Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Microbiology

Endosymbiotic theory for organelle origins

Verena Zimorski, Chuan Ku, William F Martin and Sven B Gould

Endosymbiotic theory goes back over 100 years. It explains the similarity of chloroplasts and mitochondria to free-living prokaryotes by suggesting that the organelles arose from prokaryotes through (endo)symbiosis. Gene trees provide important evidence in favour of symbiotic theory at a coarse-grained level, but the finer we get into the details of branches in trees containing dozens or hundreds of taxa, the more equivocal evidence for endosymbiotic events sometimes becomes. It seems that either the interpretation of some endosymbiotic events are wrong, or something is wrong with the interpretations of some gene trees having many leaves. There is a need for evidence that is independent of gene trees and that can help outline the course of symbiosis in eukaryote evolution. Protein import is the strongest evidence we have for the single origin of chloroplasts and mitochondria. It is probably also the strongest evidence we have to sort out the number and nature of secondary endosymbiotic events that have occurred in evolution involving the red plastid lineage. If we relax our interpretation of individual gene trees, endosymbiotic theory can tell us a lot.

Addresses
Institute of Molecular Evolution, Heinrich-Heine-University of Düsseldorf, 40225 Düsseldorf, Germany

Corresponding author: Martin, William F (bill@hhu.de, w.martin@hhu.de)
Dedicated to Klaus V Kowallik on the occasion of his 75th birthday.

Current Opinion in Microbiology 2014, **22**:38–48

This review comes from a themed issue on **Growth and development: eukaryotes**

Edited by **Michael Bölker**

http://dx.doi.org/10.1016/j.mib.2014.09.008

Introduction

Endosymbiotic theory posits that plastids and mitochondria were once free-living prokaryotes and became organelles of eukaryotic cells. The theory started with plastids [1] and was further developed for mitochondria [2]. It was rejected by cell biologists in the 1920s and revived in the 1960s [3]. The main strength of the theory is that it accounts for the physiological and biochemical similarity of organelles to prokaryotic cells [4,5]. Important evidence in support of endosymbiotic theory comes from organelle genomes. Organelles tend to retain a miniaturized prokaryotic chromosome encoding 200 proteins or less in the case of plastids [6] or 63 proteins or less in the case of mitochondria [7]. Despite that genome reduction, both organelles harbour on the order of 2000 proteins each [8,9], which are involved in a broad spectrum of pathways germane to their ancestrally prokaryotic biochemistry. The discrepancy between the number of proteins that organelles encode and the number of proteins that they harbour is generally explained by a corollary to endosymbiotic theory involving gene transfer to the nucleus, or endosymbiotic gene transfer (EGT). During the course of evolution, many genes were transferred from the organelles to the chromosomes of their host. In the early phases of organelle evolution, before the invention of the protein import apparatus that allowed plastids and mitochondria to import proteins from the cytosol, the transferred genes either became pseudogenes or became expressed as cytosolic proteins. With the advent of organelle protein import, the transferred genes could obtain the necessary expression and targeting signals to be targeted back to the organelle from which the nuclear gene was acquired [10]. For functions essential to the organelle, only the third case allowed the gene to be lost from organelle DNA [11]. This process of organelle genome reduction has resulted in an expansion of the eukaryotic nuclear gene repertoire and in reductive genome evolution in the organelle. While it has long been known that the genes retained most tenaciously by plastids and mitochondria encode for proteins involved in the electron transport chain of the bioenergetic organelle or for the ribosome required for their synthesis [12], only recently was it recognized that even within the ribosome, the same core of proteins has been retained independently by plastids and mitochondria, probably owing to constraints imposed by the process of ribosome assembly [13].

Endosymbiotic theory was also an important testing ground for molecular evolution. In the 1970s, there were competing theories to explain organelle origins. Those theories called for autogenous rather than symbiotic organelle origins and saw plastids and mitochondria as deriving from invaginations of the plasma membrane [14], from restructuring of thylakoids in a cyanobacterial ancestor of eukaryotes [15], or from budding of the nuclear membrane [16], as opposed to origins through symbiosis. They had it that the DNA in organelles stems from, and hence should be more similar in sequence to, genes encoded in nuclear DNA than to genes from free-living prokaryotes. That was a prediction that could be tested with DNA sequence comparisons. Bonen and Doolittle [17] found evidence for similarity between plastid and cyanobacterial nucleic acids, and Butow [18] found

evidence for mitochondrial genes that had been transferred to the nucleus in yeast. By about 1980, endogenous theories could be excluded and through 16S rRNA analyses, it was possible to confirm the origin of plastids from their suspected cyanobacterial ancestors [19] and to trace the origin of mitochondria to a metabolically versatile group of prokaryotes then called purple non-sulphur bacteria [20], later renamed to proteobacteria [21].

Protein import machineries as beacons for endosymbiotic events

Plastids and mitochondria each have a single origin. The strongest evidence for this comes from the protein import apparatus [22,23]. Had mitochondria become established in independent eukaryotic lineages, they would hardly have independently invented, via convergent evolution, the same core set of TIM and TOM components (translocon of the inner/outer mitochondrial membrane) that unite all mitochondria and organelles derived thereof [24•,25]. The same is true for the TIC and TOC systems (translocon of the inner/outer chloroplast membrane) of plastids [26,27]. The unity of these import machineries among mitochondria and plastids, respectively, is thus widely regarded as the best evidence we have for the single origin of these organelles, as opposed to multiple independent symbiotic origins in different lineages, even from endosymbionts so closely related as to be indistinguishable in phylogenies [28]. The establishment of a symbiotic cyanobacterium and its transition to the plastid ancestor is called primary symbiosis, it occurred perhaps some 1.2 billion years ago [29]. Subsequent to that, a number of secondary symbioses took place during evolution [30–32], in which eukaryotic algae became established as endosymbionts within eukaryotic cells, giving rise to what are called complex plastids, a term used to designate plastids surrounded by three or more membranes [33]. It is undisputed that secondary endosymbiosis occurred on at least three different occasions during eukaryote evolution: one in the lineage leading to the Euglenoids, a second independent event in the lineage leading to the Chlorarachniophytes and *at least* one more that led to the secondary plastids of red algal origin in diverse algal groups (Figure 1). For more than 20 years, the number and nature of secondary endosymbiotic events involving red algae has been heftily debated. Most of the debate has focussed on interpreting the differences between conflicting gene trees for the same groups [31,34,35,36•,37].

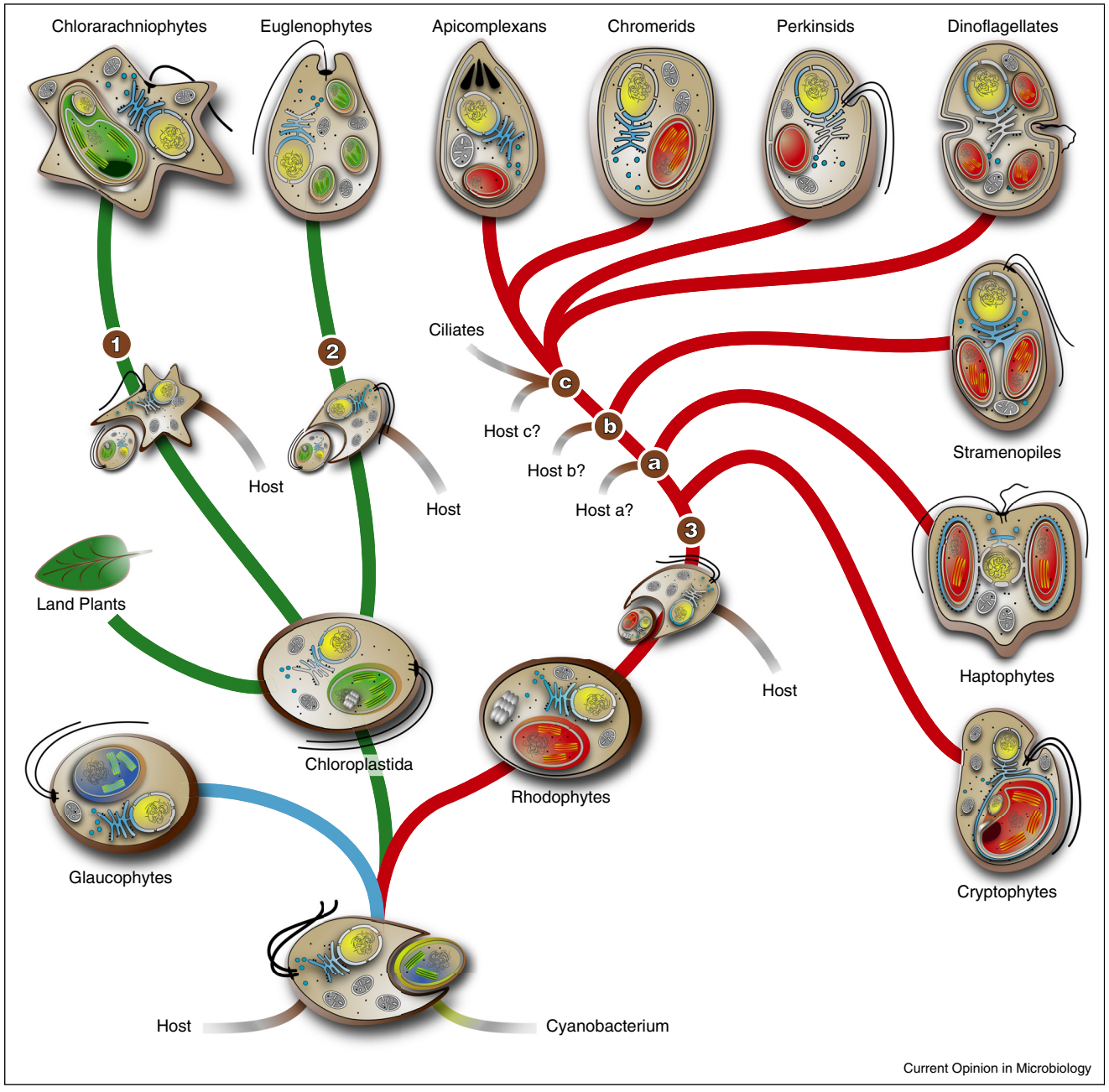
What if we step back from the trees and use the same reasoning and kind of data as the field uses to uncontentiously conclude that there was only one origin each of plastids and mitochondria? What if we look at the protein import machinery of red complex plastids of CASH lineages (Cryptophytes, Alveolates, Stramenopiles and Haptophytes)? Work in Uwe-G. Maier's group has shed light on the protein import machinery across the second outermost membrane of complex red plastids surrounded by four membranes [38,39]. That machinery is called SELMA (symbiont-specific ERAD-like machinery). SELMA is a multi-protein system that has been adopted from the symbiont's ERAD system (for endoplasmic reticulum (ER) associated degradation). In eukaryotic cells ERAD exports proteins from the ER for their degradation in the cytosol [40]. In *all* CASH plastids, a conserved N-terminal bipartite leader guides preproteins through the SELMA translocon across the second outermost membrane into the periplastidal compartment [39–42,43••].

For untangling red secondary symbioses, the crucial observation is that salient components of the SELMA are still encoded in the nucleomorph (the former nucleus of the engulfed red alga of cryptophytes; Figure 2) [38], and that protein import across the second outermost membrane of all CASH plastids involves a homologous SELMA machinery of monophyletic origin [42]. The SELMA machinery arose only once in evolution (like TIM/TOM and TIC/TOC), and it arose in the nucleus of the secondary endosymbiont that gave rise to the complex red plastid of cryptophytes (Figure 2). That tells us that all red secondary plastids are derived from the same algal endosymbiont that gave rise to cryptophyte plastids — and from that it follows that there was one single secondary endosymbiosis at the origin of the red secondary plastids (symbiosis 3 in Figure 1). So far so good, but in symbiosis it takes two to tango and a single origin of the red complex plastid still does not tell us how many hosts were involved. It could be that all CASH groups descend from the same endosymbiotic event as Cavalier-Smith suggested in the chromalveolate hypothesis [44]. Or they only share the same plastid, in which case one or more of the CASH lineages could have acquired plastids via tertiary symbiosis (like in the rhodoplex hypothesis [36•]) by engulfing a member of the ancient lineage that lead to cryptophytes (possible additional symbioses a–c in Figure 1). Should the plastid of cryptophytes also be of tertiary origin, then the secondary red alga that established SELMA has yet to be identified. Some might suggest that SELMA was passed around through lateral gene transfer (LGT), but considering its functional complexity (about a dozen or more proteins [36•]) that seems unlikely. Also note that chlorarachniophytes harbour a complex plastid still containing a nucleomorph, too, but it is of green algal origin and does not use a SELMA-like translocon [45]. Many conflicting gene trees addressing the issue of red secondary plastid origins have to be wrong, or misleading, or both.

How green are the reds, how red are the greens?

The origin of red secondary plastids highlights issues about trees and their interpretation. This can be illustrated with one recent study concerning diatoms, whose

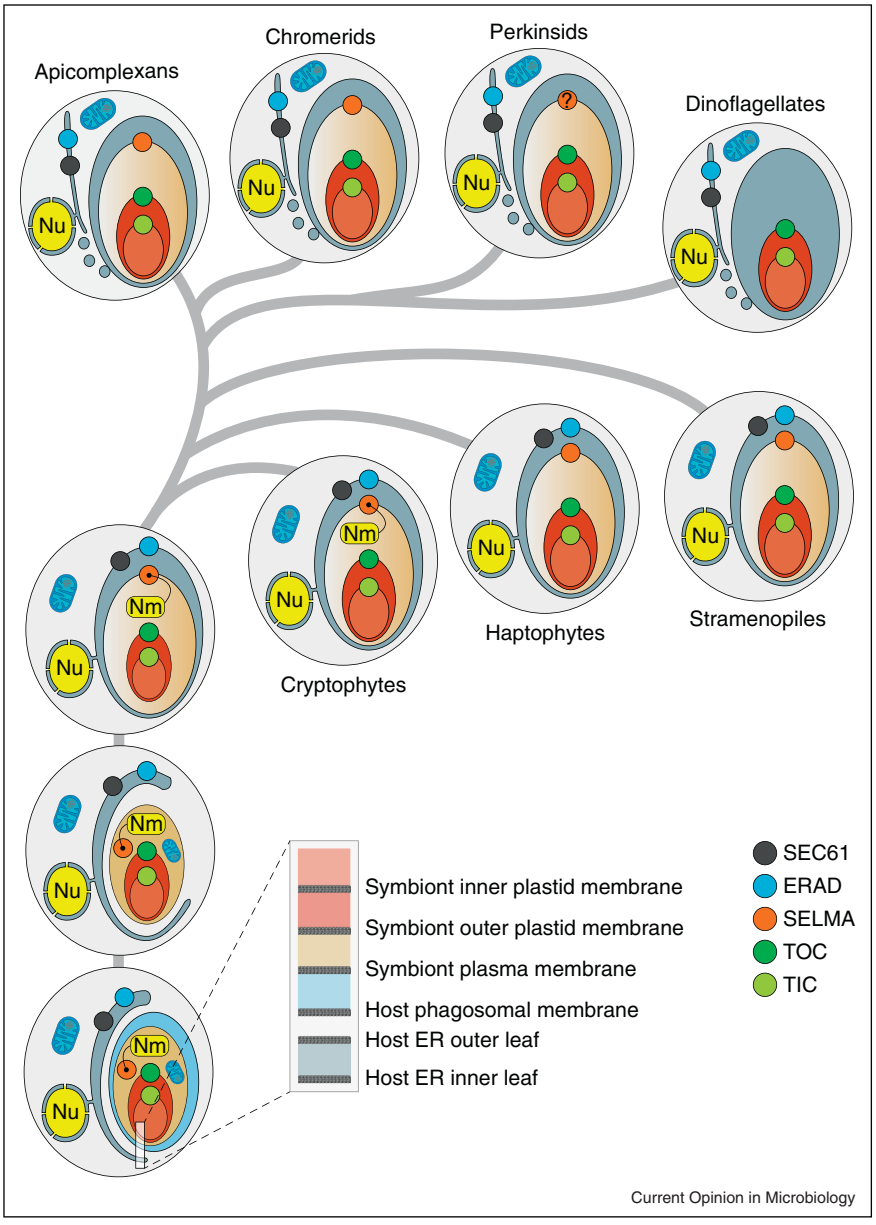
Figure 1



Plastid evolution. The initial uptake of a cyanobacterium by a heterotrophic host lead to three lineages: the Glaucophytes, Chloroplastida and Rhodophytes. Subsequently, two individual secondary endosymbiotic events involving algae of the Chloroplastida lineage and two heterotrophic hosts of unknown nature lead to the Chlorarachniophytes (symbiosis 1) and Euglenophytes (symbiosis 2). The radiation of secondary red plastids is not fully resolved, but the initial step was monophyletic, too (symbiosis 3) and connected to the origin of the SELMA translocon (see Figure 2 for details). While there is good evidence that the initial secondary plastid is of monophyletic origin, the amount of downstream-involved hosts remains uncertain (potential additional symbioses a–c). In some lineages red complex plastids could be of tertiary endosymbiotic origin. For details please refer to the text.
Modified from [30].

plastids unquestionably — based on plastid genome organization, not trees [46] — descend from red algae. Moustafa *et al.* [47] found that diatoms harbour many nuclear genes that branch with red algal homologues, as they should, *if* their plastids indeed are derived from the red lineage, which they are, and *if* many genes have been transferred from organelles to the nucleus during evolution, which has happened [48,49]. The problem is that

Figure 2



SELMA and the evolution of the CASH lineages. Schematic model for the evolution and radiation of SELMA among protists with complex red plastids. The red algal endosymbiont was initially encapsulated by a phagosomal membrane that separated it from the hosts' cytosol. That membrane was lost first, and after which a part of the hosts' endoplasmic reticulum wrapped around the endosymbiont (similar, but not identical to the 'autophagosome model' [102]). This step was accompanied by the loss of the endosymbionts' plasma membrane, mitochondrion and ER. The two eukaryotic cytosols fused and the nucleomorph (Nm)-encoded SELMA was now integrated into the inner face of the host ER membrane after the endosymbionts ER was lost. This process established the SELMA system, which is now found in all organisms with complex red plastids, but where it is now encoded in the nucleus (Nu), except for cryptophytes, where it remains Nm-encoded. Peridinin-containing dinoflagellates, whose plastids are surrounded by only three membranes, are the only exception: they appear to have lost the SELMA machinery altogether, when loosing an additional complex plastid membrane.

they found just as many diatom nuclear genes branching with green algae as with red. The same red versus green problem was observed in an independent study on *Chromera*, a photosynthetic relative of Apicomplexans [50]. And to complicate the matter, the same observation, but vice versa, was made in the genome of the chlorarachniophyte (Figure 1) *Bigelowiella natans* that houses an endosymbiont of green origin: of the 353 algal genes identified, 45 (22%) were found to branch with red algae [51••]. Hence, the results and the effects are reproducible. Some

will ask whether green plastids are frequently being replaced by red ones, and vice versa, during algal evolution, but maybe the first question we should ask is: Are trees simply fraught with systematic or random errors in such a way that diatoms end up on the green branch very often, when they really belong on the red branch [52]?

Is molecular phylogeny really that badly error prone? It well could be. In one study of a known phylogeny involving two grasses, a dicot, a gymnosperm, a liverwort and a red alga, only 40 out of 58 chloroplast encoded proteins (where there is no paralogy and no lateral gene transfer for the genes in question) recovered the true tree [53]. In a study of nine plastid genomes only 11 out of 42 genes recovered the consensus tree [54]. The simplest interpretation of such findings is that phylogeny is an imperfect art and that we should always expect some unexpected branches. The problem is that we do not know how many or which unexpected branches to expect. But the more ancient the phylogeny and the more species in the tree, the more we should expect to see spurious branches. In theory, for a tree with 38 leaves (taxa), there are roughly 10^{51} possible trees: the chances of getting the right one are the same as picking the same proton out of all the protons on Earth (6×10^{50}) twice in a row. So if we see a tree with three-dozen leaves, it is possible that many branches are wrong, we just don't know which ones are wrong or how wrong they are. Even the branches with strong bootstrap or other support values can be wrong, because support values just tell us how often the algorithm and the data produce the branch in the computer, not whether the model or the branch is correct [55]. And when the trees contain prokaryotic leaves, the problems get worse, because of LGT among prokaryotes [56].

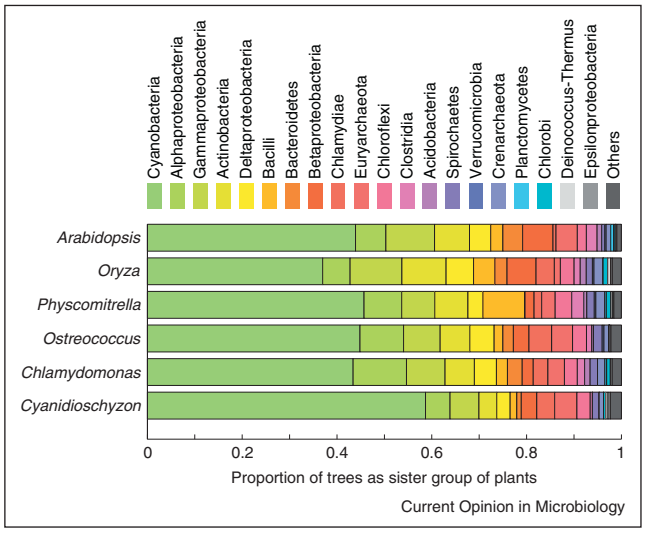
Of course, the alternative to assuming that phylogenetic trees are inherently imperfect is to assume that they are telling us the true course of history, just the way it was, every branch in every tree reflecting some past event, whose existence can be inferred because of some edge (the mathematical term for branch) that a computer produces. This is a good place to recall that plastids and mitochondria are biological entities in nature, things that we can observe and whose origins require an evolutionary explanation. By contrast, branches in phylogenetic trees are not observations of things in nature, they are things that computers generate when instructed by humans to produce them from input data — whether or not branches in phylogenetic trees require any explanation at all is debateable. Trees and branches are most effective when we use them as tools to test theories about evolution rather than as crayons to draw evolutionary history from scratch. The problem is that each tree tells a different story and if we can believe one tree all the others must be wrong, which can lead to exhausting debates of which gene tree is telling the true story, or if we look at the matter openly, whether *any* gene tree is telling the true story. Two recent developments concerning the use of gene trees in endosymbiotic theory, and the interpretation of those trees, underscore that point.

How much help did a cyanobacterium have becoming an endosymbiont?

In the genome sequence of *Chlamydia trachomatis* some genes were found that shared unexpectedly close phylogenetic relationships with plant homologs [57]. These unexpected branches were met with an array of explanations including direct LGT from eukaryotes to chlamydiae [57,58], or LGT in the other direction [59,60], indirect LGT to archaeplastids through the cyanobacterial endosymbiont [61], unrevealed relationship between archaeplastids and amoebas [54] or between cyanobacteria and chlamydiae [62], and gene transfer from mitochondria followed by differential loss [63]. Subsequent phylogenetic studies revealed a few more examples, and it was stated in its most recent formulation that '*Chlamydia-like pathogens are the second major source of foreign genes in Archaeplastida*' [64], and that the cyanobacterial origin of plastids was a symbiosis of three partners, with chlamydiae in an essential role of mediating metabolic integration of those partners [65–68].

The problem is not that modern cyanobacterial (endo)-symbioses observable in nature (lichens, cycads, *Azolla*, *Gunnera* or *Rhopalodia*) get by with just the cyanobacterium alone, with no aid from chlamydias, spirochaetes, or any other helper bacteria. The problem is also not that the benefit afforded to the host in those cyanobacterial symbioses is fixed nitrogen, not carbohydrate [69,70]. The problem is that when we look at all the trees that include prokaryotic lineages, chlamydiae no longer stand out [71]. Not much attention is paid to the overall potential gene origin in studies focusing on chlamydiae and plants alone [66,67,72]. If we apply the rationale of the chlamydial-helper hypothesis to genes apparently stemming from other prokaryotes, the endosymbiont hypothesis for plastids would be one involving many more 'helper' prokaryotes. Moreover, the '*second major*' [64], and we stress, apparent '*source*' of prokaryotic genes in plants is not chlamydia, it is alphaproteobacteria, followed by gammaproteobacteria, then actinobacteria, deltaproteobacteria, bacilli, bacteroidetes, and betaproteobacteria, behind which chlamydiae range as another meagre apparent donor (Figure 3) [71,48]. Did all of these lineages, and the lesser apparent donors, such as euryarchaeotes, clostridias, spirochaetes, planctomycetes and chlorobia help the cyanobacterium to become established as an endosymbiont or plastid? That should be the conclusion, if one takes the trees at face value. Furthermore, the genes in the different apparent donor lineage trees do not even branch with the same chlamydia, or the same proteobacteria, or for that matter of fact the same cyanobacteria. In the end, the single gene

Figure 3



Apparent prokaryotic donors of genes to plant lineages. Genes of many major prokaryotic lineages appear as nearest neighbours to archaeplastid nuclear genes in phylogenetic trees. Note that the apparent contribution of chlamydias is smaller than that of lineages such as actinobacteria, bacilli or bacteroidetes.
The figure is reproduced with permission from [71].

trees in which plants branch with cyanobacteria tell us that plastids arose from 60 or more different cyanobacteria [71]. Could that be?

An alternative would be to consider factors that are too often overlooked in studies of eukaryote gene origins in the context of organelle origins: random phylogenetic errors, limited taxon sampling, individual gene losses and LGT among prokaryotes [71,73–77]. Even if in an analysis the phylogenetic inference is completely correct and homologs from all extant organisms are included, LGT and gene losses in prokaryotes alone could still have produced the observed patterns [71,74–76,78,79]. In fact, LGT among prokaryotes is even evident in the trees in studies suggesting direct LGT (e.g. [80]), where the prokaryotic sister group of the eukaryotic clade is formed by homologs from more than one prokaryotic lineage, an observation that would not be possible had the gene never been transferred among prokaryotes. Even if the true donor was a cyanobacterium and gene phylogeny was error-free, loss of this gene or its absence from our limited sample of cyanobacteria and its transfer among prokaryotes since the origin of plastids could easily produce the pattern of apparent LGT from non-cyanobacterial sources.

Because of the single origin of plastids, the cyanobacterial ancestor of plastids was a unique prokaryotic organism. But as such, it had a pan-genome [81••]. What was the composition of its specific genome of the symbiont within that cyanobacterial pan-genome at the time of symbiosis? The best estimate probably comes from analysis of a frozen accident: the genes that plants acquired at the origin of plastids and that have persisted to the present in plant genomes. An analysis of 51 modern cyanobacterial genomes reveals 18 000 cyanobacterial gene families and 47 000 singletons [71], or a cyanobacterial pan-genome encompassing some 65 000 genes, whereby only about 5000 are found in any one cyanobacterium. Similarly, 61 strains of *Escherichia coli* have a pan-genome of about 18 000 genes, whereby only about 4500 are packaged in any given cell and only about 1000 genes (about 20% of the genome) are common to all *E. coli* strains within the species [82]. Thus, were an *E. coli* strain to become an endosymbiont today with the fate of turning into an organelle in a billion years, only about 20% of its genome would be defining for *E. coli* at the time of symbiosis, and the remainder would be shared with free-living *E. coli* strains, which would be free to generate new combinations of genes within and among species for the next billion years. In a billion years, the collection of genes that we call *E. coli* will no longer exist as an *E. coli* species complex, but most of the genes will still be around as descendant copies somewhere, just distributed among various genomes that would not be called *E. coli*. We do not know what happened a billion years and more ago, but we should keep in mind that, firstly, the genomes of the symbionts were already chimaeras; secondly, the descendants of the free-living relatives continued to experience LGT with other prokaryotes; and thirdly, phylogenetic tools are far from perfect.

An autogenous, ATP-consuming origin of mitochondria?

Another development that has unfolded around endosymbiosis could be called an issue of lumping and splitting. It centres around the origin of mitochondria. A good bit of progress has been made in understanding the role of mitochondria in eukaryote evolution in recent years. First, all eukaryote lineages are now known either to have or to have had a mitochondrion in their past [83••]. Second, the host that acquired the mitochondrion stems from a lineage that branches within the archaebacteria (or archaea), not as their sister [84••,85,86•]. Third, the presence of internalized bioenergetic membranes was the key attribute provided by mitochondrial endosymbiosis, which afforded eukaryotes many orders of magnitude more energy per gene than is available to prokaryotes [87]. Thus, while it has now been evident for some time that the common ancestor of eukaryotes possessed a mitochondrion, it is now clear why that was so: the lack of true intermediates in the prokaryote-to-eukaryote transition has a bioenergetic cause [87].

But beyond that, the origin of mitochondria is debated. Different phylogenomic analyses come to different

results regarding the nature of the free-living bacteria that are the closest relatives of mitochondria. Recent studies focussing on genes located in mitochondrial DNA, which is very AT-rich and thus prone to associate mitochondria, phylogenetically, to AT-rich proteobacteria, disagree with respect to the relationship of mitochondria to clades of free-living prokaryotes [88,89]. Different genes in mitochondrial DNA appear to trace to different sources in phylogenetic studies [90–92], as do different eukaryotic nuclear genes associated with mitochondrial functions [76,93,94]. Like in the case of plastids discussed above, such differences have causes that involve phylogenetic reconstruction, pan-genomes, and gene transfer among prokaryotes themselves [95], the relative contributions of which have however yet to be resolved. Amidst those debates, a careful and detailed survey of bioenergetic pathways and the diversity among components of the membrane-associated electron transport chain in free-living proteobacteria points to methylotrophic ancestors for mitochondria [96••], which is particularly interesting as the methylotrophs are metabolically versatile prokaryotes and have invaginations of their plasma membrane that rival the ultrastructural complexity of mitochondrial cristae [97].

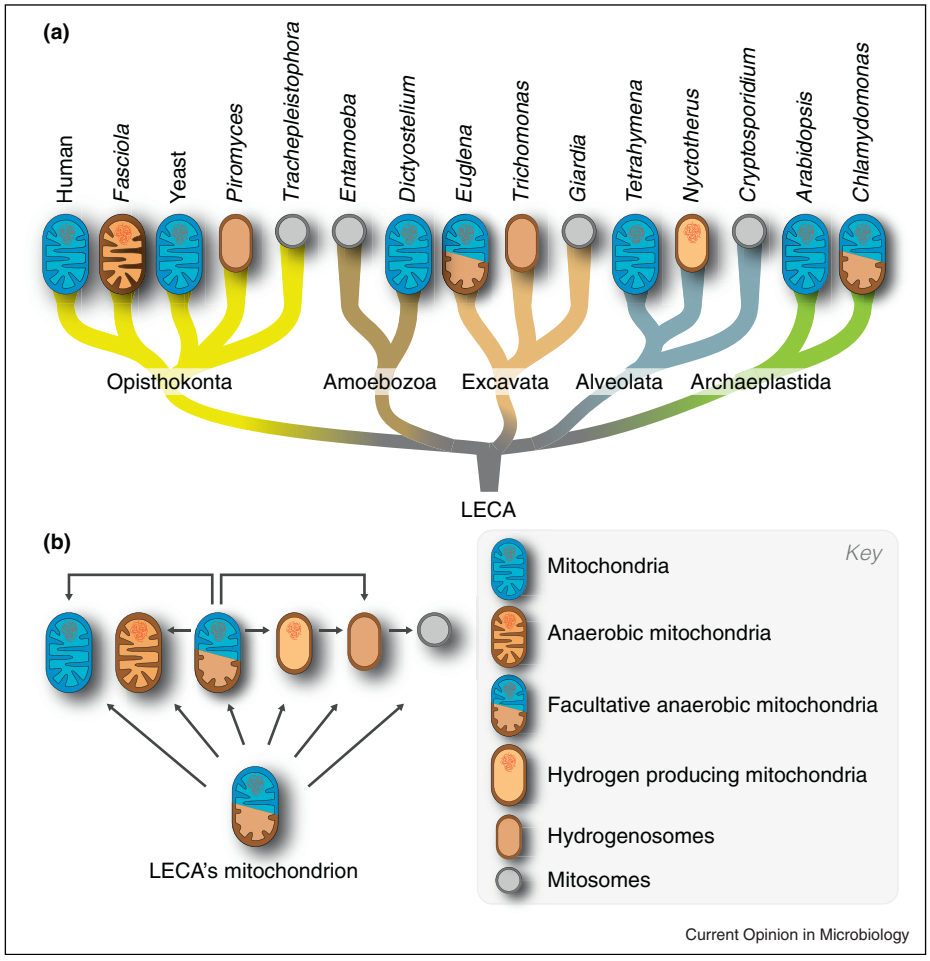
Some people still think that the main advantage of mitochondria and the key to eukaryote complexity was a roughly sixfold increase in energy yield from glucose. Indeed, with O_2-respiring mitochondria eukaryotes can harvest about 32 mol ATP per glucose, while with anaerobic mitochondria they can only glean about 5 mol ATP per glucose, and with hydrogenosomes they only harness about 4 mol ATP per glucose [98•]. But O_2 respiration cannot be the key to eukaryote complexity, for were that true, then *E. coli* and all other (facultative) aerobic prokaryotes should have become just as complex as eukaryotes, for the same reason of improved aerobic energy yield from glucose. The different manifestations of mitochondria in eukaryotes — aerobic, facultatively anaerobic, anaerobic, hydrogenosomes and mitosomes — have arisen independently as ecological specializations in different eukaryotic lineages (Figure 4), but essentially all of the genes involved in ATP production in organelles of mitochondrial origin were present in the eukaryote common ancestor [98•]. A competing alternative that the genes for anaerobic energy metabolism in eukaryotes were acquired late in eukaryotic evolution from donors that were distinct from the mitochondrion and then passed around from one eukaryote to another is favoured by some researchers [99,100], but the theory only accounts for sparse distributions of genes, which is just as simply accounted for by differential loss.

In search of one sentence on mitochondrial origin with which all prospective readers of this paper could agree, one could have recently risked: Mitochondria are organelles derived from a symbiosis between a bacterium that became the mitochondrion and a host. Yet, that formulation would not agree with the most recent view of mitochondrial evolution by Gray [101•], who was once a strong proponent of the endosymbiotic theory, but who now argues that the mitochondrial compartment was present before the organism that we call the mitochondrial endosymbiont entered the cell. His argument is that only comparatively few genes for mitochondrial proteins — 10–20% in his estimate — tend to reflect an alphaproteobacterial ancestry in single gene phylogenetic trees. The rest do not, they branch elsewhere among prokaryotic or eukaryotic homologues. From that he infers that only the genes that branch with alphaproteobacterial homologues come from endosymbiosis, while the remainder, the majority of genes whose products function in mitochondria today, were already present before the alphaproteobacterial symbiosis in an autogenously originated compartment: the pre-mitochondrion, which is envisaged as an ancestrally ATP consuming compartment. Its proteinaceous contents were specifically retargeted to the alphaproteobacterial invader, transforming it into a mitochondrion.

Gray's hypothesis, called the pre-endosymbiont hypothesis [101•], is not designed to explain the origin of mitochondria, it is designed to explain the origin of the many mitochondrial proteins that do not branch with alphaproteobacterial homologues. That is, it is designed to explain branching patterns in individual gene trees, which, as we saw in the case of chlamydiae, can be more complicated than it would seem at first glance. Like the chlamydial-helper hypothesis, the pre-endosymbiont hypothesis divides the world into, in this case, mitochondrial proteins whose trees branch with a particular group (chlamydiae, alphaproteobacteria) and those that do not. A disconcerting aspect of the theory is that it arbitrarily lumps and splits: it splits off into one bin all the mitochondrial proteins that branch with present-day alphaproteobacterial homologues and lumps together into a second bin all the ones that do not. While the former are assumed to come from the alphaproteobacterial symbiont, the origin of the latter is not addressed, they are just assumed to be present in the cell that acquired a few alphaproteobacterial genes.

The kind of transition between the pre-mitochondrion (not derived from proteobacteria) and the mitochondrion (derived from an alphaproteobacterium) that Gray envisages entails several *ad hoc* components, such as precise retargeting of all the proteins that a mitochondrion needs from the pre-mitochondrion to the mitochondrion. During the origin of plastids, the plant mitochondrion, which had its protein import apparatus in place, did not become transformed so as to become green and photosynthetic, the two compartments remained distinct, rather than showing a tendency to merge, and the plastid ended up having its own import machinery, which arose

Figure 4



Mitochondria and related organelles all have a single origin. **(a)** Different types of mitochondria-related organelles (e.g. mitosomes or hydrogenosomes) can be found in different taxa of all eukaryotic super groups, such as the Amoebozoa and the Alveolata. **(b)** The last eukaryotic common ancestor (LECA) contained a 'universal' facultative anaerobic mitochondrion of alphaproteobacterial origin and the different types of mitochondria-related organelles evolved subsequently from the common ancestor, and depending on the ecological niche the host colonized.

independently of that in mitochondria. Gray's theory is an excellent example of a thoughtful theory that is designed to explain unexpected branches in trees, but not to explain the similarity of mitochondria to bacteria. As Gray [101•] points out, it has quite a lot in common with autogenous theories for the origins of organelles, which were also not designed to explain the similarity of mitochondria to bacteria, rather they were designed to explain the presence of DNA in plastids and mitochondria [14–16].

Conclusion

Endosymbiotic theory for the origin of organelles is still by far the best tool we have to explain why chloroplasts and mitochondria are so similar to free living bacteria. Alternatives to endosymbiotic theory often share several important, but unstated assumptions: they start with the premise that endosymbiotic theory somewhere stated or predicted that *all* genes that the plant lineage acquired from cyanobacteria need to branch with present-day cyanobacterial homologues, and that *all* genes that eukaryotes acquired from mitochondria need to branch with present-day purple non-sulphur bacterial (or alphaproteobacterial) homologues in phylogenetic trees. Using that lever, one can pry loose a corollary: all genes that do not fulfil those criteria were acquired from other sources. The ensuing procedure for identifying the donor is then simple: we assume that the prokaryotic homologue and the prokaryotic rRNA gene (the basis of naming prokaryotic groups) of the genome within which the homologue of the eukaryotic gene resides, have remained linked — within the same chromosome — from the time that the gene was donated (for plastid and mitochondrial origins, about a third of Earth's history ago) until the present, and we assume that the procedure of inferring gene phylogeny is error-free. Using such assumptions, whether

explicitly stated or not, one can infer that a gene X was donated by organism Y. OK, but to be fair then the same logic needs to apply to all genes, in which case the practice of inferring gene origins directly from trees quickly turns into an affair of one endosymbiont per gene and, if we think it through in full, we would end up assuming that all prokaryotic genes having eukaryotic homologues have remained resident in the same prokaryotic chromosome together with their species name-giving rRNA for the last 1–2 billion years. Now recall that endosymbiotic theory is a lot older than the practise of building gene trees. Alternatively, endosymbiotic theory is fine but it needs to be better integrated into a modern world of microbial genomics, one where we know that the pan-genomes of prokaryotic species are much larger than any individual's genome, and where lateral gene transfer is known to transport genes across chromosomes with little respect for species (or other taxonomic) borders. In summary, we probably need to keep our expectations more relaxed when it comes to the phylogenetic behaviour of genes that eukaryotes acquired from plastids and mitochondria. If we do that, endosymbiotic theory explains a lot as it is.

Acknowledgements

We thank Andrzej Bodyl and John W. Stiller for their constructive comments on the manuscript. This work was funded by the European Research Council (grant no. 232975 to WFM) and by the Deutsche Forschungsgemeinschaft (grant no. GO1825/4-1 to SBG). CK thanks the Deutscher Akademischer Austauschdienst for a PhD stipend.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Mereschkowsky C: **Über Natur und Ursprung der Chromatophoren im Pflanzenreiche**. *Biol Centralbl* 1905, **25**:593-604 (English translation in Martin W, Kowallik KV, *Eur J Phycol* 1999, **34**:287–295).
2. Wallin IE: *Symbionticism and the Origin of Species*. London: Bailliere, Tindall and Cox; 1927, 171.
3. Sapp J: **The dynamics of symbiosis: an historical overview**. *Can J Bot* 2004, **82**:1046-1056.
4. Schnepf E: **Zur Feinstruktur von *Geosiphon pyriforme.* Ein Versuch zur Deutung cytoplasmatischer Membranen und Kompartimente**. *Arch Mikrobiol* 1964, **49**:112-131.
5. John P, Whatley FR: ***Paracoccus denitrificans* and the evolutionary origin of the mitochondrion**. *Nature* 1975, **254**:495-498.
6. Glöckner G, Rosenthal A, Valentin K: **The structure and gene repertoire of an ancient red algal plastid genome**. *J Mol Evol* 2000, **51**:382-390.
7. Burger G, Gray MW, Forget L, Lang BF: **Strikingly bacteria-like and gene-rich mitochondrial genomes throughout jakobid protists**. *Genome Biol Evol* 2013, **5**:418-438.
8. Kleine T, Maier UG, Leister D: **DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis**. *Annu Rev Plant Biol* 2009, **60**:115-138.
9. Meisinger C, Sickmann A, Pfanner N: **The mitochondrial proteome: from inventory to function**. *Cell* 2008, **134**:22-24.
10. Martin W, Herrmann RG: **Gene transfer from organelles to the nucleus: how much, what happens, and why?** *Plant Physiol* 1998, **118**:9-17.
11. Allen JF: **Control of gene-expression by redox potential and the requirement for chloroplast and mitochondrial genomes**. *J Theor Biol* 1993, **165**:609-631.
12. Allen JF: **The function of genomes in bioenergetic organelles**. *Phil Trans R Soc Lond B: Biol Sci* 2003, **358**:19-37.
13. Maier UG, Zauner S, Woehle C, Bolte K, Hempel F, Allen JF, Martin WF: **Massively convergent evolution for ribosomal protein gene content in plastid and mitochondrial genomes**. *Genome Biol Evol* 2013, **5**:2318-2329.
14. Bogorad L: **Evolution of organelles and eukaryotic genomes**. *Science* 1975, **188**:891-898.
15. Cavalier-Smith T: **The origin of nuclei and of eukaryotic cells**. *Nature* 1975, **256**:463-468.
16. Raff RA, Mahler HR: **The non symbiotic origin of mitochondria**. *Science* 1972, **177**:575-582.
17. Bonen L, Doolittle WF: **Prokaryotic nature of red algal chloroplasts**. *Proc Natl Acad Sci U S A* 1975, **72**:2310-2314.
18. Farrely F, Butow RA: **Rearranged mitochondrial genes in the yeast nuclear genome**. *Nature* 1983, **301**:296-301.
19. Giovannoni S, Turner S, Olsen G, Barns S, Lane D, Pace N: **Evolutionary relationships among cyanobacteria and green chloroplasts**. *J Bacteriol* 1988, **170**:3584-3592.
20. Yang D, Oyaizu Y, Oyaizu H, Olsen GJ, Woese CR: **Mitochondrial origins**. *Proc Natl Acad Sci U S A* 1985, **82**:4443-4447.
21. Stackebrandt E, Murray RGE, Trüper HG: ***Proteobacteria* classis nov., a name for the phylogenetic taxon that includes the "purple bacteria and their relatives"**. *Int J Syst Bacteriol* 1988, **38**:321-325.
22. Dolezal P, Likic V, Tachezy J, Lithgow T: **Evolution of the molecular machines for protein import into mitochondria**. *Science* 2006, **313**:314-318.
23. McFadden GI, van Dooren GG: **Evolution: red algal genome affirms a common origin of all plastids**. *Curr Biol* 2004, **14**:R514-R516.
24. • Zarsky V, Tachezy J, Dolezal P: **Tom40 is likely common to all mitochondria[SINGLE]**. *Curr Biol* 2012, **22**:R479-R481.
Proposes that one TOM40 unites all extant eukaryotic mitochondria and organelles of mitochondrial origin.
25. Shiflett AM, Johnson PJ: **Mitochondrion-related organelles in eukaryotic protists**. *Annu Rev Microbiol* 2010, **64**:409-429.
26. Bullmann L, Haarmann R, Mirus O, Bredemeier R, Hempel F, Maier UG, Schleiff E: **Filling the gap, evolutionarily conserved Omp85 in plastids of chromalveolates**. *J Biol Chem* 2010, **285**:6848-6856.
27. Shi LX, Theg SM: **The chloroplast protein import system: from algae to trees**. *Biochim Biophys Acta* 2013, **1833**:314-331.
28. Howe CJ, Barbrook AC, Nisbet RER, Lockhart PJ, Larkum AWD: **The origin of plastids**. *Philos Trans R Soc Lond B: Biol Sci* 2008, **363**:2678-2685.
29. Parfrey LW, Lahr DJG, Knoll AH, Katz LA: **Estimating the timing of early eukaryotic diversification with multigene molecular clocks**. *Proc Natl Acad Sci U S A* 2011, **108**:13624-13629.
30. Gould SB, Waller RF, McFadden GI: **Plastid evolution**. *Annu Rev Plant Biol* 2008, **59**:491-517.
31. Lane CE, Archibald JM: **The eukaryotic tree of life: endosymbiosis takes its TOL**. *Trends Ecol Evol* 2008, **23**:268-275.
32. Kowallik KV: **Evolution durch genomische Kombination**. In *Gott oder Darwin.* Edited by Klose J, Oehler J. Berlin, Germany: Springer Verlag; 2008:141-157.
33. Gibbs SP: **The chloroplasts of some algal groups may have evolved from endosymbiotic eukaryotic algae**. *Ann N Y Acad Sci* 1981, **361**:193-208.

34. Delwiche CF: **Tracing the thread of plastid diversity through the tapestry of life**. *Am Nat* 1999, **154**:S164-S177.

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

36. • Petersen J, Ludewig AK, Michael V, Bunk B, Jarek M, Baurain D, Brinkmann H: ***Chromera velia*, endosymbioses and the rhodoplex hypothesis–plastid evolution in cryptophytes, alveolates, stramenopiles, and haptophytes (CASH lineages)[SINGLE]**. *Genome Biol Evol* 2014, **6**:666-684.
A treatise of competing hypotheses for the evolution of red complex plastids. Proposes the 'rhodoplex hypothesis' as a possible alternative to the chromalveolate hypothesis.

37. Bodyl A, Stiller JW, Mackiewicz P: **Chromalveolate plastids: direct descent or multiple endosymbioses?** *Trends Ecol Evol* 2009, **24**:119-121.

38. Sommer MS, Gould SB, Lehmann P, Gruber A, Przyborski JM, Maier UG: **Der1-mediated pre-protein import into the periplastid compartment of chromalveolates?** *Mol Biol Evol* 2007, **24**:918-928.

39. Bolte K, Bullmann L, Hempel F, Bozarth A, Zauner S, Maier UG: **Protein targeting into secondary plastids**. *J Eukaryot Microbiol* 2009, **56**:9-15.

40. Smith MH, Ploegh HL, Weissman JS: **Road to ruin: targeting proteins for degradation in the endoplasmic reticulum**. *Science* 2011, **334**:1086-1090.

41. Gould SB: **Ariadne's thread: Guiding a precursor protein across five membranes in a cryptophyte**. *J Phycol* 2008, **44**:23-26.

42. Felsner G, Sommer MS, Gruenheit N, Hempel F, Moog D, Zauner S, Martin W, Maier UG: **ERAD components in organisms with complex red plastids suggest recruitment of a preexisting protein transport pathway for the periplastid membrane**. *Genome Biol Evol* 2011, **3**:140-150.

43. •• Stork S, Moog D, Przyborski JM, Wilhelmi I, Zauner S, Maier UG: **Distribution of the SELMA translocon in secondary plastids of red algal origin and predicted uncoupling of ubiquitin-dependent translocation from degradation[DOUBLE]**. *Eukaryot Cell* 2012, **11**:1472-1481.
The most comprehensive phylogenetic analysis of the SELMA components, with incisive perspectives on how the function evolved within its constituent components.

44. Cavalier-Smith T: **A six kingdom classification and a unified phylogeny**. In *Endocytobiology II*. Edited by Schwemmler W, Schenk HEA. Berlin, Germany: De Gruyter; 1983:1027-1034.

45. Hirakawa Y, Burki F, Keeling PJ: **Genome-based reconstruction of the protein import machinery in the secondary plastid of a chlorarachniophyte alga**. *Eukaryot Cell* 2012, **11**:324-333.

46. Stöbe B, Kowallik KV: **Gene-cluster analysis in chloroplast genomics**. *Trends Genet* 1999, **9**:344-347.

47. Moustafa A, Beszteri B, Maier UG, Bowler C, Valentin K, Bhattacharya D: **Genomic footprints of a cryptic plastid endosymbiosis in diatoms**. *Science* 2009, **324**:1724-1726.

48. Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D: **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* 2002, **99**:12246-12251.

49. Timmis JN, Ayliffe MA, Huang CY, Martin W: **Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes**. *Nat Rev Genet* 2004, **5**:123-135.

50. Woehle C, Dagan T, Martin W, Gould SB: **Red and problematic green phylogenetic signals among thousands of nuclear genes from the photosynthetic and apicomplexa-related *Chromera velia***. *Genome Biol Evol* 2011, **3**:1220-1230.

51. •• Curtis BA, Tanifuji G, Burki F, Gruber A, Irimia M, Maruyama S, Arias MC, Ball SG, Gile GH, Hirakawa Y *et al.*: **Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs[DOUBLE]**. *Nature* 2012, **492**:59-65.
Reports the genome sequences of two organisms harbouring complex plastids and a nucleomorph, elegantly comparing the independent evolution of complex cell types in both lineages. Suggests an evolutionary rationale behind the retention of nucleomorphs.

52. Dagan T, Martin W: **Microbiology. Seeing green and red in diatom genomes**. *Science* 2009, **324**:1651-1652.

53. Goremykin VV, Hansmann S, Martin WF: **Evolutionary analysis of 58 proteins encoded in six completely sequenced chloroplast genomes: Revised molecular estimates of two seed plant divergence times**. *Plant Syst Evol* 1997, **206**:337-351.

54. Martin W, Stoebe B, Goremykin V, Hapsmann S, Hasegawa M, Kowallik KV: **Gene transfer to the nucleus and the evolution of chloroplasts**. *Nature* 1998, **393**:162-165.

55. Lockhart PJ, Howe CJ, Barbrook AC, Larkum AWD, Penny D: **Spectral analysis, systematic bias, and the evolution of chloroplasts**. *Mol Biol Evol* 1999, **16**:573-576.

56. Dagan T, Martin W: **Ancestral genome sizes specify the minimum rate of lateral gene transfer during prokaryote evolution**. *Proc Natl Acad Sci U S A* 2007, **104**:870-875.

57. Stephens RS, Kalman S, Lammel C, Fan J, Marathe R, Aravind L, Mitchell W, Olinger L, Tatusov RL, Zhao Q *et al.*: **Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis***. *Science* 1998, **282**:754-759.

58. Linka N, Hurka H, Lang BF, Burger G, Winkler HH, Stamme C, Urbany C, Seil I, Kusch J, Neuhaus HE: **Phylogenetic relationships of non-mitochondrial nucleotide transport proteins in bacteria and eukaryotes**. *Gene* 2003, **306**:27-35.

59. Greub G, Raoult D: **History of the ADP/ATP-translocase-encoding gene, a parasitism gene transferred from a Chlamydiales ancestor to plants 1 billion years ago**. *Appl Environ Microbiol* 2003, **69**:5530-5535.

60. Royo J, Gímez E, Hueros G: **CMP-KDO synthetase: a plant gene borrowed from Gram-negative eubacteria**. *Trends Genet* 2000, **16**:432-433.

61. Schmitz-Esser S, Linka N, Collingro A, Beier CL, Neuhaus HE, Wagner M, Horn M: **ATP/ADP translocases: a common feature of obligate intracellular amoebal symbionts related to chlamydiae and rickettsiae**. *J Bacteriol* 2004, **186**:683-691.

62. Brinkman FSL, Blanchard JL, Cherkasov A, Av-Gay Y, Brunham RC, Fernandez RC, Finlay BB, Otto SP, Ouellette BFF, Keeling PJ *et al.*: **Evidence that plant-like genes in *Chlamydia* species reflect an ancestral relationship between Chlamydiaceae, cyanobacteria, and the chloroplast**. *Genome Res* 2002, **12**:1159-1167.

63. Amiri H, Karlberg O, Andersson SGE: **Deep origin of plastid/parasite ATP/ADP translocases**. *J Mol Evol* 2003, **56**:137-150.

64. Facchinelli F, Colleoni C, Ball SG, Weber AP: **Chlamydia, cyanobiont, or host: who was on top in the ménage à trois?** *Trends Plant Sci* 2013, **18**:673-679.

65. Subtil A, Collingro A, Horn M: **Tracing the primordial Chlamydiae: extinct parasites of plants?** *Trends Plant Sci* 2014, **19**:36-43.

66. Huang J, Gogarten JP: **Did an ancient chlamydial endosymbiosis facilitate the establishment of primary plastids?** *Genome Biol* 2007, **8**:R99.

67. Moustafa A, Reyes-Prieto A, Bhattacharya D: **Chlamydiae has contributed at least 55 genes to plantae with predominantly plastid functions**. *PLoS ONE* 2008, **3**:e2205.

68. Cencil U, Nitschke F, Steup M, Minassian BA, Colleoni C, Ball SG: **Transition from glycogen to starch metabolism in Archaeplastida**. *Trends Plant Sci* 2014, **19**:18-28.

69. Kneip C, Lockhart P, Voss C, Maier UG: **Nitrogen fixation in eukaryotes — new models for symbiosis**. *BMC Evol Biol* 2007, **7**:55.

70. Raven JA: **Evolution of cyanobacterial symbioses**. In *Cyanobacteria in Symbiosis.* Edited by Rai AN, Bergman B, Rasmussen U. Dordrecht, The Netherlands: Kluwer Academic Publishers; 2002:326-346.

71. Dagan T, Roettger M, Stucken K, Landan G, Koch R, Major P, Gould SB, Goremykin VV, Rippka R, de Marsac NT *et al.*: **Genomes of stigonematalean cyanobacteria (subsection V) and the evolution of oxygenic photosynthesis from prokaryotes to plastids**. *Genome Biol Evol* 2013, **5**:31-44.

72. Becker B, Hoef-Emden K, Melkonian M: **Chlamydial genes shed light on the evolution of photoautotrophic eukaryotes**. *BMC Evol Biol* 2008, **8**:203.

73. Lange BM, Rujan T, Martin W, Croteau R: **Isoprenoid biosynthesis: The evolution of two ancient and distinct pathways across genomes**. *Proc Natl Acad Sci U S A* 2000, **97**:13172-13177.

74. Rujan T, Martin W: **How many genes in *Arabidopsis* come from cyanobacteria? An estimate from 386 protein phylogenies**. *Trends Genet* 2001, **17**:113-120.

75. Martin WF, Roettger M, Kloesges T, Thiergart T, Woehle C, Gould SB, Dagan T: **Modern endosymbiotic theory: getting lateral gene transfer into the equation**. *J Endocyt Cell Res* 2012, **23**:1-5.

76. Thiergart T, Landan G, Schenk M, Dagan T, Martin WF: **An evolutionary network of genes present in the eukaryote common ancestor polls genomes on eukaryotic and mitochondrial origin**. *Genome Biol Evol* 2012, **4**:466-485.

77. Stiller JW: **Experimental design and statistical rigor in phylogenomics of horizontal and endosymbiotic gene transfer**. *BMC Evol Biol* 2011, **11**:259.

78. Ochman H, Lawrence JG, Groisman EA: **Lateral gene transfer and the nature of bacterial innovation**. *Nature* 2000, **405**:299-304.

79. Wolf YI, Koonin EV: **Genome reduction as the dominant mode of evolution**. *Bioessays* 2013, **35**:829-837.

80. Suzuki K, Miyagishima S: **Eukaryotic and eubacterial contributions to the establishment of plastid proteome estimated by large-scale phylogenetic analyses**. *Mol Biol Evol* 2010, **27**:581-590.

81. •• Beck C, Knoop H, Axmann IM, Steuer R: **The diversity of cyanobacterial metabolism: genome analysis of multiple phototrophic microorganisms[DOUBLE]**. *BMC Genomics* 2012, **13**:56.

Reports on the nature of the cyanobacterial pangenome showing that the number of newly identified cyanobacterial genes continues to increase with each new cyanobacterial genome sequenced.

82. Lukjancenko O, Wassenaar TM, Ussery DW: **Comparison of 61 sequenced *Escherichia coli* genomes**. *Microb Ecol* 2010, **60**:708-720.

83. •• McInerney JO, O'Connell M, Pisani D: **The hybrid nature of the eukaryota and a consilient view of life on Earth[DOUBLE]**. *Nat Rev Microbiol* 2014, **12**:449-455.

An insightful perspective on the chimaeric nature of eukaryotes as true genomic hybrids of an archaeal host and a bacterial symbiont: the ancestor of mitochondria.

84. •• Williams TA, Foster PG, Cox CJ, Embley TM: **An archaeal origin of eukaryotes supports only two primary domains of life[DOUBLE]**. *Nature* 2013, **504**:231-236.

An incisive overview of current theories for eukaryote origins and tests of their predictions using phylogenetic methods. Presents strong evidence that the host lineage for the origin of mitochondria stems from within the archaea, not as the sister to archaea as in the traditional 'three domains' rRNA tree.

85. Williams TA, Embley M: **Archaeal ''dark matter'' and the origin of eukaryotes**. *Genome Biol Evol* 2014, **6**:474-481.

86. • Guy L, Saw JH, Ettema TJG: **The archaeal legacy of eukaryotes: a phylogenomic perspective[SINGLE]**. *Cold Spring Harb Perspect Biol* 2014, **6** http://dx.doi.org/10.1101/cshperspect.a016022.

An important overview of the discovery and occurrence of eukaryotic cytoskeletal components in certain groups of archaea currently called the TACK superphylum.

87. Lane N, Martin W: **The energetics of genome complexity**. *Nature* 2010, **467**:929-934.

88. Thrash JC, Boyd A, Huggett MJ, Grote J, Carini P, Yoder RJ, Robbertse B, Spatafora JW, Rappe MS, Giovannoni SJ: **Phylogenomic evidence for a common ancestor of mitochondria and the SAR11 clade**. *Sci Rep* 2011, **1**:13.

89. Brindefalk B, Ettema TJ, Viklund J, Thollesson M, Andersson SG: **A phylometagenomic exploration of oceanic alphaproteobacteria reveals mitochondrial relatives unrelated to the SAR11 clade**. *PLoS ONE* 2011, **6**:e24457.

90. Esser C, Ahmadinejad N, Wiegand C, Rotte C, Sebastiani F, Gelius-Dietrich G, Henze K, Kretschmann E, Richly E, Leister D *et al.*: **A genome phylogeny for mitochondria among alpha-proteobacteria and a predominantly eubacterial ancestry of yeast nuclear genes**. *Mol Biol Evol* 2004, **21**:1643-1660.

91. Abhishek A, Bavishi A, Bavishi A, Choudhary M: **Bacterial genome chimaerism and the origin of mitochondria**. *Can J Mircobiol* 2011, **57**:49-61.

92. Georgiades K, Raoult D: **The rhizome of *Reclinomonas americana*, *Homo sapiens*, *Pediculus humanus* and *Saccharomyces cerevisiae* mitochondria**. *Biol Direct* 2011, **6**:55.

93. Atteia A, Adrait A, Brugière S, van Lis R, Tardif M, Deusch O, Dagan T, Kuhn L, Gontero B, Martin W *et al.*: **A proteomic survey of *Chlamydomonas reinhardtii* mitochondria sheds new light on the metabolic plasticity of the organelle and on the nature of the a-proteobacterial mitochondrial ancestor**. *Mol Biol Evol* 2009, **29**:1533-1548.

94. Rochette NC, Brochier-Armanet C, Gouy M: **Phylogenomic test of the hypotheses for the evolutionary origin of eukaryotes**. *Mol Biol Evol* 2014, **31**:832-845.

95. Le PT, Pontrarotti P, Raoult D: **Alphaproteobacteria species as a source and target of lateral sequence transfers**. *Trends Microbiol* 2014, **22**:147-156.

96. •• Degli Esposti M, Chouaia B, Comandatore F, Crotti E, Sassera D, Lievens PM, Bandi C: **Evolution of mitochondria reconstructed from the energy metabolism of living bacteria[DOUBLE]**. *PLoS ONE* 2014, **9**:e96566.

A comprehensive comparative survey of membrane bioenergetics in alphaproteobacteria that uncovers new and intruiging links in the evolutionary history of mitochondria.

97. Cavanaugh CM, Wirsen CO, Jannasch HW: **Evidence for methylotrophic symbionts in a hydrothermal vent mussel (bivalvia: mytilidae) from the mid-atlantic ridge**. *Appl Environ Microbiol* 1992, **58**:3799-3803.

98. • Müller M *et al.*: **Biochemistry and evolution of anaerobic energy metabolism in eukaryotes[SINGLE]**. *Microbiol Mol Biol Rev* 2012, **76**:444-495.

A comprehensive review of energy metabolism in eukaryotic anaerobes focusing on the role of mitochondria in well-studied anaerobic model organisms.

99. Stairs CW, Eme L, Brown MW, Mutsaers C, Susko E, Dellaire G, Soanes DM, van der Giezen M, Roger AJ: **A SUF Fe-S cluster biogenesis system in the mitochondrion-related organelles of the anaerobic protist *Pygsuia***. *Curr Biol* 2014, **24**:1176-1186.

100. Hug LA, Stechmann A, Roger AJ: **Phylogenetic distributions and histories of proteins involved in anaerobic pyruvate metabolism in eukaryotes**. *Mol Biol Evol* 2010, **27**:311-324.

101 • Gray MW: **The pre-endosymbiont hypothesis: a new perspective on the origin and evolution of mitochondria[SINGLE]**. *Cold Spring Harb Perspect Biol* 2014, **6** http://dx.doi.org/10.1101/cshperspect.a016097.

A proposal to modify the endosymbiont hypothesis for the origin of mitochondria through the suggestion that a compartment similar to the mitochondrion already existed in the host that acquired mitochondrial endosymbiont, whereby this preexisting compartment contained all the proteins of modern mitochondria that do not branch with alphaproteobacterial homolgues in single gene trees.

102. Melkonian M: **Phylogeny of photosynthetic protists and their plastids**. *Verh Dtsch Zool Ges* 1996, **89**:71-96.