Sali A, and Rout MP (2011). On a bender-BARs, ESCRTs, COPs, and finally getting your coat. *J. Cell Biol.* **193**: 963–972.

Flannagan RS, Jaumouillé V, and Grinstein S (2012). The cell biology of phagocytosis. *Annu. Rev. Pathol.* **7**: 61–98.

Franzén LG, Rochaix JD, and von Heijne G (1989). Chloroplast transit peptides from the green alga *Chlamydomonas reinhardtii* share features with both mitochondrial and higher plant chloroplast presequences. *FEBS Lett.* **260**: 165–168.

Fraser CM, Eisen JA, and Salzberg SL (2000). Microbial genome sequencing. *Nature* **406**: 799–803.

Freeman SA and Grinstein S (2014). Phagocytosis: receptors, signal integration, and the cytoskeleton. *Immunol. Rev.* **262**: 193–215.

Garg SG and Gould SB (2016). The role of charge in protein targeting evolution. *Trends Cell Biol.* **26**: 894–905.

Garg SG and Martin WF (2016). Mitochondria, the cell cycle, and the origin of sex via a syncytial eukaryote common ancestor. *Genome Biol. Evol.* **8**: 1950–1970.

Garg S, Stölting J, Zimorski V, Rada P, Tachezy J, Martin WF, and Gould SB (2015). Conservation of transit peptide-independent protein import into the mitochondrial and hydrogenosomal matrix. *Genome Biol. Evol.* **7**: 2716–2726.

Gatta AT and Levine TP (2016). Piecing together the patchwork of contact sites. *Trends Cell Biol.* **10.1016/j.tcb.2016.08.010**.

Gauthier NC, Rossier OM, Mathur A, Hone JC, and Sheetz MP (2009). Plasma membrane area increases with spread area by exocytosis of a GPI-anchored protein compartment. *Mol. Biol. Cell* **20**: 3261–3272.

Geissler A, Krimmer T, Bömer U, Guiard B, Rassow J, and Pfanner N (2000). Membrane potential-driven protein import into mitochondria. . *Mol. Biol. Cell* **11**: 3977–3991.

Geissler A, Rassow J, Pfanner N, and Voos W (2001). Mitochondrial import driving forces: enhanced trapping by matrix Hsp70 stimulates translocation and reduces the membrane potential dependence of loosely folded preproteins. *Mol. Cell. Biol.* **21**: 7097–7104.

Germot A, Philippe H, and Le Guyader H (1996). Presence of a mitochondrial-type 70-kDa heat shock protein in *Trichomonas vaginalis* suggests a very early mitochondrial endosymbiosis in eukaryotes. *Proc. Natl. Acad. Sci. U.S.A.* **93**: 14614–14617.

Van der Giezen M (2009). Hydrogenosomes and mitosomes: conservation and evolution of functions. *J. Eukaryot. Microbiol.* **56**: 221–231.

Gilmore R, Blobel G, and Walter P (1982). Protein translocation across the endoplasmic reticulum. I. Detection in the microsomal membrane of a receptor for the signal recognition particle. *J. Cell Biol.* **95**: 463–469.

Glick BS and Nakano A (2009). Membrane traffic within the golgi apparatus. *Annu. Rev. Cell Dev. Biol.* **25**: 113–132.

Goldberg AV, Molik S, Tsaousis AD, Neumann K, Kuhnke G, Delbac F, Vivares CP, Hirt RP, Lill R, and Embley TM (2008). Localization and functionality of microsporidian iron–sulphur cluster assembly proteins. *Nature* **452**: 624–628.

Goodenough U and Heitman J (2014). Origins of eukaryotic sexual reproduction. *Cold Spring Harb. Perspect. Biol.* **6**: a016154–a016154.

Gould SB, Garg SG, and Martin WF (2016). Bacterial vesicle secretion and the evolutionary origin of the eukaryotic endomembrane system. *Trends Microbiol.* **24**: 525–534.

Gray MW (1983). The bacterial ancestry of plastids and mitochondria. *BioScience* **33**: 693–699.

Gray MW and Doolittle WF (1982). Has the endosymbiont hypothesis been proven? *Microbiol. Rev.* **46**: 1–42.

Gribaldo S, Poole AM, Daubin V, Forterre P, and Brochier-Armanet C (2010). The origin of eukaryotes and their relationship with the Archaea: are we at a phylogenomic impasse? *Nat. Rev. Microbiol.* **8**: 743–752.

Gross J and Bhattacharya D (2010). Uniting sex and eukaryote origins in an emerging oxygenic world. *Biol. Direct* **5**: 53.

Guy L and Ettema TJG (2011). The archaeal 'TACK' superphylum and the origin of eukaryotes. *Trends Microbiol.* **19**: 580–587.

Haas A (2007). The phagosome: compartment with a license to kill. *Traffic* **8**: 311–330.

Hartwell LH and Weinert TA (1989). Checkpoints: controls that ensure the order of cell cycle events. *Science* **246**: 629–634.

Havird JC, Hall MD, and Dowling DK (2015). The evolution of sex: A new hypothesis based on mitochondrial mutational erosion: Mitochondrial mutational erosion in ancestral eukaryotes would favor the evolution of sex, harnessing nuclear recombination to optimize compensatory nuclear coadaptation. *BioEssays* **37**: 951–958.

von Heijne G (1986). Mitochondrial targeting sequences may form amphiphilic helices. *EMBO J.* **5**: 1335–471.

von Heijne G and Abrahmsèn L (2001). Species-specific variation in signal peptide design implications for protein secretion in foreign hosts. *FEBS Lett.* **244**: 439–446.

Heitman J (2015). Evolution of sexual reproduction: a view from the fungal kingdom supports an evolutionary epoch with sex before sexes. *Fungal Biol. Rev.* **29**: 108–117.

Herrmann JM and Riemer J (2014). Three approaches to one problem: protein folding in the periplasm, the endoplasmic reticulum, and the intermembrane space. *Antioxid. Redox Signal.* **21**: 438–456.

Heywood P and Magee PT (1976). Meiosis in protists. Some structural and physiological aspects of meiosis in algae, fungi, and protozoa. *Bacteriol Rev* **40**: 190–240.

Hooke R (1665). *Micrographia or Some Physiological Descriptions of Minute Bodies Made by Magnifying Glasses with Observations and Inquiries Thereupon*. Jo. Martyn and Ja. Allestry, London.

Hörandl E (2009). A combinational theory for maintenance of sex. *Heredity* **103**: 445–457.

Hörandl E and Hadacek F (2013). The oxidative damage initiation hypothesis for meiosis. *Plant Reprod.* **26**: 351–367.

Horner DS, Hirt RP, Kilvington S, Lloyd D, and Embley TM (1996). Molecular data suggest an early acquisition of the mitochondrion endosymbiont. *Proc. R. Soc. B* **263**: 1053–1059.

Hsu SC and Inoue K (2009). Two evolutionarily conserved essential beta-barrel proteins in the chloroplast outer envelope membrane. *Biosci Trends* **3**: 168–178.

Hurst LD and Nurse P (1991). A note on the evolution of meiosis. *J. Theor. Biol.* **150**: 561–563.

Husnik F and McCutcheon JP (2016). Repeated replacement of an intrabacterial symbiont in the tripartite nested mealybug symbiosis. *Proc. Natl. Acad. Sci. U.S.A.* **113**: E5416–24.

Husnik F, Nikoh N, Koga R, Ross L, Duncan RP, Fujie M, Tanaka M, Satoh N, Bachtrog D, Wilson ACC, von Dohlen CD, Fukatsu T, and McCutcheon JP (2013). Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. *Cell* **153**: 1567–1578.

Hutson V and Law R (1993). Four steps to two sexes. *Proc. R. Soc. B* **253**: 43–51.

Iwanowski D (1892). "*Über die Mosaikkrankheit der Tabakspflanze*". Bulletin Scientifique publié par l'Académie Impériale des Sciences de Saint-Pétersbourg / Nouvelle Serie III (in German and Russian). St. Petersburg. 35: 67—70. Translated into English in Johnson, J., Ed. (1942) *Phytopathological classics*. *Am. Phytopathol. Soc* **7**: 27–30.

Jones HE, Holland IB, and Campbell AK (2002). Direct measurement of free Ca^{2+} shows different regulation of Ca^{2+}between the periplasm and the cytosol of *Escherichia coli*. *Cell Calcium* **32**: 183–192.

Karlin-Neumann GA and Tobin EM (1986). Transit peptides of nuclear-encoded chloroplast proteins share a common amino acid framework. *EMBO J.* **5**: 9–13.

Kennelly PJ (2014). Protein Ser/Thr/Tyr phosphorylation in the Archaea. *J. Biol. Chem.* **289**: 9480–9487.

Klenk HP and Doolittle WF (1994). Evolution: Archaea and eukaryotes versus bacteria? *Curr. Biol.* **4**: 920–922.

Klinger CM, Spang A, Dacks JB, and Ettema TJG (2016). Tracing the archaeal origins of eukaryotic membrane-trafficking system building blocks. *Mol. Biol. Evol.* **33**: 1528–1541.

Klingl A (2014). S-layer and cytoplasmic membrane – exceptions from the typical archaeal cell wall with a focus on double membranes. *Front. Microbiol.* **5**: 351.

Knoll AH (2011). The multiple origins of complex multicellularity. *Annu. Rev. Earth Planet. Sci.* **39**: 217–239.

Koch GL (1990). The endoplasmic reticulum and calcium storage. *BioEssays* **12**: 527–531.

Kondrashov AS (1988). Deleterious mutations and the evolution of sexual reproduction. *Nature* **336**: 435–440.

Kondrashov AS (1993). Classification of hypotheses on the advantage of amphimixis. *J. Hered.* **84**: 372–387.

Kondrashov AS (1994). The asexual ploidy cycle and the origin of sex. *Nature* **370**: 213–216.

Kondrashov AS (1997). Evolutionary genetics of life cycles. *Ann. Rev. Ecol. Syst.* **28**: 391–435.

Kondrashov AS and Crow JF (1991). Haploidy or diploidy: which is better? *Nature* **351**: 314–315.

Koumandou VL, Wickstead B, Ginger ML, Van der Giezen M, Dacks JB, and Field MC (2013). Molecular paleontology and complexity in the last eukaryotic common ancestor. *Crit. Rev. Biochem. Mol. Biol.* **48**: 373–396.

Kreil G (1981). Transfer of proteins across membranes. *Annu. Rev. Biochem.* **50**: 317–348.

Krylov DM, Nasmyth K, and Koonin EV (2003). Evolution of eukaryotic cell cycle regulation: stepwise addition of regulatory kinases and late advent of the CDKs. *Curr. Biol.* **13**: 173–177.

Lake JA, Henderson E, Oakes M, and 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**: 3786–3790.

Lane N (2014). Bioenergetic constraints on the evolution of complex life. *Cold Spring Harb. Perspect. Biol.* **6**: a015982–a015982.

Lane N (2015). The unseen world: reflections on Leeuwenhoek (1677) 'Concerning little animals'. *Phil. Trans. R. Soc. Lond. B.* **370**: 20140344–20140344.

Lane N and Martin W (2010). The energetics of genome complexity. *Nature* **467**: 929–934.

Lane N and Martin WF (2016). Mitochondria, complexity, and evolutionary deficit spending. *Proc. Natl. Acad. Sci. U.S.A.* **113**: E666–E666.

Lang A, John Peter AT, and Kornmann B (2015). ER–mitochondria contact sites in yeast: beyond the myths of ERMES. *Curr. Opin. Cell Biol.* **35**: 7–12.

Lechevalier H (1972). Dmitri Iosifovich Ivanovski (1864-1920). *Bacteriol Rev* **36**: 135–145.

Lewis Jr WM (1987). The cost of sex. In SC Stearns (Ed.) *The Evolution of Sex and its Consequences*, pp. 33–57. The evolution of sex and its consequences. Birkhäuser Verlag, Basel.

Lindås AC and Bernander R (2013). The cell cycle of archaea. *Nat. Rev. Microbiol.* **11**: 627–638.

Lippincott-Schwartz J, Roberts TH, and Hirschberg K (2000). Secretory protein trafficking and organelle dynamics in living cells. *Annu. Rev. Cell Dev. Biol.* **16**: 557–589.

Lister R, Hulett JM, Lithgow T, and Whelan J (2005). Protein import into mitochondria: origins and functions today (Review). *Mol. Membr. Biol.* **22**: 87–100.

Lithgow T and Schneider A (2010). Evolution of macromolecular import pathways in mitochondria, hydrogenosomes and mitosomes. *Phil. Trans. R. Soc. Lond. B.* **365**: 799–817.

Lodish H (2008). *Molecular Cell Biology*. Macmillan.

Loidl J (2016). Conservation and variability of meiosis across the eukaryotes. *Annu. Rev. Genet.* **doi: 10.1146/annurev-genet-120215-035100**.

Luzio JP, Pryor PR, and Bright NA (2007). Lysosomes: fusion and function. *Nat. Rev. Mol. Cell. Biol.* **8**: 622–632.

Lynch M, Butcher D, Burger R, and Gabriel W (1993). The mutational meltdown in asexual populations. *J. Heredity* **84**: 339–344.

Lynch M and Marinov GK (2015). The bioenergetic costs of a gene. *Proc. Natl. Acad. Sci. U.S.A.* **112**: 15690–15695.

Maccecchini ML, Rudin Y, Blobel G, and Schatz G (1979). Import of proteins into mitochondria - precursor forms of the extra-mitochondrially made F1-ATPase subunits in yeast. *Proc. Natl. Acad. Sci. U.S.A.* **76**: 343–347.

Maguire MP (1992). The evolution of meiosis. *J. Theor. Biol.* **154**: 43–55.

Makarova KS, Yutin N, Bell SD, and Koonin EV (2010). Evolution of diverse cell division and vesicle formation systems in Archaea. *Nat. Rev. Microbiol.* **8**: 731–741.

Margulis L and Sagan D (1986). *Origins of Sex: Three Billion Years of Genetic Recombination*. Yale University Press, New Haven.

Markov AV and Kaznacheev IS (2016). Evolutionary consequences of polyploidy in prokaryotes and the origin of mitosis and meiosis. *Biol. Direct* **11**: 1.

Marshansky V, Rubinstein JL, and Grueber G (2014). Eukaryotic V-ATPase: Novel structural findings and functional insights. *BBA - Bioenergetics* **1837**: 857–879.

Martijn J and Ettema TJG (2013). From archaeon to eukaryote: the evolutionary dark ages of the eukaryotic cell. *Biochem. Soc. Trans.* **41**: 451–457.

Martin W (2010). Evolutionary origins of metabolic compartmentalization in eukaryotes. *Phil. Trans. R. Soc. Lond. B.* **365**: 847–855.

Martin WF, Garg S, and Zimorski V (2015). Endosymbiotic theories for eukaryote origin. *Phil. Trans. R. Soc. Lond. B.* **370**: 20140330.

Martin W and Kowallik K (1999). Annotated english translation of Mereschkowsky's 1905 paper 'Über Natur und Ursprung der Chromatophoren im Pflanzenreiche'. *European Journal of Phycology* **34**: 287–295.

Martin J, Mahlke K, and Pfanner N (1991). Role of an energized inner membrane in mitochondrial protein import. Delta psi drives the movement of presequences. *J. Biol. Chem.* **266**: 18051–18057.

Martin W and Müller M (1998). The hydrogen hypothesis for the first eukaryote. *Nature* **392**: 37–41.

Masai H, Matsumoto S, You Z, Yoshizawa-Sugata N, and Oda M (2010). Eukaryotic chromosome DNA replication: where, when, and how? *Annu. Rev. Biochem.* **79**: 89–130.

Mast FD, Barlow LD, Rachubinski RA, and Dacks JB (2014). Evolutionary mechanisms for establishing eukaryotic cellular complexity. *Trends Cell Biol.* **24**: 435–442.

Matlin KS (2013). History of the Signal Hypothesis. *eLS* .

May T and Soll J (2000). 14-3-3 proteins form a guidance complex with chloroplast precursor proteins in plants. *Plant Cell* **12**: 53–64.

Maynard Smith J (1978). *The Evolution of Sex*. Cambridge University Press, Cambridge.

McBride HM, Silvius JR, and Shore GC (1995). Insertion of an uncharged polypeptide into the mitochondrial inner membrane does not require a trans-bilayer electrochemical potential: effects of positive charges. *BBA - Biomembranes* **1237**: 162–168.

McInerney JO, O'Connell MJ, and Pisani D (2014). The hybrid nature of the Eukaryota and a consilient view of life on Earth. *Nat. Rev. Microbiol.* **12**: 449–455.

Mentel M, Zimorski V, Haferkamp P, Martin W, and Henze K (2008). Protein import into hydrogenosomes of *Trichomonas vaginalis* involves both N-terminal and internal targeting signals: a case study of thioredoxin reductases. *Eukaryot. Cell* **7**: 1750–1757.

Merdanovic M, Clausen T, Kaiser M, Huber R, and Ehrmann M (2011). Protein quality control in the bacterial periplasm. *Annu. Rev. Microbiol.* **65**: 149–168.

Mereschkowsky C (1905). Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. *Biol. Centralbl.* **25**: 593–604.

Meyer DI, Krause E, and Dobberstein B (1982). Secretory protein translocation across membranes—the role of the 'docking protein'. *Nature* **297**: 647–650.

Mindell JA (2012). Lysosomal acidification mechanisms. *Annu. Rev. Physiol.* **74**: 69–86.

Missiakas D and Raina S (1997). Protein folding in the bacterial periplasm. *J. Bacteriol.* **179**: 2465–2471.

Mohanty A and McBride HM (2013). Emerging roles of mitochondria in the evolution, biogenesis, and function of peroxisomes. *Front. Physiol.* **4**: 268.

Morgan DO (1997). Cyclin-dependent kinases: engines, clocks, and microprocessors. *Annu. Rev. Cell Dev. Biol.* **13**: 261–291.

Mulkidjanian AY, Makarova KS, Galperin MY, and Koonin EV (2007). Inventing the dynamo machine: the evolution of the F-type and V-type ATPases. *Nat. Rev. Microbiol.* **5**: 892–899.

Muller HJ (1964). The relation of recombination to mutational advance. *Mutat. Res.* **106**: 2–9.

Müller M (1973). Peroxisomes and hydrogenosomes in protozoa. *J. Histochem. Cytochem.* **21**: 955–957.

Müller M, Mentel M, van Hellemond JJ, Henze K, Woehle C, Gould SB, Yu RY, Van der Giezen M, Tielens AGM, and Martin WF (2012). Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. *Microbiol. Mol. Biol. Rev.* **76**: 444–495.

Murcha MW, Elhafez D, Lister R, Tonti-Filippini J, Baumgartner M, Philippar K, Carrie C, Mokranjac D, Soll J, and Whelan J (2006). Characterization of the preprotein and amino acid transporter gene family in Arabidopsis. *Plant Physiol.* **143**: 199–212.

Murley A and Nunnari J (2016). The emerging network of mitochondria-organelle contacts. *Mol. Cell* **61**: 648–653.

Murray AW (2004). Recycling the cell cycle: cyclins revisited. *Cell* **116**: 221–234.

Nasmyth K (1995). Evolution of the cell-cycle. *Phil. Trans. R. Soc. Lond. B.* **349**: 271–281.

Nei M (2013). *Mutation-Driven Evolution*. OUP Oxford.

Neupert W (2015). A perspective on transport of proteins into mitochondria: a myriad of open questions. *J. Mol. Biol.* **427**: 1135–1158.

Niederman RA (2016). Development and dynamics of the photosynthetic apparatus in purple phototrophic bacteria. *BBA - Bioenergetics* **1857**: 232–246.

Norbury C and Nurse P (1992). Animal cell cycles and their control. *Annu. Rev. Biochem.* **61**: 441–468.

Novak B, Csikasz-Nagy A, Gyorffy B, Nasmyth K, and Tyson JJ (1998). Model scenarios for evolution of the eukaryotic cell cycle. *Phil. Trans. R. Soc. Lond. B.* **353**: 2063–2076.

Olmos Y, Hodgson L, Mantell J, Verkade P, and Carlton JG (2015). ESCRT-III controls nuclear envelope reformation. *Nature* **522**: 236–239.

Otto SP (2009). The evolutionary enigma of sex. *Am. Nat.* **174**: S1–S14.

Parfrey LW, Lahr DJG, Knoll AH, and Katz LA (2011). Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proc. Natl. Acad. Sci. U.S.A.* **108**: 13624–13629.

Peeters E, Rosalie P C Driessen, Werner F, and Dame RT (2015). The interplay between nucleoid organization and transcription in archaeal genomes. *Nat. Rev. Microbiol.* **13**: 333–341.

Pereira SFF, Goss L, and Dworkin J (2011). Eukaryote-like serine/threonine kinases and phosphatases in bacteria. *Microbiol. Mol. Biol. Rev.* **75**: 192–212.

Perrot VV, Richerd SS, and Valéro MM (1991). Transition from haploidy to diploidy. *Nature* **351**: 315–317.

Perry AJ, Hulett JM, Likic VA, Lithgow T, and Gooley PR (2006). Convergent evolution of receptors for protein import into mitochondria. *Curr. Biol.* **16**: 221–229.

Pfanner N and Geissler A (2001). Versatility of the mitochondrial protein import machinery. *Nat. Rev. Mol. Cell. Biol.* **2**: 339–349.

Pinevich AV (1997). Intracytoplasmic membrane structures in bacteria. *Endocyt. Cell Res.* **12**: 9–40.

Pollard TD (2003). The cytoskeleton, cellular motility and the reductionist agenda. *Nature* **422**: 741–745.

Poole AM and Gribaldo S (2014). Eukaryotic origins: how and when was the mitochondrion acquired? *Cold Spring Harb. Perspect. Biol.* **6**: a015990–a015990.

Portier P (1918). *Les Symbiotes*. Masson: Paris, France.

Portillo MC, Leff JW, Lauber CL, and Fierer N (2013). Cell size distributions of soil bacterial and archaeal taxa. *Appl. Environ. Microbiol.* **79**: 7610–7617.

Rada P, Makki AR, Zimorski V, Garg S, Hampl V, Hrdý I, Gould SB, and Tachezy J (2015). N-terminal presequence-independent import of phosphofructokinase into hydrogenosomes of *Trichomonas vaginalis*. *Eukaryot. Cell* **14**: 1264–1275.

Raikov IB (1994). The diversity of forms of mitosis in protozoa - a comparative review. *Eur. J. Protistol.* **30**: 253–269.

Ramesh MA, Malik SB, and Logsdon JM Jr (2005). A phylogenomic inventory of meiotic genes. *Curr. Biol.* **15**: 185–191.

Raymann K, Brochier-Armanet C, and 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**: 6670–6675.

Reeve JN, Sandman K, and Daniels CJ (1997). Archaeal histones, nucleosomes, and transcription initiation. *Cell* **89**: 999–1002.

Remmert M, Biegert A, Linke D, and Lupas AN (2010). Evolution of outer membrane β-barrels from an ancestral $\beta\beta$ hairpin. *Mol. Biol. Evol.* **27**: 1348–1358.

Reyes-Lamothe R, Nicolas E, and Sherratt DJ (2012). Chromosome replication and segregation in bacteria. *Annu. Rev. Genet.* **46**: 121–143.

Reynders E, Foulquier F, Annaert W, and Matthijs G (2011). How Golgi glycosylation meets and needs trafficking: the case of the COG complex. *Glycobiology* **21**: 853–863.

Richardson LGL, Paila YD, Siman SR, Chen Y, Smith MD, and Schnell DJ (2014). Targeting and assembly of components of the TOC protein import complex at the chloroplast outer envelope membrane. *Front. Plant Sci.* **5**: 1513.

Roger AJ, Clark CG, and Ford Doolittle W (1996). A possible mitochondrial gene in the early-branching amitochondriate protist *Trichomonas vaginalis*. *Proc. Natl. Acad. Sci. U.S.A.* **93**: 14618–14622.

Sagan L (1967). On the origin of mitosing cells. *J. Theor. Biol.* **14**: 255–274.

Santarella-Mellwig R, Pruggnaller S, Roos N, Mattaj IW, and Devos DP (2013). Three-dimensional reconstruction of bacteria with a complex endomembrane system. *PLoS Biol.* **11**: –e1001565.

Sapp J (2005). The prokaryote-eukaryote dichotomy: meanings and mythology. *Microbiol. Mol. Biol. Rev.* **69**: 292–305.

Saw JH, Spang A, Zaremba-Niedzwiedzka K, Juzokaite L, Dodsworth JA, Murugapiran SK, Colman DR, Takacs-Vesbach C, Hedlund BP, Guy L, and Ettema TJG (2015). Exploring microbial dark matter to resolve the deep archaeal ancestry of eukaryotes. *Phil. Trans. R. Soc. Lond. B.* **370**: 20140328–630.

Sazer S, Lynch M, and Needleman D (2014). Deciphering the evolutionary history of open and closed mitosis. *Curr. Biol.* **24**: R1099–103.

Schatz G and Mason TL (1974). The biosynthesis of mitochondrial proteins. *Annu. Rev. Biochem.* **43**: 51–87.

Schlacht A, Herman EK, Klute MJ, Field MC, and Dacks JB (2014). Missing pieces of an ancient puzzle: evolution of the eukaryotic membrane-trafficking system. *Cold Spring Harb. Perspect. Biol.* **6**: a016048–a016048.

Schleiff EE and Becker TT (2010). Common ground for protein translocation: access control for mitochondria and chloroplasts. *Nat. Rev. Mol. Cell. Biol.* **12**: 48–59.

Schleiff E, Maier UG, and Becker T (2011). Omp85 in eukaryotic systems: one protein family with distinct functions. *Biol. Chem.* **392**: 21–27.

Shi LX and Theg SM (2013). Energetic cost of protein import across the envelope membranes of chloroplasts. *Proc. Natl. Acad. Sci. U.S.A.* **110**: 930–935.

Shinohara A, Ogawa H, and Ogawa T (1992). Rad51 protein involved in repair and recombination in *S. cerevisiae* is a RecA-like protein. *Cell* **69**: 457–470.

Slesarev AI, Belova GI, Kozyavkin SA, and Lake JA (1998). Evidence for an early prokaryotic origin of histones H2A and H4 prior to the emergence of eukaryotes. *Nucleic Acids Res.* **26**: 427–430.

Sogin ML (1989). Evolution of eukaryotic microorganisms and their small subunit ribosomal RNAs. *Am. Zool.* **29**: 487–499.

Solari AJ (2002). Primitive forms of meiosis: the possible evolution of meiosis. *Biocell* **26**: 1–13.

Soll J and Schleiff E (2004). Protein import into chloroplasts. *Nat. Rev. Mol. Cell. Biol.* **5**: 198–208.

Soubannier V, McLelland GL, Zunino R, Braschi E, Rippstein P, Fon EA, and McBride HM (2012). A Vesicular Transport Pathway Shuttles Cargo from Mitochondria to Lysosomes. *Curr. Biol.* **22**: 135–141.

Spang A, Saw JH, Jørgensen SL, Zaremba-Niedzwiedzka K, Martijn J, Lind AE, van Eijk R, Schleper C, Guy L, and Ettema TJG (2015). Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* **521**: 173–179.

Speijer D (2016). What can we infer about the origin of sex in early eukaryotes? *Phil. Trans. R. Soc. Lond. B.* **371**: 20150530.

Speijer D, Lukes J, and Eliáš M (2015). Sex is a ubiquitous, ancient, and inherent attribute of eukaryotic life. *Proc. Natl. Acad. Sci. U.S.A.* **112**: 8827–8834.

Stanier RY (1970). Some aspects of the biology of cells and their possible evolutionary significance. In HP Charles and B Knight (Eds.) *Symp Soc Gen Microbiol*, pp. 1–38. Cambridge University Press, Cambridge.

Stanier RY and van Niel CB (1962). The concept of a bacterium. *Archiv. Mikrobiol.* **42**: 17–35.

Stanley P (2011). Golgi glycosylation. *Cold Spring Harb. Perspect. Biol.* **3**: a005199–a005199.

Steppuhn J and Herrmann RG (1989). Domain structure of mitochondrial and chloroplast targeting peptides. *Eur. J. Biochem.* **180**: 535–545.

Sugiura A, McLelland GL, Fon EA, and McBride HM (2014). A new pathway for mitochondrial quality control: mitochondrial-derived vesicles. - PubMed - NCBI. *EMBO J.* **33**: 2142–2156.

Takeuchi N, Kaneko K, and Koonin EV (2014). Horizontal gene transfer can rescue prokaryotes from Muller's ratchet: benefit of DNA from dead cells and population subdivision. *G3* **4**: 325–339.

Thangudu RR, Manoharan M, Srinivasan N, Cadet F, Sowdhamini R, and Offmann B (2008). Analysis on conservation of disulphide bonds and their structural features in homologous protein domain families. *BMC Struct.Biol.* **8**: 55.

Tong J, Doležal P, Selkrig J, Crawford S, Simpson AGB, Noinaj N, Buchanan SK, Gabriel K, and Lithgow T (2011). Ancestral and derived protein import pathways in the mitochondrion of *Reclinomonas americana. Mol. Biol. Evol.* **28**: 1581–1591.

Toro E and Shapiro L (2010). Bacterial chromosome organization and segregation. *Cold Spring Harb. Perspect. Biol.* **2**: a000349.

Villeneuve AM and Hillers KJ (2001). Whence meiosis? *Cell* **106**: 647–650.

Voeltz GK, Rolls MM, and Rapoport TA (2002). Structural organization of the endoplasmic reticulum. *EMBO Rep.* **3**: 944–950.

Voulhoux R and Tommassen J (2004). Omp85, an evolutionarily conserved bacterial protein involved in outer-membrane-protein assembly. *Res Microbiol* **155**: 129–135.

Wallin IE (1925). On the nature of mitochondria. IX. Demonstration of the bacterial nature of mitochondria. *American Journal of Anatomy* **36**: 131–149.

Walter P, Gilmore R, and Blobel G (1984). Protein translocation across the endoplasmic reticulum. *Cell* **38**: 5–8.

Watson P and Stephens DJ (2005). ER-to-Golgi transport: Form and formation of vesicular and tubular carriers. *BBA - Mol. Cell Res.* **1744**: 304–315.

Whiddon K and Konopka MC (2015). Rapid assessment of intracytoplasmic membranes in bacteria by fluorescence microscopy. *Biophys. J.* **108**: 78a.

Whitfeld PR and Bottomley W (1983). Organization and structure of Chloroplast genes . *Ann. Rev. Plant Physiol.* **34**: 279–310.

Wilkins AS and Holliday R (2008). The evolution of meiosis from mitosis. *Genetics* **181**: 3–12.

Williams GC (1975). *Sex and evolution*. Princeton University Press, Princeton.

Williams TA, Foster PG, Cox CJ, and Embley TM (2013). An archaeal origin of eukaryotes supports only two primary domains of life. *Nature* **504**: 231–236.

Williams TA, Foster PG, Nye TMW, Cox CJ, and Embley TM (2012). A congruent phylogenomic signal places eukaryotes within the Archaea. *Proc. R. Soc. B* **279**: 4870–4879.

Wilson C, Venditti R, Rega LR, Colanzi A, D'Angelo G, and De Matteis MA (2011). The Golgi apparatus: an organelle with multiple complex functions. *Biochem. J.* **433**: 1–9.

Woese CR and Fox GE (1977). Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proc. Natl. Acad. Sci. U.S.A.* **74**: 5088–5090.

Wu LG, Hamid E, Shin W, and Chiang HC (2014). Exocytosis and endocytosis: modes, functions, and coupling mechanisms. *Annu. Rev. Physiol.* **76**: 301–331.

Yang D, Oyaizu Y, Oyaizu H, Olsen GJ, and Woese CR (1985). Mitochondrial origins. *Proc. Natl. Acad. Sci. U.S.A.* **82**: 4443–4447.

Yutin N, Wolf MY, Wolf YI, and Koonin EV (2009). The origins of phagocytosis and eukaryogenesis. *Biol. Direct* **4**: 9.

Zeth K and Thein M (2010). Porins in prokaryotes and eukaryotes: common themes and variations. *Biochem. J.* **431**: 13–22.

Zimorski V, Major P, Hoffmann K, Brás XP, Martin WF, and Gould SB (2013). The N-terminal sequences of four major hydrogenosomal proteins are not essential for import into hydrogenosomes of *Trichomonas vaginalis*. *J. Eukaryot. Microbiol.* **60**: 89–97.

Acknowldgements

When I met Bill for the first time in the grounds of the Nation Center for Biological Sciences in Bengaluru, India I had no clue as to the amazing journey that would entail. It took all of eight minutes from “Excuse me professor” to a question that went on to re-defined my life. There are a million things to thank for but most important is the freedom he gave me. The freedom to explore avenues of science most challenging and guiding my ever poor and confused intellect. At a time when I had lost all hope from science Bill managed to rekindle the scientist within me for which I will remain ever grateful and strive to repay that kindness.

Being trained in the laboratory when Bill asked would I like an informatics project or a lab project my choice led me to Sven’s doorstep. To someone who taught me the difference between thinking up an experiment to do and the difficulties that go into realizing it in practice a mere thank you is not enough. It has always been exciting talking about cell biology and its evolution with Sven and I hope that continues.

Finding a house in Germany is just as hard as finding friends for a culturally shocked timid guy from India. Fortunately I found both in Shiju. From arranging an apartment to a working internet connection to teaching me programming, the list goes on and will continue to do so. Shiju made me feel welcome in Germany and the Lab to which I will always be thankful

No one that I have ever met in the lab has been more indispensable than Verena. From locating machines and manuals from decades ago to making sure my haphazard way of working in the lab does not burn the building down I am indebted to her. On the other side of things Mayo is just as indispensable when it comes to my laptop (surviving without which is impossible). I have troubled him on so many occasions on so many things. I am thankful for his patience in dealing with a noob like me.

Gary, Houda, Re-young and Kathrin were terrific friends when they were in the lab and constantly kept pushing me and providing a level of moral support that I continue to miss. Filipa is irreplaceable both as an office-mate and as a friend. Jan and Sophie, your sense of dedication towards science and science fiction is admirable, while it was indeed sad to see you leave I hope we can settle debts in the future as we are not done yet! Cessa, well Thank you for everything past and everything yet to come. Lastly I would like to thank Doris, for making the dreaded German bureaucracy non-existent. Molevol and everyone associated with it will always remain special to me, it is a place where the proverbial boy became a man.

I would also like to thank Esther for tolerating me and all my idiosyncrasies and promising to continue to do so. Finally, I would like to thank my parents and family for being patient and that unwavering belief in me.