Current Biology
Review

Endosymbiosis and Eukaryotic Cell Evolution

John M. Archibald
Department of Biochemistry and Molecular Biology, Dalhousie University, Sir Charles Tupper Medical Building, 5850 College Street, Halifax, Nova Scotia, B3H 4R2, Canada, and Program in Integrated Microbial Biodiversity, Canadian Institute for Advanced Research, Toronto, Ontario, M5G 1Z8, Canada
Correspondence: John.Archibald@Dal.Ca
http://dx.doi.org/10.1016/j.cub.2015.07.055

Understanding the evolution of eukaryotic cellular complexity is one of the grand challenges of modern biology. It has now been firmly established that mitochondria and plastids, the classical membrane-bound organelles of eukaryotic cells, evolved from bacteria by endosymbiosis. In the case of mitochondria, evidence points very clearly to an endosymbiont of α-proteobacterial ancestry. The precise nature of the host cell that partnered with this endosymbiont is, however, very much an open question. And while the host for the cyanobacterial progenitor of the plastid was undoubtedly a fully-fledged eukaryote, how — and how often — plastids moved from one eukaryote to another during algal diversification is vigorously debated. In this article I frame modern views on endosymbiotic theory in a historical context, highlighting the transformative role DNA sequencing played in solving early problems in eukaryotic cell evolution, and posing key unanswered questions emerging from the age of comparative genomics.

Introduction

There are two kinds of cellular life forms on Earth — prokaryotes and eukaryotes. How the latter evolved from the former is a mystery that has intrigued biologists for the better part of a century. In the early 1960s, Stanier, Douderoff, and Adelberg referred to the prokaryote–eukaryote divide as "the greatest single evolutionary discontinuity to be found in the present-day world" [1]. It is a sentiment with which many researchers today would agree — many, but not all. The terms 'prokaryote' and 'eukaryote', first introduced in 1938 by the little-known Frenchman Edouard Chatton [2,3], have come in and out of fashion as knowledge of the microbial biosphere has improved.

Together with the advent of ribosomal RNA (rRNA)-based molecular systematics, the 'discovery' of the archaea in the 1970s by Carl Woese and colleagues [4] led to widespread acceptance of the so-called 'three-domains' view of life in which there are two prokaryotic domains, Bacteria and Archaea, each as distinct from the other as they are from the domain Eukarya [5]. In the 1980s and 90s, phylogenetic analyses of anciently duplicated protein genes suggested that archaea and eukaryotes share a more recent common ancestor than either does with bacteria (e.g., [6–8] but see [9]). With the root of the tree of life placed between bacteria on one side and archaea-plus-eukaryotes on the other, the evolutionary significance of the prokaryote–eukaryote dichotomy became the subject of intense debate [5,10,11]. Norman Pace has argued that the word 'prokaryote' is obsolete, an impediment towards understanding the true nature of the living world [12].

But while genome sequence data have shown that the archaea are indeed a distinct biological entity [13], the relationships between bacteria, archaea and eukaryotes are not as clear-cut as once thought. Twenty years of comparative genomics have led to the realization that prokaryotic and eukaryotic genomes are evolutionary mosaics; the phylogenetic signals contained within them are complex, so complex that using genomic data to test competing hypotheses of cell evolution is extremely difficult. Nevertheless, recent phylogenetic analyses of ubiquitous and presumed-to-be vertically inherited 'core' genes suggest that there might in fact be only two primary domains of life, Bacteria and Archaea, with the eukaryotes having emerged from within the latter (e.g., [14,15]). This scenario is a revival of James Lake's 'eocyte hypothesis', first proposed in the 1980s on the basis of ribosome morphology [16].

Taxonomic issues aside, the eukaryotic grade of cellular organization must surely have arisen from some sort of 'simpler' prokaryotic one, and the question of how eukaryotes first evolved is as relevant today as when Stanier and colleagues first underscored the problem more than 50 years ago. And there is one fact upon which everyone agrees: endosymbiosis, the bringing together of distinct cells, one inside the other, has been an important factor in eukaryotic evolution. Precisely how, how often, and why are the outstanding questions of the day. Here I explore past and present views on the role of endosymbiosis in the evolution of eukaryotic cells and their organelles, largely from the perspective of molecular biology and genomics. I begin with a brief sketch of endosymbiosis research in the pre-molecular era, as it is from here that the conceptual framework for testing the endosymbiont hypothesis for the origins of mitochondria and plastids emerged.

Serial Endosymbiosis Theory

The roots of modern endosymbiotic theory run deep and tangled. It was founded on the concept of *symbiosis* — from the Greek 'together' and 'living' — which emerged largely from the study of lichens. In 1867, the Swiss biologist Simon Schwendener put forth the heretical notion that lichens were composite beings comprised of a fungus and an alga [17]. Lichens were the 'problem child' of nineteenth century systematists, 'organisms' that by nature did not fit into the classification schemes of the day [18]. In 1879 the German Anton de Bary defined symbiosis as "the living together of unlike organisms" [19], and bolstered by the pioneering work of Poland's Franz

Kamienski and Germany's Albert Frank in the late 1800s on mycorrhizal fungi and their intimate relationship with the roots of plants [20,21], symbiosis gradually became recognized as a legitimate, albeit perplexing, biological phenomenon.

The Russian botanist Constantin Mereschkowsky played an important role in developing the concept of symbiogenesis, "the origin of organisms through the combination and unification of two or many beings entering into symbiosis" [22]. Among other things, Mereschkowsky was an authority on the subcellular architecture of diatom algae and their 'chromatophores' (plastids), and he was well versed in lichen biology (see [23] for review). In a 1905 publication, translated from German into English by Martin and Kowallik in 1999 [24], Mereschkowsky put forth an explicit and compelling case for an endosymbiotic origin for plastids. It was the first of its kind.

In his landmark 1905 paper, Mereschkowsky reviewed the present state of knowledge with respect to symbiosis, including lichens and certain amoebae found to have green algae living inside them. He also documented what he saw as the "great and obvious" similarities between plastids and free-living 'cyanophyceae' — cyanobacteria. And he emphasized the "continuity of chromatophores": as demonstrated by the German Andreas Schimper and the Swiss Carl Wilhelm von Nägeli, among others, plastids are not made *de novo* by the cell, but rather come from preexisting organelles by division. This fact, Mereschkowsky argued, strongly suggested that plastids had once been free-living organisms, as had been alluded to in a footnote by Schimper in 1883 [25].

Mereschkowsky is widely hailed as the 'founding father' of endosymbiotic theory. It is nevertheless interesting to note that he firmly dismissed the possibility that mitochondria might have evolved by endosymbiosis, and it is important to recognize that he is one of a number of researchers who helped lay the foundation for endosymbiotic thinking long before the advent of molecular sequencing. These include Schimper, Mereschkowsky's fellow Russian 'symbiogeneticists' Andrey Famintsyn and Boris Kozo-Polyansky, the Frenchman Paul Portier, and the American Ivan Wallin (see [18,26] and references therein). With the benefit of hindsight, it is fascinating to explore their individual contributions and consider the extent to which they fit with the 'facts' of modern biology. One constant is the modest impact their research appears to have had at the time it was carried out, Mereschkowsky's efforts included. For much of the 18th century symbiosis was largely ignored as a potential source of evolutionary innovation. The burgeoning disciplines of genetics and cell biology were simply not equipped to properly account for it.

The English-speaking world was introduced to the concept of endosymbiosis largely through the works of the American biologist Lynn Margulis [18,26]. By her own telling, Margulis was heavily influenced by her cytology Professor Hans Ris and her graduate supervisor Walter Plaut; both knew the history of symbiosis research in Germany and Russia, scorned though it was [27]. Her earliest scientific contributions were modest. Studying amoebae and algae in the 1950s and 60s, she added to the body of knowledge suggesting the presence of extranuclear DNA in eukaryotes [28,29]. Her mentors Ris and Plaut published microscopic evidence for DNA in plastids in 1962 [30], and complementary data for mitochondria came a year later from Margit and Sylvan Nass [31].

Drawing on information taken from diverse areas of biology and geoscience, including genetics, bacteriology, cell biology, ecology, and paleontology, Margulis formulated a bold, broad-sweeping hypothesis for the evolution of eukaryotic life. Symbiosis featured heavily. In her classic 1967 paper, *On the origin of mitosing cells* (published as Lynn Sagan), Margulis proposed that "…mitochondria, the (9+2) basal bodies of the flagella, and the photosynthetic plastids can all be considered to have derived from free-living cells, and the eukaryotic cell is the result of the evolution of ancient symbioses" [32]. Her 1970 book entitled *Origin of Eukaryotic Cells* [33] brought symbiosis in general, and endosymbiosis in particular, to the scientific mainstream. Unlike her predecessors, Margulis was in the right place at the right time — her ideas were center stage when the tools of molecular biology were coming online and could be used to test them [26]. Margulis was nevertheless not the sole contemporary advocate for the importance of endosymbiosis in eukaryotic cell evolution. For example, the Norwegian Jostein Goksøyr [34] and the American Peter Raven [35] published complementary proposals in 1967 and 1970, respectively.

If not endosymbiosis, what? A more conservative view, endorsed early on by Richard Klein and Arthur Cronquist [36] and later by Tom Cavalier Smith [37], was that mitochondria and plastids had evolved from within the confines of a photosynthetic eukaryotic cell, which had itself evolved in a vertical fashion from a cyanobacterium-like prokaryote. Margulis referred to this scenario as the 'botanical myth' [38]. She argued that it was more parsimonious to assume that oxygenic photosynthesis was not an ancestral eukaryotic feature but rather had evolved more recently via the endosymbiotic uptake of a cyanobacterium by a heterotrophic eukaryote.

Various other non-endosymbiotic models for organelle evolution were also on the table, and not unreasonably so. The views of Lawrence Bogorad [39], Rudolph Raff and Henry Mahler [40], and Thomas Uzzell and Christina Spolsky [41] were similar to those of Klein *et al.* in emphasizing intracellular compartmentalization as the mechanism by which membrane-bound mitochondria and plastids had evolved; they differed in their assumptions about the nature of the prokaryotic stock that spawned the earliest eukaryotic cells. All such models were grounded in the belief that endosymbiosis was unnecessarily radical: "In our opinion [Raff and Mahler] there is no *a priori* reason why the eucaryotic cell, which has proved capable of remarkable evolutionary innovations, should have originated as a collage of procaryotic cells and parts of cells rather than having evolved in a more direct manner from a particularly advanced type of procaryotic cell" [40]. In 1974, Max Taylor framed these competing hypotheses as the 'autogenous' and 'xenogenous' (or foreign) models of eukaryotic evolution, the latter forming the foundation of his 'Serial Endosymbiosis Theory' [26,42].

The first nucleic acid sequences brought to bear on the endosymbiont hypothesis were obtained in the mid-1970s using the laborious 'RNA cataloguing' technique developed by Carl Woese, the same approach that led to the discovery of the archaea [4]. As recounted in more detail elsewhere [26,43], Linda Bonen, Ford Doolittle, Woese and colleagues obtained snippets of rRNA sequences from algal plastids and cyanobacteria and demonstrated a strong evolutionary link between them [44–46]. Mitochondrial rRNA fragments were shown to be demonstrably

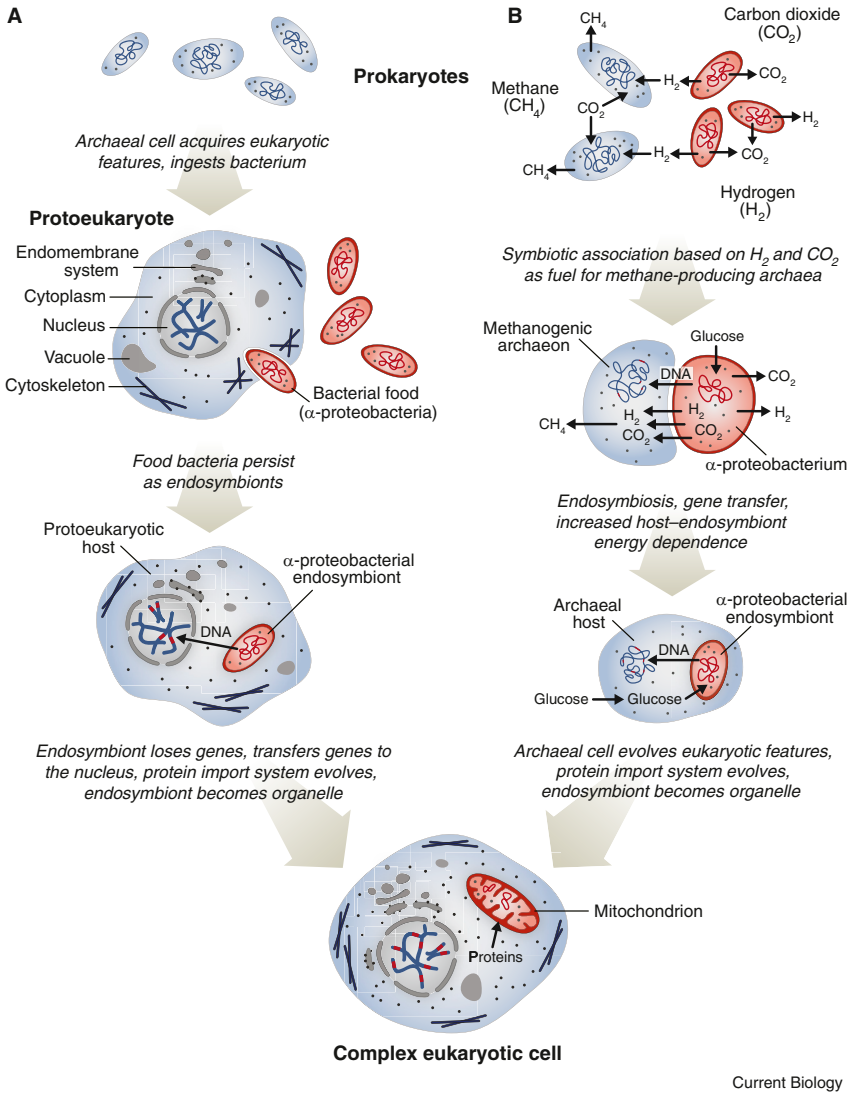

Figure 1. Two competing evolutionary scenarios for the origin of eukaryotic cells and their mitochondria.

(A) The traditional view posits that the bulk of eukaryotic cellular complexity arose in a stepwise fashion prior to the endosymbiotic uptake (by phagocytosis) of the α-proteobacterium that became the mitochondrion. (B) The hydrogen hypothesis [70] invokes a metabolic symbiosis between methane-producing archaea and α-proteobacteria. In this scenario eukaryotic cellular complexity arises after endosymbiosis. Both models involve extensive gene transfer from the α-proteobacterium to the archaeal host and the evolution of a system for targeting nucleus-encoded proteins to the endosymbiont-turned-organelle. Figure modified from [26].

The Eukaryotic Cell: From Whom and How?

Of course, knowing that mitochondria and plastids evolved by endosymbiosis did not solve the problem of eukaryotic evolution, far from it. Forty years have passed since the first organellar sequence data were analyzed and there is still no consensus as to how the complex suite of eukaryotic features — nucleus, endomembrane system, cytoskeleton, mitosis, and so on — evolved from a prokaryotic cell. This is not for lack of ideas or interest. The numerous models for the evolution of eukaryotes proposed before, during and after the molecular sequencing revolution have been explored elsewhere (e.g., [51–53]) and summarized recently and authoritatively by Martin *et al.* [54]. I will not attempt that feat here, except to say that they vary considerably with respect to the partner cells involved and the specific aspects of eukaryotic cell biology and physiology that they seek to explain. I will instead compare and contrast two fundamentally different sorts of models: those holding that the salient features of present-day eukaryotes evolved prior to the endosymbiotic origin of the mitochondrion, and those in which both the host and endosymbiont were prokaryotic cells, i.e., the mitochondrion evolved concomitantly with the eukaryotic cell itself (Figure 1). I will refer to these as the 'mitochondrion-late' and 'mitochondrion-early' scenarios.

prokaryotic shortly thereafter [47], although at the time they were not obviously affiliated with a specific bacterial lineage (a phylogenetic connection between mitochondria and α proteobacteria was shown in 1985 by Woese's group [48]). Margulis' hypothesis that eukaryotic flagella (or undulipodia as she eventually preferred to call them) were of endosymbiotic origin [32,33,38] could not be directly tested; DNA was not, and has never been, found associated with these motility organelles.

Proof of the endosymbiont hypothesis for the origins of mitochondria and plastids was a consilience. I have emphasized the molecular sequencing revolution, which revealed that mitochondrial and plastid rRNA sequences were more similar to those of bacteria — and each to *different* bacteria — than they were to nuclear rRNAs. But it was consideration of a wealth of evidence on the biochemistry and molecular biology of organelles collected by researchers around the world that ultimately led to the death of the autogenous model. By the mid-1980s it was clear that endosymbiosis was the only reasonable explanation for the data in hand [49,50].

Cavalier-Smith's Archezoa hypothesis is perhaps the best-known 'mitochondrion-late' scenario, having served as the main framework for research on eukaryotic cell evolution for much of the 1980s and 90s. Simply put, the Archezoa were 'eukaryotes without mitochondria' [55,56]. Cavalier-Smith argued that certain extant lineages of anaerobic single-celled eukaryotes were 'living fossils', direct descendants of organisms that had diverged from the main eukaryotic line before the endosymbiotic origin of mitochondria but after the evolution of the nucleus and cytoskeleton. Indeed, Cavalier-Smith considered the presence of a cytoskeleton in the host cell to be a prerequisite for

the phagocytotic uptake of the bacterial endosymbiont that became the mitochondrion, as did Christian de Duve and Roger Stanier many years prior [57,58] (Figure 1A).

Early molecular phylogenetic analyses that included archezoan sequences generated much excitement in placing amitochondriate lineages such as diplomonads (e.g., *Giardia*), parabasalids (e.g., *Trichomonas*) and microsporidians (e.g., *Encephalitozoon*) as the deepest offshoots of the eukaryotic tree, far below the branches leading to the more familiar plants, animals and fungi (e.g., [59–61]). Such topologies were consistent with a 'late' mitochondrial acquisition, fueling hopes that the study of archezoan taxa would provide important clues to the genesis of eukaryote-specific molecular and cell biological features.

By the turn of the century, however, the Archezoa hypothesis was in dire straights, for two main reasons. First, it gradually became clear that the basal placement of archezoan taxa in molecular trees was a 'long branch attraction' artifact: with the sparse taxonomic sampling and simplistic tree-building algorithms available at the time, the long, uninterrupted branches leading to the archezoan species were artificially attracted to one another and to the base of the eukaryotic tree by the prokaryotic outgroup sequences (e.g., [62–64]). Confidence in the idea that anaerobic archezoa were more ancient than aerobic, mitochondrion-containing lineages was eroded.

The second (and final) nail in the archezoan coffin was much more damaging: the Archezoa no longer appeared to exist. With much hard work and ingenuity, researchers showed that Cavalier-Smith's archezoan protists have double membrane-bound organelles that do not look like mitochondria but nevertheless very clearly *evolved from* mitochondria. These mitochondrion-related organelles (MROs) include the tiny 'mitosomes' of diplomonads and the hydrogen-producing 'hydrogenosomes' of parabasalids and anaerobic ciliates (see [65,66] and references therein). While some MROs have a genome and produce ATP, others do not. Still others possess genomic and biochemical features in common with both mitochondria and hydrogenosomes, blurring the distinction between them [67]. An important process common to *all* mitochondria and MROs is the synthesis of iron–sulfur clusters, and while the full range of metabolisms exhibited by MROs in nature is still unknown, it now appears that textbook aerobic mitochondria such as our own represent one extreme of an evolutionary continuum [68,69].

The realization that all known eukaryotes evolved from a mitochondrion-bearing ancestor does not exclude the possibility that Archezoa existed at some point in time but subsequently went extinct (or have yet to be discovered); many or most of the key features of the eukaryotic cell could still have arisen before the mitochondrion. But what it has done is to inspire fresh perspectives on eukaryogenesis, free of the assumption that the eukaryotic cell must have become complex prior to the evolution of the mitochondrion. None are more provocative than the hydrogen hypothesis of William Martin and Miklós Müller [70], the classic 'mitochondrion-early' scenario.

The hydrogen hypothesis (Figure 1B) posits that there never was an amitochondriate phase in eukaryotic evolution: the mitochondrion and the eukaryotic cell evolved in concert with one another, the result of a symbiosis between two prokaryotes. The host was an autotrophic methanogenic archaeon and the symbiont was a metabolically sophisticated α-proteobacterium capable of living both anaerobically and aerobically. In the absence of oxygen, the α-proteobacterium produced hydrogen gas and carbon dioxide as waste, which served to fuel the anaerobic metabolism of the methanogen with which it was closely associated. The α-proteobacterium eventually came to reside within the methanogen (how this happened is not specified but examples of prokaryotes living within other prokaryotes are known (e.g., [71])) and the transfer of genes from endosymbiont to host cemented the relationship, providing the methanogen with the means to import organic molecules from the environment and carry out glycolysis [70].

The hydrogen hypothesis is pleasingly explicit from a metabolic perspective: it provides a selective explanation for why the host needed its endosymbiont (hydrogen) and why endosymbiotic gene transfer was an essential part of the equation (it converted the autotrophic host into a heterotroph). And unlike most phagotrophy-based models for eukaryotic evolution, it points to a specific archaeal lineage — the methanogens or some other hydrogen-dependent archaeon — as having given rise to the nucleocytoplasmic component of today's eukaryotic cell. Unfortunately, although testable in principle, this aspect of the hypothesis is not easily confirmed or refuted. Nuclear genome sequences possess a complex mix of prokaryotic signatures, and making sense of them is complicated greatly by the fact that prokaryotic genomes are themselves highly mosaic due to horizontal gene transfer (HGT) [72,73].

A more controversial aspect of the hydrogen hypothesis is what it says about the ancestral state of mitochondrial biochemistry. While the raison d'être of the mitochondrion is traditionally assumed to revolve around oxygen and aerobic respiration (e.g., [33]), the hydrogen hypothesis holds that the proto-mitochondrion had — and kept hold of — everything it needed to give rise not only to aerobic mitochondria but to anaerobic, hydrogenosome-style organelles as well. If so, one prediction is that (when present) mitochondrial enzymes for anaerobic energy metabolism such as pyruvate–ferredoxin oxidoreductase and iron–iron hydrogenase should be of α-proteobacterial ancestry and specifically related to one another in evolutionary trees. Here too, however, the phylogenetic data are open to interpretation and complicated by genome mosaicism in prokaryotes. The evidence that prokaryote-to-eukaryote HGT has facilitated the 'recent' adaptation of protists to low oxygen environments is increasingly compelling (e.g., [74–76]). But it is nevertheless also true that genes for anaerobic enzymes are cropping up in a much broader range of eukaryotes than expected, and not just in highly derived, parasitic lineages [69,77]. Because we do not know how much data we are missing we cannot yet tell with certainty whether such genes were present in the eukaryotic common ancestor, acquired secondarily on multiple occasions by HGT, or a combination of the two.

Membranes and Missing Links

Regardless of what the primordial mitochondrion was or was not capable of, gene inventory studies paint a very clear picture of what the last eukaryotic common ancestor was like: it was a biochemically sophisticated, highly compartmentalized, gene-rich cell capable of phagocytosis, mitosis and sexual recombination

[52]. This leaves frustratingly open the critical question of how prokaryotes became eukaryotes in the first place and, indeed, why eukaryote-style complexity appears to have evolved only once in the four billion year history of life.

Lane and Martin recently put forth a thought-provoking argument for why prokaryotes will forever remain 'simple' and 'small', and why eukaryotic cells and mitochondria evolved hand in hand: "Prokaryotic genome size is constrained by bioenergetics. The endosymbiosis that gave rise to mitochondria restructured the distribution of DNA in relation to bioenergetic membranes [the cellular membranes in prokaryotes, mitochondrial membranes in eukaryotes], permitting a remarkable 200,000-fold expansion in the number of genes expressed. This vast leap in genomic capacity was strictly dependent on mitochondrial power, and prerequisite to eukaryotic complexity…" [78]. The evidence marshaled in support of these bold statements is complex, and includes consideration of cell volume, genome size, ploidy levels, total energy budgets, and the 'energy available per gene' in diverse prokaryotic and eukaryotic cells. The devil is very much in the details, but if Lane and Martin are on the right track, then there can be no true intermediates on a road from prokaryotes to eukaryotes that does not somehow involve a mitochondrion; it was, they argue, an evolutionary journey that was impossible to take without the energy that the mitochondrion ultimately provided.

Will we ever know for sure? What researchers in the field have long been craving are genuine 'missing links', modern-day organisms whose biology has potential to bridge the prokaryote–eukaryote divide and tell us which among the competing eukaryogenesis scenarios best fit the data. Recent genomics-enabled explorations of diverse aquatic environments have rekindled enthusiasm for the possibility that such bridge organisms might actually exist.

Efforts have focused on the wealth of microbial diversity springing up around the Crenarchaeota, one of the two classical archaeal lines (the other being Euryarchaeota) and to which eukaryotes appear specifically allied in rigorous molecular phylogenies of slowly evolving, universally distributed proteins [14,15]. These new organisms include the so-called 'DPANN' taxa (e.g., Diapherotrites, Aenigmarchaeota and Nanoarchaeota) and members of the TACK superphylum (Thaumarchaeota, Aigarchaeota, Crenarchaeota, and Korarchaeota) (see [79] and references therein). Precisely how these lineages relate to one another and to crenarchaeotes and euryarchaeotes is still being worked out. But what is particularly interesting is that they possess a handful of 'eukaryotic signature proteins' (ESPs) [80] — molecular hallmarks of eukaryotic cells that would otherwise clearly distinguish them from bacteria and archaea. These include homologs of the cytoskeletal proteins actin and tubulin, components of the ubiquitin-based protein degradation system, and certain translation elongation factors (see [15] for review).

With all the excitement comes a need for caution. One concern is that the distribution of these ESPs amongst TACK and DPANN taxa is somewhat patchy: it is formally possible that at least some of the ESPs in prokaryotes do not represent ancestral features at all, but rather instances of 'recent' horizontal acquisition. On the face of it, eukaryote-to-prokaryote gene transfer would seem unlikely, but it should not be dismissed outright. It is, for example, the only reasonable explanation for the highly localized distribution of tubulin genes in members of the bacterial genus *Prosthecobacter* [81,82]. And one should always be cautious when interpreting phylogenetic trees built from alignments containing proteins with very low sequence identity (as is the case for eukaryotic tubulins and their prokaryotic FtsZ homologs [82]).

Another intriguing development on the 'missing link' front is the discovery of the 'Lokiarchaeota', an archaeal lineage found lurking in deep marine sediments [83,84]. The 'Loki' genomes encode most of the patchily distributed ESPs found in TACK and DPANN members, and a host of others as well. Of particular interest is a plethora of genes for small GTPases and genes for protein components of the ESCRT system (endosomal sorting complexes required for transport). Together these protein families speak to the possible existence of an endomembrane system in Lokiarchaeota. The presence of a cytoskeleton was also proposed on the basis of genes for bona fide actin homologs and gelsolin-like domain-containing proteins (the gelsolins are regulators of actin filament dynamics in eukaryotes) [84].

On the basis of phylogenomic analyses, the Lokiarchaeota are the closest relatives of eukaryotes presently known. But while their suite of ESPs certainly is impressive, the extent to which these organisms actually 'bridge the gap' [84] between prokaryotes and eukaryotes is debatable. The identification of prokaryotes with some of the molecular seeds of core eukaryotic cell biological processes is significant, but a lot hinges on what these proteins — the GTPases, the ESCRT components and so on — actually do in the cell. At present the main barrier to further progress is the lack of cultured representatives of Lokiarchaeota and their closest TACK relatives: these organisms are currently defined by genomes stitched together from metagenomic data. This obstacle will no doubt soon be overcome, paving the way for laboratory experimentation. Regardless of what the Lokiarchaeota ultimately tell us about the prokaryote-to-eukaryote transition, there will be much to learn from this exciting new phase of research on the biology of archaea.

Eukaryotic Photosynthesis: Origin and Spread

Like mitochondria, the evolution of plastids was a singularity — plastids evolved from endosymbiotic cyanobacteria only once in the history of eukaryotic life. There are, however, important differences in the evolutionary trajectories of the two organelles (and as discussed below, the situation is complicated by the existence of cyanobacterium-derived photosynthetic 'chromatophores' in a little-known amoeba named *Paulinella*). Whereas mitochondria and MROs have (or at least are assumed to have) evolved in a strictly vertical fashion since they first arose, plastid evolution has involved both vertical inheritance and horizontal spread. Determining the relative impact of these two distinct modes of organelle acquisition has proved daunting. Despite the availability of genome sequence data from diverse algal lineages, key 'when' and 'how' questions about the evolution of eukaryotic photosynthesis remain.

What we do know is that the host for the cyanobacterial progenitor of the plastid was a mitochondrion-containing eukaryote, a single-celled heterotrophic organism capable of ingesting prey by phagocytosis. The so-called 'primary' endosymbiotic origin of plastids appears to have taken place in a common ancestor shared by three eukaryotic lineages: red algae, glaucophyte

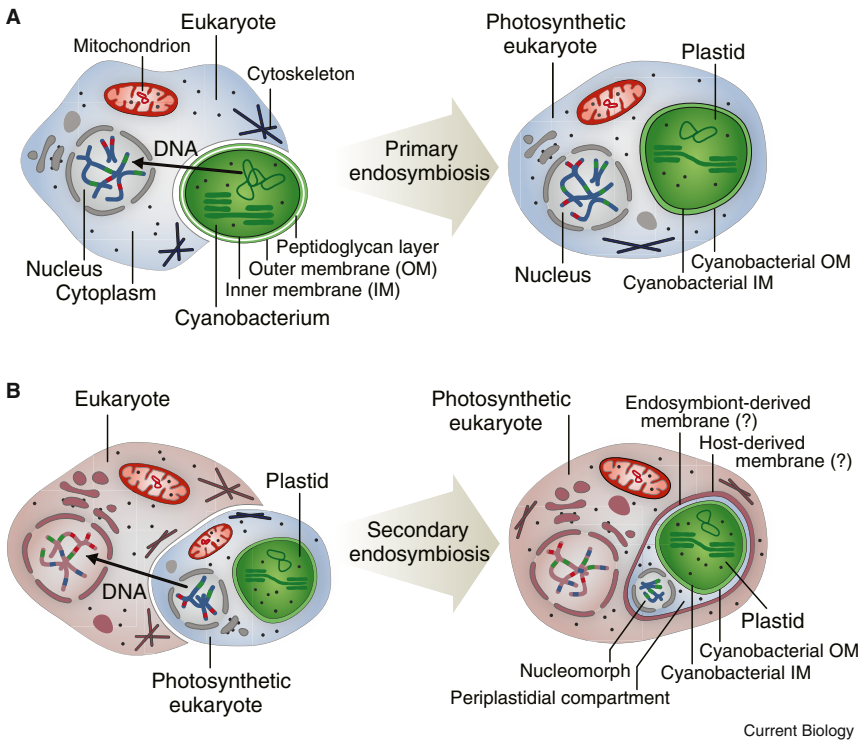

Figure 2. Endosymbiosis and plastid evolution.
(A) Primary endosymbiosis involves the uptake of a cyanobacterium by a non-photosynthetic eukaryote. The process involves endosymbiont to host DNA transfer and the evolution of a protein import apparatus. Primary plastids are surrounded by two membranes. The peptidoglycan layer present in the cyanobacterial progenitor of the plastid has been retained in glaucophyte algae but was lost in red and green algae. (B) Secondary endosymbiosis occurs when a primary plastid-bearing alga is ingested by a non-photosynthetic eukaryote. Genes of both prokaryotic and eukaryotic ancestry are transferred from the endosymbiont nucleus to the secondary host nucleus. In cryptophyte and chlorarachniophyte algae, the endosymbiont nucleus persists as a 'nucleomorph' residing in the periplastidial compartment (derived from the cytoplasm of the engulfed alga). Nucleomorphs have been lost in other secondary plastid-bearing algae. Secondary plastids are characterized by the presence of three or four membranes. Figure modified from [26].

algae, and green algae (it is from within the green line that land plants emerged) [85]. This was not always thought to be the case. Extant plastids are remarkably diverse in morphology and pigmentation, and in the 1970s proponents of endosymbiosis took this diversity as support for the notion that red and green algal plastids had evolved from different cyanobacteria (e.g., [35,86,87]). Forty years on, evidence for a singular origin of primary plastids is robust and multifaceted; it includes consideration of plastid genome size, architecture, and content, and rests heavily on the fact that all primary plastid-bearing organisms use the same multi-subunit translocon (the TIC–TOC complex) to direct nucleus-encoded proteins to the organelle (see [88,89] for review).

What were the metabolic factors that contributed to the evolution of plastids more than 1.5 billion years ago [90,91]? The most obvious benefit of a cyanobacterial endosymbiont to a heterotrophic eukaryote is oxygenic photosynthesis and the carbohydrate it provides. Another possibility is nitrogen fixation, an idea not without merit in light of gene content similarities between gene-rich heterocyst-forming cyanobacteria and photosynthetic eukaryotes [72]. The possible existence of a third player in the primary endosymbiotic origin of plastids is actively being debated, more specifically a (no longer present) *Chlamydia*-like pathogen proposed to have been impacting glycogen metabolism in the eukaryotic host at around the time the cyanobacterium came on the scene (see [92–94] and references therein for discussion). Whatever the reason(s), endosymbiotic gene transfer (EGT), together with the evolution of the TIC–TOC translocon, led to the establishment of the primary plastid and the very first autotrophic eukaryote.

Primary plastids reside in the cytosol of their eukaryotic hosts and are surrounded by two membranes (Figure 2A), both of which appear to be of cyanobacterial ancestry. However, the plastids of the model lab alga *Euglena* have three membranes, and various other algae have three- or four membrane-bound plastids [88]. These supernumerary membranes were an enigma until the 1970s and 80s when Sarah Gibbs, Max Taylor, Dennis Greenwood and colleagues recognized them for what they are: the calling card of 'secondary endosymbiosis', i.e., the spread of plastids from one eukaryote to another [42,95]. This process has given rise to some of the most ecologically significant algal lineages on Earth, including diatoms, haptophytes, and bloom-forming dinoflagellates [85].

The mechanics of evolving a 'complex' plastid are reasonably well understood. During the course of endosymbiont integration, a second round of EGT takes place, this time from the primary host nucleus to that of the secondary host (Figure 2B), and additional protein import mechanisms and machineries evolve on top of the pre-existing TIC–TOC-based system, facilitating the passage of nucleus-encoded proteins across the three or four plastid membranes [88]. Unlike primary plastids, secondary plastids reside within the lumen of the host's endomembrane system.

Various other aspects of complex plastid evolution are still enigmatic. The problem lies not in figuring out the nuts and bolts, but rather in determining how many times such organelles have arisen. Comparative genomic data strongly support the idea that two secondary endosymbioses involving capture of distinct green algae have occurred during the course of eukaryotic evolution, one leading to euglenids and the other to the chlorarachniophyte algae [96,97]. In contrast, all bets are off with respect to the origin(s) of red algal-derived complex plastids, which are exceptionally diverse and patchily distributed [85,96,98]. Cavalier-Smith's 'chromalveolate hypothesis' posits a single ancient capture of a red alga by an ancestor of 'chromist' algae and 'alveolates', followed by multiple plastid losses in organisms such as ciliates and oomycetes

[99]. This long-standing hypothesis has taken a beating in recent years but is nevertheless worthy of a fresh look from the perspective of protein import and membrane dynamics [100].

The alternative view is that subsequent to a single red algal secondary endosymbiosis, perhaps in an ancestor of cryptophyte algae, one or more cryptic 'tertiary' endosymbiotic events served to spread this red algal-derived plastid further afield (e.g., [101–104]). This model takes into account discrepancies between the host- and organelle-associated features of the organisms in question, including the relative strength of phylogenetic signals in their nuclear, mitochondrial and plastid genomes [105]. These additional endosymbioses are as yet hypothetical, but tertiary endosymbiosis is itself very real, having occurred on multiple occasions within dinoflagellate algae [106].

Which is 'easier', evolutionarily speaking, plastid gain by secondary/tertiary endosymbiosis or plastid loss? It depends on whom you ask and which characters one chooses to weigh most heavily. We do know that like mitochondria, plastids are hardwired into the cell's metabolic circuitry; they are the site of diverse biochemical processes that are not directly linked to photosynthesis, including amino acid, isoprenoid, and fatty acid biosynthesis [107]. For this reason, even when photosynthesis is lost (which is not uncommon) plastids usually stick around. Nevertheless, two clear examples of 'recent' plastid loss have been documented, in the apicomplexan *Cryptosporidium* [108,109] and, more recently, in the dinoflagellate *Hematodinium* [110]. It is important to note that both of these organisms are parasites, a lifestyle that is known to influence what a cell can and cannot get away with. Regardless, plastid loss clearly can happen and should not be discounted.

At the present time at least three secondary endosymbioses must be invoked to account for the diversity of extant complex plastids, two involving green algae and one a red alga. Additional higher-order endosymbioses on the red side are considered necessary by those attempting to reconcile comparative genomic and phylogenomic data with plastid-associated features (e.g., [102,104,111]). The nuclear genomes of complex algae are mosaics of genes whose individual histories have the potential to recount past endosymbiotic events (Figure 2B). The sobering reality, however, is that inferring the big picture of plastid evolution from gene trees has proven to be as problematic as the events themselves are ancient [112,113]. Some prefer to focus on cell biology and leave "gene tree conflicts deliberately unexplained" [100].

The Essence of an Organelle

Running parallel to the debate over anciently evolved plastids is vigorous discussion about the endosymbiont-to-organelle transition. What *is* an organelle and how do we know one when we see it? An unassuming freshwater amoeba by the name of *Paulinella chromatophora* has taken center stage. Discovered in 1894 by the German Robert Lauterborn, *Paulinella* was found to possess one or two blue–green pigmented bodies per cell. Lauterborn was struck by the similarities between these plastid-like structures and cyanobacteria, so much so that he is said to have mused "on the possible endosymbiotic origin of the chromatophores (i.e. plastids) without explicitly advancing this hypothesis (as did Mereschkovsky 10 years later)" [114].

Fast-forward more than 100 years and what we know about the chromatophores of *Paulinella* is this: they are obligate subcellular entities that divide synchronously with their host; they evolved from cyanobacteria but are not specifically related to canonical plastids; and they are organelles. How so? The chromatophore genome is ~1 megabase pairs in size, much larger than a plastid genome but significantly reduced relative to the *Synechococcus* cyanobacteria from which it evolved ~60–200 million years ago (see [115] and references therein for recent review). More than 30 genes of chromatophore origin have been found in the nuclear genome of the amoeba (most of which appear to be involved in photosynthesis), and experiments have shown that at least some of these nucleus-encoded proteins are targeted to the chromatophore post-translationally [116]. This last point is significant: according to Cavalier-Smith and John Lee, organelles have a protein import system, endosymbionts do not [117,118]. So the chromatophore is, technically speaking, a photosynthetic organelle, albeit one in the early stages of establishment.

Another interesting and even more recent endosymbiosis involves a particular group of cyanobacteria living inside rhopalodeacean diatoms. These 'spheroid bodies' are closely related to free-living, nitrogen-fixing species of the cyanobacterial genus *Cyanothece* (the endosymbiosis was apparently established less than 15 million years ago [119]). An intriguing aspect of this obligate relationship is that the spheroid bodies are no longer photosynthetic; their genomes are substantially reduced relative to *Cyanothece* and are in the process of jettisoning their photosynthesis genes [120,121]. This is perhaps not surprising, given that photosynthate is not something that would be particularly valuable to their plastid-bearing diatom hosts. What the spheroid bodies *do* provide is fixed nitrogen, a precious commodity for a eukaryotic cell. EGT and protein import have yet to be demonstrated in this system; whether the spheroid body is an organelle in the strictest sense [117] is thus unknown. But it is worth asking how much it matters. The spheroid body has evolved to carry out a specialized function for the diatom and has become metabolically dependent on it in return [121].

The distinction between endosymbiont and organelle has recently become even fuzzier with new data from the world of sap-feeding insects and their 'nutritional symbionts'. The bacteria are housed within specialized insect cells called bacteriocytes where they synthesize certain amino acids upon which their host is dependent for life. The symbionts themselves are famous for having pushed genome reduction to remarkable extremes — their genomes are sometimes substantially smaller than those of organelles such as the plastids of red algae [122]. Occasionally there are two symbionts whose minimal gene sets complement one another; only together are they capable of making up for the metabolic deficiencies of their host [123]. Interestingly, Nakabachi *et al.* [124] have shown that a bacterial-derived gene in the pea aphid genome encodes a protein that is targeted to its *Buchnera* symbiont (discussed in [125]). This discovery provides a small piece of the complex puzzle of how insect symbionts

are able to survive with such limited gene sets. It also illustrates the problem of definitions in biology.

Eukaryotic Evolution: Looking Back, Moving Forward

> "Evolutionary hypotheses are correctly interpreted as products of the data they set out to explain, but they are less often recognized as being heavily influenced by other factors. One of these is the history of preceding thought…"
>
> Keeling 2014 [126].

While research on mitochondria and plastids has served us well in framing the problem of how endosymbionts become organelles, it has the potential to constrain our thinking. Mitochondria and plastids each evolved only once during the history of life and we want to understand why and how. Using all of the modern tools available, continued exploration of the biochemical, genetic, and cell biological interactions underlying all manner of recently established endosymbiotic relationships will be valuable in and of itself. And it cannot fail to help us in our quest to understand how these classical organelles evolved.

More generally, it is worth considering the profound influence of endosymbiotic theory on the field of cellular evolution over the past 40 years. If endosymbiosis so readily explains the origins of mitochondria and plastids, might it also account for the evolution of other subcellular entities? Indeed, virtually every membrane-bound or membrane-associated component of the eukaryotic cell has at one time or another been suggested to be of endosymbiotic origin, flagellum, peroxisome, endoplasmic reticulum (ER), and nucleus included (see [54] and references therein). In the case of the peroxisome, for example, de Duve [127] proposed that this mysterious metabolic organelle evolved from an ancient bacterial endosymbiont that lost its genome and somehow ended up surrounded by a single membrane (unlike mitochondria and primary plastids, which have two). However, current data on protein import and organelle biogenesis are more consistent with the notion that peroxisomes are derivations of the ER (see [128,129] for discussion). So where did *it* come from?

The concept of nucleus as endosymbiont goes back to Mereschkowsky in the early 1900s [130] and various modern incarnations of this idea have appeared in the literature. Unfortunately, such hypotheses often suffer from a lack of cell biological realism (e.g., by ignoring the continuous nature of the inner and outer surfaces of the nuclear envelope) and it has not been possible to distinguish them from non-endosymbiotic models using genome sequence data [54]. In the absence of evidence to the contrary, we are left trying to explain why and how the nucleus and endomembrane system arose autogenously.

One intriguing possibility is that the nucleus evolved in response to an invasion of mobile genetic elements, more specifically the migration of group II introns from the bacterial ancestor of the mitochondrion to the archaeal host cell during organelle establishment [131]. Self-splicing group II introns are common in prokaryotic and organellar genomes, and are widely held to be the progenitors of spliceosomal introns [132]. According to Martin and Koonin, as the former type of intron evolved into the latter, co-transcriptional protein synthesis led to a situation in which introns were translated before they could be spliced out, a serious problem for the cell were it to involve more than a handful of genes. One solution would be to physically separate the process of transcription from translation: "The rapid, fortuitous spread of introns following the origin of mitochondria is adduced as the selective pressure that forged nucleus–cytosol compartmentalization" [131].

As for the origin of the nuclear envelope itself, it is typically thought to stem from invaginations of a prokaryotic host cell's plasma membrane; over time these invaginations became increasingly differentiated and connected to the nascent cytoskeleton (e.g., [128]). Baum and Baum have recently turned this sequence of events 'inside-out', suggesting that "…eukaryotes evolved from a prokaryotic cell with a single bounding membrane that extended extracellular protrusions that fused to give rise to the cytoplasm and endomembrane system" [133]. It was these protrusions that surrounded the bacterial ancestors of mitochondria. For his part, Cavalier-Smith has gone to great lengths to integrate cell biological and genomic information, arguing for the co-evolution of the endomembrane system, mitosis, and cytoskeleton [134].

At present there is no convincing evidence to support the notion that endosymbiosis played a role in the origin of any cytosolic compartments other than mitochondria and plastids. The challenge now is to come up with testable hypotheses for the autogenous origin of such compartments that are compatible with our understanding of cell biology and genome evolution. There is value in looking back at the history of cell evolution research. There is also a lot to be gained from attempting to divorce oneself from the past while looking forward at cell biological problems with modern data. And there is clearly much about the evolution of the eukaryotic cell that still needs to be worked out. In doing so, we should enjoy the view from both perspectives.

ACKNOWLEDGEMENTS

I thank Daniel Moog, Sven Gould, and Jan de Vries for helpful discussion, and William Martin, Klaus Kowallik and Jan Sapp for published translations. Apologies are made to those whose primary research could not be cited due to space constraints. Research on endosymbiosis in the Archibald Lab is supported by the Canadian Institutes of Health Research, the Natural Sciences and Engineering Research Council of Canada, and the Centre for Comparative Genomics and Evolutionary Bioinformatics at Dalhousie University. J.M.A. acknowledges long-term support from the Canadian Institute for Advanced Research.

REFERENCES

1. Stanier, R.Y., Douderoff, M., and Adelberg, E. (1963). The Microbial World, Second Edition (Englewood Cliffs, N.J: Prentice-Hall).
2. Sapp, J. (2005). The prokaryote-eukaryote dichotomy: meanings and mythology. Microbiol. Mol. Biol. Rev. *69*, 292–305.
3. Chatton, E. (1938). Titre et Travaux Scientifique (1906-1937) de Edouard Chatton (Scottano, Italy: Sette).
4. Woese, C.R., and Fox, G.E. (1977). Phylogenetic structure of the prokaryotic domain. The primary kingdoms. Proc. Natl. Acad. Sci. USA *74*, 5088–5090.
5. Woese, C.R., Kandler, O., and Wheelis, M.L. (1990). Towards a natural system of organisms, proposal for the domains Archaea, Bacteria, and Eucarya. Proc. Natl. Acad. Sci. USA *87*, 4576–4579.

6. Baldauf, S.L., Palmer, J.D., and Doolittle, W.F. (1996). The root of the universal tree and the origin of eukaryotes based on elongation factor phylogeny. Proc. Natl. Acad. Sci. USA *93*, 7749–7754.

7. Gogarten, J.P., Kibak, H., Dittrich, P., Taiz, L., Bowman, E.J., Bowman, B.J., Manolson, M.F., Poole, R.J., Date, T., Oshima, T., Konishi, J., *et al*. (1989). Evolution of the vacuolar H^+-ATPase: implications for the origin of eukaryotes. Proc. Natl. Acad. Sci. USA *86*, 6661–6665.

8. Iwabe, N., Kuma, K., Hasegawa, M., Osawa, S., and Miyata, T. (1989). Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. Proc. Natl. Acad. Sci. USA *86*, 9355–9359.

9. Philippe, H., and Forterre, P. (1999). The rooting of the universal tree of life is not reliable. J. Mol. Evol. *49*, 509–523.

10. Mayr, E. (1998). Two empires or three? Proc. Natl. Acad. Sci. USA *95*, 9720–9723.

11. Woese, C.R. (1998). Default taxonomy: Ernst Mayr's view of the microbial world. Proc. Natl. Acad. Sci. USA *95*, 11043–11046.

12. Pace, N.R. (2006). Time for a change. Nature *441*, 289.

13. Makarova, K.S., and Koonin, E.V. (2003). Comparative genomics of Archaea: how much have we learned in six years, and what's next? Genome Biol. *4*, 115.

14. Williams, T.A., and Embley, T.M. (2014). Archaeal "dark matter" and the origin of eukaryotes. Genome Biol. Evol. *6*, 474–481.

15. Williams, T.A., Foster, P.G., Cox, C.J., and Embley, T.M. (2013). An archaeal origin of eukaryotes supports only two primary domains of life. Nature *504*, 231–236.

16. Lake, J.A., Henderson, E., Oakes, M., and Clark, M.W. (1984). Eocytes: a new ribosome structure indicates a kingdom with a close relationship to eukaryotes. Proc. Natl. Acad. Sci. USA *81*, 3786–3790.

17. Honneger, R. (2002). Simon Schwendener (1829-1919) and the dual hypothesis of lichens. Bryologist *103*, 307–313.

18. Sapp, J. (1994). Evolution By Association: A History of Symbiosis (New York: Oxford University Press).

19. de Bary, A. (1879). Die Erscheinung der Symbiose (Privately printed in Strasbourg).

20. Berch, S.M., Massicotte, H.B., and Tackaberry, L.E. (2005). Re-publication of a translation of 'The vegetative organs of *Monotropa hypopitys* L.' published by F. Kamienski in 1882, with an update on *Monotropa* mycorrhizas. Mycorrhiza *15*, 323–332.

21. Frank, A.B., and Trappe, J.M. (2005). On the nutritional dependence of certain trees on root symbiosis with belowground fungi (an English translation of A.B. Frank's classic paper of 1885). Mycorrhiza *15*, 267–275.

22. Mereschkowsky, C. (1920). La plante considerée comme un complexe symbiotique. Societé des Sciences Naturelles de l'Ouest de la France, Nante, Bulletin *6*, 17–98.

23. Sapp, J., Carrapico, F., and Zolotonosov, M. (2002). Symbiogenesis: the hidden face of Constantin Merezhkowsky. Hist. Phil. Life Sci. *24*, 413–440.

24. Martin, W., and Kowallik, K.V. (1999). Annotated English translation of Mereschkowsky's 1905 paper 'Uber Natur and Ursprung der Chromatophoren im Pflanzenreiche'. Eur. J. Phycol. *34*, 287–295.

25. Schimper, A.F.W. (1883). Ueber die Entwickelung der Chlorophyllkörner und Farbkörper. Bot. Zeit. *41*, 105–114, 121–131, 137–146, 153–162.

26. Archibald, J.M. (2014). One Plus One Equals One: Symbiosis and the Evolution of Complex Life (Oxford University Press).

27. Margulis, L. (1998). Symbiotic Planet (Basic Books).

28. Plaut, W., and Alexander Sagan, L. (1958). Incorporation of thymidine in the cytoplasm of *Amoeba proteus*. J. Biophys. Biochem. Cytol. *4*, 843–846.

29. Sagan, L. (1965). An unusual pattern of tritiated thymidine incorporation in *Euglena*. J. Protozool. *12*, 105–109.

30. Ris, H., and Plaut, W. (1962). Ultrastructure of DNA-containing areas in the chloroplast of *Chlamydomonas*. J. Cell Biol. *13*, 383–391.

31. Nass, M.M.K., and Nass, S. (1963). Intramitochondrial fibers with DNA characteristics. I. Fixation and electron staining reactions. J. Cell. Biol. *19*, 593–611.

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

33. Margulis, L. (1970). Origin of Eukaryotic Cells (Yale University Press).

34. Goksøyr, J. (1967). Evolution of eucaryotic cells. Nature *214*, 1161.

35. Raven, P.H. (1970). A multiple origin for plastids and mitochondria. Science *169*, 641–646.

36. Klein, R., and Cronquist, A. (1967). A consideration of the evolutionary and taxonomic significance of some biochemical, micromorphological and physiological characters in the Thallophytes. Quart. Rev. Biol. *42*, 105–296.

37. Cavalier-Smith, T. (1975). The origin of nuclei and of eukaryotic cells. Nature *256*, 463–467.

38. Margulis, L. (1981). Symbiosis in Cell Evolution (W. H. Freeman and Company).

39. Bogorad, L. (1975). Evolution of organelles and eukaryotic genomes. Science *188*, 891–898.

40. Raff, R.A., and Mahler, H.R. (1972). The non symbiotic origin of mitochondria. Science *177*, 575–582.

41. Uzzell, T., and Spolsky, C. (1974). Mitochondria and plastids as endosymbionts: a revival of special creation? Am. Sci. *62*, 334–343.

42. Taylor, F.J.R. (1974). Implications and extensions of the serial endosymbiosis theory of the origin of eukaryotes. Taxon *23*, 229–258.

43. Sapp, J. (2009). The New Foundations of Evolution (Oxford University Press).

44. Bonen, L., and Doolittle, W.F. (1975). On the prokaryotic nature of red algal chloroplasts. Proc. Natl. Acad. Sci. USA *72*, 2310–2314.

45. Bonen, L., and Doolittle, W.F. (1976). Partial sequences of 16S rRNA and the phylogeny of blue-green algae and chloroplasts. Nature *261*, 669–673.

46. Zablen, L.B., Kissil, M.S., Woese, C.R., and Buetow, D.E. (1975). Phylogenetic origin of the chloroplast and prokaryotic nature of its ribosomal RNA. Proc. Natl. Acad. Sci. USA *72*, 2418–2422.

47. Bonen, L., Cunningham, R.S., Gray, M.W., and Doolittle, W.F. (1977). Wheat embryo mitochondrial 18S ribosomal RNA: evidence for its prokaryotic nature. Nucleic Acids Res. *4*, 663–671.

48. Yang, D., Oyaizu, Y., Oyaizu, H., Olsen, G.J., and Woese, C.R. (1985). Mitochondrial origins. Proc. Natl. Acad. Sci. USA *82*, 4443–4447.

49. Gray, M.W. (1992). The endosymbiont hypothesis revisited. Int. Rev. Cytol. *141*, 233–357.

50. Gray, M.W., and Doolittle, W.F. (1982). Has the endosymbiont hypothesis been proven? Microbiol. Rev. *46*, 1–42.

51. Embley, T.M., and Martin, W. (2006). Eukaryotic evolution, changes and challenges. Nature *440*, 623–630.

52. Koonin, E.V. (2010). The origin and early evolution of eukaryotes in the light of phylogenomics. Genome Biol. *11*, 209.

53. O'Malley, M.A. (2010). The first eukaryote cell: an unfinished history of contestation. Stud. Hist. Philos. Biol. Biomed. Sci. *41*, 212–224.

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

55. Cavalier-Smith, T. (1983). A 6-kingdom classification and a unified phylogeny. In Endocytobiology, W. Schwemmler, and J.E.A. Schenk, eds. (Berlin: de Gruyter), pp. 1027–1034.

56. Cavalier-Smith, T. (1987). Eukaryotes with no mitochondria. Nature *326*, 332–333.

57. de Duve, C. (1969). Evolution of the peroxisome. Ann. NY Acad. Sci. USA *168*, 369–381.

58. Stanier, R.Y. (1970). Some aspects of the biology of cells and their possible evolutionary significance. In Organization and Control in Prokaryotic and Eukaryotic Cells: 20th symposium of the Society for General Microbiology, H.P. Charles, and B.D. Knight, eds. (London: Cambridge University Press), pp. 1–38.

59. Hashimoto, T., Nakamura, Y., Kamaishi, T., and Hasegawa, M. (1997). Early evolution of eukaryotes inferred from the amino acid sequences of elongation factors 1a and 2. Arch. Protistenkd. *148*, 287–295.

60. Leipe, D.D., Gunderson, J.H., Nerad, T.A., and Sogin, M.L. (1993). Small subunit ribosomal RNA+ of *Hexamita inflata* and the quest for the first branch in the eukaryotic tree. Mol. Biochem. Parasitol. *59*, 41–48.

61. Sogin, M.L., Gunderson, J.H., Elwood, H.J., Alonso, R.A., and Peattie, D.A. (1989). Phylogenetic meaning of the kingdom concept: an unusual ribosomal RNA from *Giardia lamblia*. Science *243*, 75–77.

62. Hirt, R.P., Logsdon, J.M., Jr., Healy, B., Dorey, M.W., Doolittle, W.F., and Embley, T.M. (1999). Microsporidia are related to Fungi: evidence from the largest subunit of RNA polymerase II and other proteins. Proc. Natl. Acad. Sci. USA *96*, 580–585.

63. Inagaki, Y., Susko, E., Fast, N.M., and Roger, A.J. (2004). Covarion shifts cause a long-branch attraction artifact that unites microsporidia and archaebacteria in EF-1alpha phylogenies. Mol. Biol. Evol. *21*, 1340–1349.

64. Philippe, H., Lopez, P., Brinkmann, H., Budin, K., Germot, A., Laurent, J., Moreira, D., Muller, M., and Le Guyader, H. (2000). Early-branching or fast-evolving eukaryotes? An answer based on slowly evolving positions. Proc. Biol. Sci. *267*, 1213–1221.

65. Embley, T.M., van der Giezen, M., Horner, D.S., Dyal, P.L., and Foster, P.G. (2002). Mitochondria and hydrogenosomes are two forms of the same fundamental organelle. Phil. Trans. R. Soc. Lond. B *358*, 191–203.

66. Tovar, J. (2007). Mitosomes of parasitic protozoa: biology and evolutionary significance. In Origin of Mitochondria and Hydrogenosomes, W.F. Martin, and M. Muller, eds. (Berlin: Springer-Verlag), pp. 277–300.

67. Stechmann, A., Hamblin, K., Perez-Brocal, V., Gaston, D., Richmond, G.S., van der Giezen, M., Clark, C.G., and Roger, A.J. (2008). Organelles in *Blastocystis* that blur the distinction between mitochondria and hydrogenosomes. Curr. Biol. *18*, 580–585.

68. Hjort, K., Goldberg, A.V., Tsaousis, A.D., Hirt, R.P., and Embley, T.M. (2010). Diversity and reductive evolution of mitochondria among microbial eukaryotes. Phil. Trans. R. Soc. Lond. B *365*, 713–727.

69. Müller, M., Mentel, M., van Hellemond, J.J., Henze, K., Woehle, C., Gould, S.B., Yu, R.Y., van der Giezen, M., Tielens, A.G., and Martin, W.F. (2012). Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. Microbiol. Mol. Biol. Rev. *76*, 444–495.

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

71. Husnik, F., Nikoh, N., Koga, R., Ross, L., Duncan, R.P., Fujie, M., Tanaka, M., Satoh, N., Bachtrog, D., Wilson, A.C., *et al*. (2013). Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. Cell *153*, 1567–1578.

72. Dagan, T., Roettger, M., Stucken, K., Landan, G., Koch, R., Major, P., Gould, S.B., Goremykin, V.V., Rippka, R., Tandeau de Marsac, N., *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.

73. Esser, C., Martin, W., and Dagan, T. (2007). The origin of mitochondria in light of a fluid prokaryotic chromosome model. Biol. Lett. *3*, 180–184.

74. Leger, M.M., Gawryluk, R.M., Gray, M.W., and Roger, A.J. (2013). Evidence for a hydrogenosomal-type anaerobic ATP generation pathway in *Acanthamoeba castellanii*. PLoS One *8*, e69532.

75. Nyvltova, E., Stairs, C.W., Hrdy, I., Ridl, J., Mach, J., Paces, J., Roger, A.J., and Tachezy, J. (2015). Lateral gene transfer and gene duplication played a key role in the evolution of *Mastigamoeba balamuthi* hydrogenosomes. Mol. Biol. Evol. *32*, 1039–1055.

76. Takishita, K., Chikaraishi, Y., Leger, M.M., Kim, E., Yabuki, A., Ohkouchi, N., and Roger, A.J. (2012). Lateral transfer of tetrahymanol-synthesizing genes has allowed multiple diverse eukaryote lineages to independently adapt to environments without oxygen. Biol. Direct *7*, 5.

77. Hug, L.A., Stechmann, A., and Roger, A.J. (2010). Phylogenetic distributions and histories of proteins involved in anaerobic pyruvate metabolism in eukaryotes. Mol. Biol. Evol. *27*, 311–324.

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

79. Eme, L., and Doolittle, W.F. (2015). Microbial diversity: a bonanza of phyla. Curr. Biol. *25*, R227–R230.

80. Hartman, H., and Fedorov, A. (2002). The origin of the eukaryotic cell: a genomic investigation. Proc. Natl. Acad. Sci. USA *99*, 1420–1425.

81. Pilhofer, M., Rosati, G., Ludwig, W., Schleifer, K.H., and Petroni, G. (2007). Coexistence of tubulins and ftsZ in different *Prosthecobacter* species. Mol. Biol. Evol. *24*, 1439–1442.

82. Yutin, N., and Koonin, E.V. (2012). Archaeal origin of tubulin. Biol. Direct *7*, 10.

83. Embley, T.M., and Williams, T.A. (2015). Evolution: Steps on the road to eukaryotes. Nature *521*, 169–170.

84. Spang, A., Saw, J.H., Jorgensen, S.L., Zaremba-Niedzwiedzka, K., Martijn, J., Lind, A.E., van Eijk, R., Schleper, C., Guy, L., and Ettema, T.J. (2015). Complex archaea that bridge the gap between prokaryotes and eukaryotes. Nature *521*, 173–179.

85. Reyes-Prieto, A., Weber, A.P., and Bhattacharya, D. (2007). The origin and establishment of the plastid in algae and plants. Annu. Rev. Genet. *41*, 147–168.

86. Schwartz, R.M., and Dayhoff, M.O. (1978). Origins of prokaryotes, eukaryotes, mitochondria, and chloroplasts. Science *199*, 395–403.

87. Lewin, R.A., and Withers, N.W. (1975). Extraordinary pigment composition of a prokaryotic alga. Nature *256*, 735–737.

88. Gould, S.B., Waller, R.F., and McFadden, G.I. (2008). Plastid evolution. Annu. Rev. Plant Biol. *59*, 491–517.

89. Kim, E., and Archibald, J.M. (2009). Diversity and evolution of plastids and their genomes. In The Chloroplast-Interactions with the Environment, H. Aronsson, and A.S. Sandelius, eds. (Berlin: Springer-Verlag), pp. 1–39.

90. Parfrey, L.W., Lahr, D.J., Knoll, A.H., and Katz, L.A. (2011). Estimating the timing of early eukaryotic diversification with multigene molecular clocks. Proc. Natl. Acad. Sci. USA *108*, 13624–13629.

91. Yoon, H.S., Hackett, J.D., Ciniglia, C., Pinto, G., and Bhattacharya, D. (2004). A molecular timeline for the origin of photosynthetic eukaryotes. Mol. Biol. Evol. *21*, 809–818.

92. Ball, S.G., Colleoni, C., Kadouche, D., Ducatez, M., Arias, M.C., and Tirtiaux, C. (2015). Toward an understanding of the function of Chlamydiales in plastid endosymbiosis. Biochim. Biophys. Acta *1847*, 495–504.

93. Deschamps, P. (2014). Primary endosymbiosis: have cyanobacteria and Chlamydiae ever been roommates? Acta Soc. Bot. Pol. *83*, 291–302.

94. Zimorski, V., Ku, C., Martin, W.F., and Gould, S.B. (2014). Endosymbiotic theory for organelle origins. Curr. Opin. Microbiol. *22C*, 38–48.

95. Gibbs, S.P. (2006). Looking at life: from binoculars to the electron microscope. Annu. Rev. Plant Biol. *57*, 1–17.

96. Keeling, P.J. (2013). The number, speed, and impact of plastid endosymbioses in eukaryotic evolution. Annu. Rev. Plant Biol. *64*, 583–607.

97. Rogers, M.B., Gilson, P.R., Su, V., McFadden, G.I., and Keeling, P.J. (2007). The complete chloroplast genome of the chlorarachniophyte *Bigelowiella natans*: evidence for independent origins of chlorarachniophyte and euglenid secondary endosymbionts. Mol. Biol. Evol. *24*, 54–62.

98. Archibald, J.M. (2009). The puzzle of plastid evolution. Curr. Biol. *19*, R81–R88.

99. Cavalier-Smith, T. (1999). Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. J. Eukaryot. Microbiol. *46*, 347–366.

100. Gould, S.B., Maier, U.G., and Martin, W.F. (2015). Protein import and the origin of red complex plastids. Curr. Biol. *25*, R515–R521.

101. Bodył, A. (2006). Did the peridinin plastid evolve through tertiary endosymbiosis? A hypothesis. Eur. J. Phycol. *41*, 435–448.

102. Petersen, J., Ludewig, A.K., Michael, V., Bunk, B., Jarek, M., Baurain, D., and Brinkmann, H. (2014). *Chromera velia*, endosymbioses and the rhodoplex hypothesis–plastid evolution in cryptophytes, alveolates, stramenopiles, and haptophytes (CASH lineages). Genome Biol. Evol. *6*, 666–684.

103. Sanchez-Puerta, M.V., and Delwiche, C.F. (2008). A hypothesis for plastid evolution in chromalveolates. J. Phycol. *44*, 1097–1107.

104. Stiller, J.W., Schreiber, J., Yue, J., Guo, H., Ding, Q., and Huang, J. (2014). The evolution of photosynthesis in chromist algae through serial endosymbioses. Nat. Comm. *5*, 5764.

105. Baurain, D., Brinkmann, H., Petersen, J., Rodriguez-Ezpeleta, N., Stechmann, A., Demoulin, V., Roger, A.J., Burger, G., Lang, B.F., and Philippe, H. (2010). Phylogenomic evidence for separate acquisition of plastids in cryptophytes, haptophytes, and stramenopiles. Mol. Biol. Evol. *27*, 1698–1709.

106. Gaget, P., Bodył, A., Mackiewicz, P., and Stiller, J.W. (2014). Tertiary plastid endosymbioses in dinoflagellates. In Endosymbiosis, W. Löffelhardt, ed. (Wein: Springer-Verlag), pp. 233–290.

107. Seeber, F., and Soldati-Favre, D. (2010). Metabolic pathways in the apicoplast of apicomplexa. Int. Rev. Cell Mol. Biol. *281*, 161–228.

108. Abrahamsen, M.S., Templeton, T.J., Enomoto, S., Abrahante, J.E., Zhu, G., Lancto, C.A., Deng, M., Liu, C., Widmer, G., Tzipori, S., *et al*. (2004). Complete genome sequence of the apicomplexan, *Cryptosporidium parvum*. Science *304*, 441–445.

109. Huang, J., Mullapudi, N., Lancto, C.A., Scott, M., Abrahamsen, M.S., and Kissinger, J.C. (2004). Phylogenomic evidence supports past endosymbiosis, intracellular and horizontal gene transfer in *Cryptosporidium parvum*. Genome Biol. *5*, R88.

110. Gornik, S.G., Febrimarsa, Cassin, A.M., MacRae, J.I., Ramaprasad, A., Rchiad, Z., McConville, M.J., Bacic, A., McFadden, G.I., Pain, A., *et al*. (2015). Endosymbiosis undone by stepwise elimination of the plastid in a parasitic dinoflagellate. Proc. Natl. Acad. Sci. USA *112*, 5767–5772.

111. Archibald, J.M. (2015). Genomic perspectives on the birth and spread of plastids. Proc. Natl. Acad. Sci. USA. http://dx.doi.org/10.1073/pnas.1421374112.

112. Ku, C., Nelson-Sathi, S., Roettger, M., Garg, S., Hazkani-Covo, E., and Martin, W.F. (2015). Endosymbiotic gene transfer from prokaryotic pangenomes: Inherited chimerism in eukaryotes. Proc. Natl. Acad. Sci. USA. http://dx.doi.org/10.1073/pnas.1421385112.

113. Stiller, J.W. (2011). Experimental design and statistical rigor in phylogenomics of horizontal and endosymbiotic gene transfer. BMC Evol. Biol. *11*, 259.

114. Melkonian, M., and Mollenhauer, D. (2005). Robert Lauterborn (1869-1952) and his *Paulinella chromatophora*. Protist *156*, 253–262.

115. Nowack, E.C. (2014). *Paulinella chromatophora*—rethinking the transition from endosymbiont to organelle. Acta Soc. Botan. Polon. *83*, 387–397.

116. Nowack, E.C., and Grossman, A.R. (2012). Trafficking of protein into the recently established photosynthetic organelles of *Paulinella chromatophora*. Proc. Natl. Acad. Sci. USA *109*, 5340–5345.

117. Cavalier-Smith, T., and Lee, J.J. (1985). Protozoa as hosts for endosymbioses and the conversion of symbionts into organelles. J. Protozool. *32*, 376–379.

118. Theissen, U., and Martin, W. (2006). The difference between organelles and endosymbionts. Curr. Biol. *16*, R1016–R1017.

119. Nakayama, T., Ikegami, Y., Nakayama, T., Ishida, K.-I., Inagaki, Y., and Inouye, I. (2011). Spheroid bodies in Rhopalodiacean diatoms were derived from a single endosymbiotic cyanobacterium. J. Plant Res. *124*, 93–97.

120. Kneip, C., Voss, C., Lockhart, P.J., and Maier, U.G. (2008). The cyanobacterial endosymbiont of the unicellular algae *Rhopalodia gibba* shows reductive genome evolution. BMC Evol. Biol. *8*, 30.

121. Nakayama, T., Kamikawa, R., Tanifuji, G., Kashiyama, Y., Ohkouchi, N., Archibald, J.M., and Inagaki, Y. (2014). Complete genome of a nonphotosynthetic cyanobacterium in a diatom reveals recent adaptations to an intracellular lifestyle. Proc. Natl. Acad. Sci. USA *111*, 11407–11412.

122. Moran, N.A., and Bennett, G.M. (2014). The tiniest tiny genomes. Annu. Rev. Microbiol. *68*, 195–215.

123. Bennett, G.M., and Moran, N.A. (2013). Small, smaller, smallest: the origins and evolution of ancient dual symbioses in a phloem-feeding insect. Genome Biol. Evol. *5*, 1675–1688.

124. Nakabachi, A., Ishida, K., Hongoh, Y., Ohkuma, M., and Miyagishima, S. (2014). Aphid gene of bacterial origin encodes a protein transported to an obligate endosymbiont. Curr. Biol. *24*, R640–R641.

125. McCutcheon, J.P., and Keeling, P.J. (2014). Endosymbiosis: protein targeting further erodes the organelle/symbiont distinction. Curr. Biol. *24*, R654–R655.

126. Keeling, P.J. (2014). The impact of history on our perception of evolutionary events: endosymbiosis and the origin of eukaryotic complexity. Cold Spring Harb. Perspect. Biol. *6*, a016196.

127. de Duve, C. (1982). Peroxisomes and related particles in historical perspective. Ann. NY Acad. Sci. *386*, 1–4.

128. de Duve, C. (2007). The origin of eukaryotes: a reappraisal. Nat. Rev. Genet. *8*, 395–403.

129. Gabaldon, T. (2010). Peroxisome diversity and evolution. Phil. Trans. R. Soc. Lond. B *365*, 765–773.

130. Mereschkowsky, C. (1910). Theorie der zwei Plasmaarten als Grundlage der Symbiogenesis, einer neuen Lehre von der Entstehung der Organismen. Biol. Centralbl. *30*, 353–442.

131. Martin, W., and Koonin, E.V. (2006). Introns and the origin of nucleus-cytosol compartmentalization. Nature *440*, 41–45.

132. Rogozin, I.B., Carmel, L., Csuros, M., and Koonin, E.V. (2012). Origin and evolution of spliceosomal introns. Biol. Direct *7*, 11.

133. Baum, D.A., and Baum, B. (2014). An inside-out origin for the eukaryotic cell. BMC Biol. *12*, 76.

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