ACADÉMIE
DES SCIENCES
INSTITUT DE FRANCE

Comptes Rendus
Biologies

Steven Ball, Denis Baurain, Marie Leleu, Ingrid Lafontaine, Ugo Cenci, Christophe Colleoni, Olivier Vallon, Gilbert Greub, Andreas Weber and Debashish Bhattacharya

Obligate intracellular bacterial pathogens as major players in the metabolic integration of organelles

Volume 348 (2025), p. 107-135

Online since: 11 June 2025

https://doi.org/10.5802/crbiol.176

The Comptes Rendus. Biologies *are a member of the Mersenne Center for open scientific publishing*
www.centre-mersenne.org — e-ISSN : 1768-3238

ACADÉMIE DES SCIENCES
INSTITUT DE FRANCE

Comptes Rendus. Biologies
2025, Vol. 348, p. 107-135
https://doi.org/10.5802/crbiol.176

Review article

Obligate intracellular bacterial pathogens as major players in the metabolic integration of organelles

Steven Ball [*,a], Denis Baurain [b], Marie Leleu [a,b], Ingrid Lafontaine [c], Ugo Cenci [a], Christophe Colleoni [a], Olivier Vallon [c], Gilbert Greub [d], Andreas Weber [e] and Debashish Bhattacharya [f]

[a] Unité de Glycobiologie Structurale et Fonctionnelle UMR8576 CNRS/Université de Lille, Bâtiment C9, Cité Scientifique, 59655 Villeneuve d'Ascq, France

[b] Laboratoire de Phylogénomique des Eucaryotes Bât. B22 Quartier Vallée 1 chemin de la Vallée 4, 4000 Liège 1, Belgium

[c] Unité de Biologie du chloroplaste et perception de la lumière chez les micro-algues UMR7141 CNRS/Sorbonne Université, Institut de Biologie Physico-Chimique, 13, rue Pierre et Marie Curie, 75005 Paris, France

[d] Institut de microbiologie de l'Université de Lausanne, Département des laboratoires Bureau IMUL 02-222, Bugnon 48, CH-1011 Lausanne, Switzerland

[e] Institute for Plant Biochemistry, Heinrich-Heine-University, Geb. 26.14.02, Room 106, Universitaetsstrasse 1, D-40225 Duesseldorf, Germany

[f] Department of Biochemistry and Microbiology, Rutgers, The State University of New Jersey, Foran Hall, Room 102, 59 Dudley Road, New Brunswick, NJ 08901-8520, USA

E-mail: steven.ball@univ-lille.fr (S. G. Ball)

Abstract. Endosymbionts are very common in nature, offering multiple occasions to recapitulate events that have led to the generation of mitochondria and plastids. However, both these organelles are unique because they are thought to derive from two individual events that gave rise to all eukaryotes and the plastids in algae and plants (excluding *Paulinella chromatophora*), respectively. This review focuses on the differences and similarities existing between extant endosymbionts and the two major endosymbiont derived organelles: the mitochondria and plastids. Emphasis is put on recent developments that point to the major role of intracellular pathogens in the establishment of these organelles. We argue that metabolic integration of bacterial endosymbionts into mitochondria and plastids required an unusually high degree of preadaptation not shared by most extant endosymbionts. We propose that this was achieved by either recruiting intracellular bacterial pathogens as "helper genomes" providing needed gene products, or by selecting endosymbionts destined to become organelles directly from such obligate intracellular bacteria.

Keywords. Endosymbiosis, Mitochondria, Plastids, Chlamydia, Rickettsiales, Ménage à trois hypothesis.

Funding. CNRS, Université de Lille CNRS, ANR grants "Mathtest" (ANR-18-CE13-0027) "expendo" (ANR-14-CE11-0024) and "ménage à trois" (ANR-12-BSV2-0009), Région Hauts de France and European project Alibiotech, ANR grant "Mathtest" (ANR-18-CE13-0027).
Note. Steven Ball is the recipient of the 2022 George Morel prize.

Manuscript received 6 February 2025, accepted 26 March 2025.

[*] Corresponding author

1. Introduction: comparing extant endosymbionts with endosymbiont-derived organelles

Unlike the primary endosymbioses that gave rise to mitochondria and plastids (excluding Paulinella, see below), which are single events involving two different bacterial donors, endosymbiosis of bacteria by eukaryotes is common, particularly in invertebrates and protists. These diverse endosymbioses rely on mutualism, commensalism, or parasitic interactions. Such common endosymbionts can be distinguished from mitochondria and plastids by the apparent absence within the endosymbiont compartment of proteins that are not encoded by the residual symbiont genome itself. This correlates with the absence of a specialized protein targeting machinery that directs host nuclear-encoded proteins to the ancestral bacterial compartment or membranes. Some of the associations of extant endosymbionts are obligate, whereas others are facultative. Endosymbiont-derived organelles such as mitochondria and plastids are derived from obligate mutualistic endosymbionts, which have managed to target to the symbiont compartment proteins now encoded by the host. Obligate endosymbionts and endosymbiont-derived organelles share many properties. Upon entering symbiosis with the eukaryotic host, ancestral bacteria were in all cases isolated from their "free-living" progenitors, and unable to exchange genes with them through recombination. As a result, these genomes accumulated mutations leading to pseudogenes and, ultimately, gene losses, a phenomenon known as "Muller's ratchet" [1]. The presence of a nutrient-rich cytosol within the eukaryotic host likely further favoured the loss of all functions that are not essential for symbiosis and maintenance through the selection of symbiotic genes [2]. Some of the extant obligate endosymbiont genomes, such as that of *Nasuia deltocephalinicola*, an amino acid supplying symbiont of hemipteran insects, have thus undergone reduction. It went from a genome several Mbp in size in the hypothetical ancestral Burkholderiales β-proteobacterium to a 112 kbp relic genome with 137 predicted protein-coding genes, which is comparable to or less than what is found in endosymbiont-derived organelles [3,4]. This correlates with a significant increase in genome AT content and changes in the genetic code, traits also present in mitochondria. Hence, these endosymbionts show convergent evolution with respect to genome size and nucleotide content in mitochondria and plastids. However, unlike standard endosymbionts, organelle genome size and gene number are not robust indicators of the organelle proteome complexity. Both mitochondria and plastid symbioses have relied on the presence of electron transport systems coupled to specific complex membrane structures. The number of genes required to optimize the function or to encode components of these complex machineries vastly exceed the number of genes involved, for example, in amino acid metabolism, which form the basis of the symbiosis uniting the aforementioned insect endosymbionts to their host. To avoid the deleterious effects of Muller's ratchet-induced accumulation of mutations affecting the fitness of their genes, the organelle-generating endosymbionts relied increasingly on the host genetic system. This took the form of either the host providing host-derived functions or the host genome being the recipient of lateral gene transfers providing compromised symbiont functions. In most cases, the foreign gene source is the endosymbiont itself, thereby generating **EGTs** (endosymbiotic gene transfers). However, in a significant number of cases, the source of the gene is from another bacterial origin not directly related to the symbiont. LGTs that specifically impact the host genome at the time of organelle endosymbiosis are hereby termed **ERGTs** (Endosymbiosis-Related lateral Gene Transfer). In this context, the foreign substitute of the original symbiont gene only accounts for part of the observed ERGT signal. Equally important are ERGTs corresponding to genes not previously encoded by the symbiont, which bring novel functions required during the process of organelle-host metabolic integration. In the case of EGTs and replacement ERGTs, the now nuclear-encoded organelle-derived genes escaped Muller's ratchet, due to the population size and the gene recombination and repair systems of the host. Nevertheless, accumulation of mutations followed by pseudogenization and gene losses was the hallmark of the early phases of reductive genome evolution in organelles. In time, and in addition, both organelle genomes have evolved mechanisms that attenuate the ratchet, including, among others, an increased genome copy number and efficient recombination and repair systems [5,6].

However, the present evidence gathered from bacterial non-organellar endosymbionts and their insect host genomes suggests that if any EGT or ERGT can eventually be associated with such symbioses, these will likely be less frequent. Nevertheless, some extant endosymbionts not related to mitochondria or plastids are presently suspected [7] and even recognized to be able to display either limited [8,9] or even extensive symbiont targeting of host encoded proteins [10,11]. This further blurs the distinction between true organelles and endosymbionts. This review will focus on the ancient and recently gathered evidence pointing to an active role of intracellular bacterial pathogens in the organelle-generating endosymbioses. For a more general outlook at the process of primary endosymbiosis and alternative hypotheses, the reader is redirected to more general, excellent and comprehensive reviews [12–15].

2. Consequences of Muller's ratchet on our understanding of the initial symbiosis biochemistry

Pseudogenization followed by gene loss is very rapid, frequent, and universal in all past and present endosymbiotic events. This, in turn, limits the number of ways in which the nature of the symbiotic relationship between endosymbionts and their host can be altered. For instance, extant mitochondria and plastids are obviously providing oxidative phosphorylation and oxygenic photosynthesis, respectively, to their host; however, it has been and is proposed that other symbiotic relationships initiated the partnership. For instance, it has been proposed that the symbiosis uniting the ancestral mitochondrion to its archaea-derived host was based on the supply of molecular hydrogen to the host under anoxia [16]. The molecular supply of hydrogen by a symbiont hydrogenase can indeed form the basis of a very efficient symbiosis under anaerobiosis. We argue, however, that if such a symbiotic link drove mitochondrial endosymbiosis for any significant length of time, Muller's ratchet would have made it quickly irreversible. Indeed, the production of molecular hydrogen most often acts as an electron valve for recycling NADH during fermentation. Hydrogen production would thus occur in conditions where oxidative phosphorylation is prevented. Free-living facultative anaerobic α-proteobacteria do not respire and produce molecular hydrogen simultaneously, but they do so at distinct times under different conditions in their changing environments. Upon endosymbiosis, within a strictly anaerobic archaea-like organism, unless oxygen consumption has some symbiotic advantage (in which case, however, symbiosis would not rely solely or even chiefly on hydrogen production), Muller's ratchet would have quickly and irreversibly affected the capacity of a hydrogen-producing symbiont to switch to oxidative phosphorylation in the constant nutrient-rich but anoxic environment of the eukaryote ancestors' cytosol. Nitrogen fixation has also been proposed as a starter symbiosis for plastid endosymbiosis [17] on the grounds that nitrogen fixation is the most frequently found symbiosis involving extant cyanobacteria and protists. Yet irreversible losses of oxygenic photosynthesis are also evidenced dramatically in the extant cyanobacterial endosymbionts that fix nitrogen [18,19]. These are derived from photosynthetically competent unicellular diazotrophic cyanobacteria with large genomes (around 5 Mbp with up to 5000 ORFs). The extant endosymbionts fix nitrogen for their photosynthetic diatom or Prymnesiophyceae (haptophyte) hosts. These symbionts, called spheroid bodies within diatoms, are presently at comparable stages of genome reduction (from 1.5 to 3 Mbp genomes with approximately 1200 to 1700 deduced ORFs), due to mutations followed by deletions in several distinct metabolic pathways [20].

These gene mutations and losses encompass essential elements of ad minima PSII (haptophyte symbionts) or of both PSI and PSII (spheroid bodies) and of the Calvin cycle and respiration, among many other pathways. It is proposed that the bacteria have evolved an essentially fermentative metabolism and import several metabolites from the nutrient-rich host cytosol to ensure anoxygenic photosynthesis through PSI in haptophytes. These recently acquired endosymbionts testify to the irreversible nature imposed by Muller's ratchet on the evolution of endosymbionts. Even at the earliest stages of their evolution, these nitrogen-fixing symbionts had definitively lost the opportunity to revert to oxygenic photosynthesis and become plastids because of the loss of essential proteins of PSII.

Interestingly, in the case of *Braarudosphaera bigelowii*, a Prymnesiophyceae containing such diazotrophic endosymbionts, extensive targeting of host-encoded proteins into the symbiont seems to occur, leading to the proposal that the latter would define a true organelle: the nitroplast [21].

We argue that a facultative hydrogen-producing anaerobic α-proteobacterium would have experienced a similar evolution of its genome towards anaerobic metabolism through mutations and gene losses, precluding its evolution into an oxygen-consuming organelle, such as the mitochondrion, at a very early stage. In other words, if a bacterium engages in an endosymbiotic relationship, it had better get it right from the start. Once a particular evolutionary path is followed, there is no turning back! An important consequence of these considerations is that both the nature of the initial bacterium and that of the initial symbiosis biochemistry will irreversibly affect the outcome of endosymbiosis a fact that is overlooked or underestimated by many researchers in this field. These considerations do not mean that the initial ancestral symbiosis that subsequently led to mitochondria and plastids necessarily entailed the export of ATP through oxidative phosphorylation and carbohydrates through oxygenic photosynthesis, respectively. However, they do mean that the initial symbiotic relationship required biochemical conditions that must have necessarily relied on respiration and photosynthesis and thus conserved all the components of oxidative phosphorylation and oxygenic photosynthesis to prevent their irreversible loss through Muller's ratchet. These could indeed lead later to the evolution of ATP and photosynthate supply by mitochondria and plastids, respectively. In this respect, the respiratory consumption and protection from oxygen by respiring bacteria in the anoxic host cytosol, or the supply of oxygen to protist ancestors by photosynthetic symbionts, qualifies, among other examples, as possible initial symbiotic relationships. On the other hand, production of hydrogen or nitrogen fixation by α-proteobacterial or cyanobacterial ancestors does not, as neither requires the maintenance of oxidative phosphorylation or oxygenic photosynthesis (on the contrary, most nitrogenases and hydrogenases are highly sensitive to inhibition by oxygen) in the intracellular conditions that prevailed for the progenitors of the mitochondrion or plastid.

3. Endosymbiotic and endosymbiosis-related lateral gene transfers: the organelle solutions to Muller's ratchet

Both mitochondria and plastids have evolved sophisticated protein targeting machineries that allowed the transfer of genes present in the symbiont genome to the host nuclear genome (for reviews concerning plastids see Ballabani et al. [22], Flores-Pérez and Jarvis [23]). Interestingly, mitochondria and plastids use analogous systems based on N-terminal targeting peptides of 20 to 70 amino acids present on the proteins targeted to organelles.

These peptides are of similar amino acid composition, except for an additional segment of approximately 16 residues found on plastidial targeting peptides. Most importantly, all such organelle-targeting peptides share the ability to spontaneously form an amphiphilic α-helix when interacting with their target organelle membranes. These structural and functional properties are also shared by a subclass of Helical-Amphiphilic-Ribosomal antimicrobial (HA-RAMP) peptides, which are ubiquitous in both prokaryotes and eukaryotes. HA-RAMPs are an ancient antimicrobial defence system against invasion by intracellular pathogenic bacteria. The phagotrophic habit of ancestral eukaryotes would have required the evolution of strong antibacterial immunity very early on. Membrane permeabilization by HA-RAMPs is well documented and could explain their antibacterial activity through the unselective leakage of essential metabolites [24]. In a recent insightful paper, it was proposed that the acquisition of resistance to eukaryotic HA-RAMPs through the import of such peptides by intracellular bacteria, and their proteolytic destruction within the bacteria by a selective protease, would have formed a possible mechanism explaining the emergence of a complex and efficient protein import machinery [25]. Indeed, eukaryotic or bacterial genes (of endosymbiont or foreign bacterial origin) transferred to the eukaryote nucleus, when fortuitously located downstream of an eukaryotic HA-RAMP gene directed against such bacteria, would have resulted in the inadvertent targeting of the hybrid protein to the evolving symbiont, followed by the selective degradation of the HA-RAMP through the bacterial protease and the possible folding of the protein remnant.

Such antimicrobial peptide resistance mechanisms have been well documented in bacteria ([26–28]; reviewed in [29]). Furthermore, two out of the four suspected peptidases active during mitochondrial and plastidial endosymbiosis display a bacterial phylogenetic affinity in agreement with this hypothesis [30].

In any case, whatever the origin of these targeting machineries of mitochondria and plastids, both targeting systems currently require a distinct set of chaperones specialized in the translocation of the unfolded targeting peptide-protein to the outer membrane of the organelle [22,23], followed by import, peptide cleavage, and subsequent folding. We argue that native bacterial and host proteins have not evolved under such constraints and may have, by chance, been quite differently suited to following such a complex route of unfolding, folding, and interaction with a specific set of chaperones. This could explain why, whenever a native endosymbiont protein was poorly suited to follow this route, a foreign alternative may have been selected faster. Hence, the genes that were selected for transfer to protect the evolving organelle from the impact of Muller's ratchet became a mix of original endosymbiont and foreign genes. The origin of the gene depended on the one hand on functional constraints and, on the other hand, on the rate of endosymbiont compared to foreign DNA transfer. At first glance, these conditions would apparently favour the transfer of endosymbiont genes by endosymbiotic gene transfer (EGT). However, the origin of the selected gene equally depended on the number of mutations that must have followed the transfer, to facilitate the targeting of the protein. In many instances other than the photosynthetic machinery per se, or a few core proteins needed for replication or expression of genes encoded in the endosymbiont genome, there was little need to reproduce the original cyanobacterial protein network with all its interactions. A more integrated alternative, allowing for a significant amount of intraorganellar proteome chimerism was favoured because this happened faster and was suitable. Thus, both the EGT of the endosymbiont genes and the ERGT of foreign bacterial genes, or the redirection of the product of corresponding host genes to the plastid, allowed for the loss of the corresponding gene in the symbiont genome and achieved protection of these organelle functions from mutation, inactivation, and loss.

4. Are antimicrobial peptides at the core of non-organellar endosymbioses?

It has recently been shown that extant nitrogen-fixing endosymbionts of legumes, such as the α-proteobacteria of the Rhizobiales order (reviewed in Mergaert [31]), or the actinobacteria of the Frankia genus [32] both interact with cysteine-rich antimicrobial peptides released by the plant host. Targeting of these symbiotic peptides to symbiont membranes has been demonstrated to be required for symbiosis. These antimicrobial peptides are probably derived from the mechanisms of antibacterial host immunity described above that interfere with a number of bacterial processes, such as membrane permeability and cell division. During symbioses, at sublethal doses, the peptides limit uncontrolled cell division and induce a differentiation process that is compatible with nitrogen fixation and host cell viability. Sometimes this differentiation is cytotoxic for the symbiont, and occasionally the bacteria do not recover from these interactions (reviewed in Mergaert [31]). In the case of the *Frankia* symbiosis, it has been demonstrated that reduced nitrogen-containing molecules such as glutamine and glutamic acid are released through interaction with host-encoded peptides [32]. Hence, as with HA-RAMPs, membrane interaction with these toxic peptides provokes an increased permeability of the bacterial membranes, possibly leading to the non-specific leakage of the two amino acids. These anyhow account for a substantial proportion of the metabolites present in the Frankia cytosol during symbiosis. Therefore leakage specificity is not an issue for this symbiosis. It was recently suggested that the apparent absence or paucity of transporters on bacterial membranes of insect endosymbionts [33,34] could be explained if one assumes the existence of antibacterial peptide-mediated non-specific leakage of molecules and metabolites [31,35]. This would happen in insect endosymbioses in the small volume defined by the lumen of the host-controlled symbiosome membrane that contains the symbionts within the bacteriome. Limited by this restricted volume, this leakage would not be lethal or even detrimental to the bacteria. Hence, here again, sublethal alterations

of membrane permeability followed by non-specific leakage could be occurring. Interestingly, such a system would obviate the need for a complex protein targeting machinery to locate crucial transporters on the bacterial envelopes. We believe that this property may account for the high frequency and diversity of endosymbionts observed in both animals and other eukaryotes. We are in agreement with those who have proposed this function to account for endosymbiosis onset and function in extant non-organellar endosymbiosis such as the nitrogen-fixing *Frankia alni*/Alnus symbiosis [32] or even the few organellar-generated endosymbionts that are not derived from, nor apparently assisted by, obligate intracellular pathogens such as the Paulinella chromatophores [15,31,35]. Indeed, despite all efforts, no convincing case for the presence of metabolite transporters targeted to the chromatophore membranes could be made in Paulinella [36] leaving microbial peptide-induced membrane permeability as the sole possible mechanism for the installation of the symbiotic metabolic fluxes.

5. Are antimicrobial peptides at the core of endosymbiosis of the future organelles?

The host's demonstrated use of antimicrobial peptides to control nitrogen fixation by bacterial endosymbionts has led to the recent proposal that such interactions may offer an alternative to the requirement for transporters to explain the onset of symbiosis, in the case of the mitochondria or primary plastids [32,35,37]. It must be pointed out, however, that tinkering with membrane permeability to ensure nonspecific leakage of metabolites is highly unlikely to have benefited symbionts that heavily relied on membrane integrity for the normal function of membrane-embedded electron transport chains. These chains must have been at the core of the symbiotic relationship in the case of the future mitochondria in order to obviate their loss through Müller's ratchet. We therefore propose that, on the contrary, the use of antibacterial peptides by a majority of extant endosymbioses not involving intracellular bacterial pathogens distinguishes them from endosymbiont-derived "true" organelles. This proposal could explain the apparent contradiction between the high frequency of endosymbiosis observed in nature and the extreme rarity of primary endosymbiosis of eukaryotic organelles, as evidenced by the apparent monophyly of these events. This implies that the bacteria used as the progenitors of both the mitochondria and plastids must have had the full suite of transporters required to establish symbiosis at the very onset of the symbiotic interactions. This also distinguishes the Paulinella chromatophore from the primary plastid, and limits the frequency of primary endosymbiosis of bacteria destined to become organelles. The level of preadaptation required for the progenitors of mitochondrion or plastid had to be much higher than for many extant endosymbionts. In this respect, the bacteria that offer the highest level of preadaptation to living within eukaryotes are by far the obligatory intracellular bacterial pathogens [38].

6. The restricted diversity but high connectivity and flexibility of the obligate intracellular pathogens

Bacteria that have lost the ability to multiply out of their eukaryotic hosts are constantly evolving from facultative intracellular bacteria. Some of these evolutionary transitions are fairly recent, as exemplified by the leprosy agent *Mycobacterium leprae*, the aphid symbiont *Buchnera aphidicolae*, or the Q-fever pathogen *Coxiella burnettii*, which have evolved through genome reduction of a facultative intracellular bacterial pathogen of the order Mycobacteriales, Enterobacteriales or Legionellales, respectively, in animals [39,40]. Some of these are, however, far more ancient and define large and diverse groups of bacteria in which most members of bacterial orders are intracellular. These are defined by the α-proteobacteria Rickettsiales and the PVC supergroup member order Chlamydiales [41,42]. Representatives of these two orders can be found multiplying in hosts from the whole spectrum of the eukaryotic domain and not only within the animal kingdom, as initially thought. This makes them candidates for very ancient transitions to intracellular life, possibly as far back as, or prior to, the time of plastid and mitochondrial endosymbioses. While reductive genome evolution has been at work at some point to explain their relatively simple genomes, the most ancestrally derived members of Rickettsiales or Chlamydiales still encode over 2000 protein-coding candidate genes, a respectable number compared

to the amino acid or vitamin supplying insect bacterial symbionts [41,43]. The mitochondria have often been phylogenetically found closely related to the Rickettsiales order among α-proteobacteria. What has been and still is controversial is their precise relative position, nested within [43] or sisters to Rickettsiales [44]. This is of importance, since a nested topology would strongly suggest that mitochondria evolved from an ancestral bacterium that already had an obligate intracellular lifestyle. Apparently, the latest evidence points at mitochondria being, at best, sisters of Rickettsiales, with one study questioning the sisterhood [45] and another placing mitochondria only as sisters of α-proteobacteria as a whole [46]. For the time being, the prospects of clarifying the origin of mitochondria through phylogenetics alone seem remote. However, because of the considerations summarized in this review, we do not view the close phylogenetic placement of mitochondria relative to the Rickettsiales as coincidental.

At variance with the non-Wolbachia insect endosymbionts with highly reduced genomes discussed above, obligate intracellular pathogens have retained the potential to infect eukaryotes and multiply by horizontal transmission, often implying additional selection for the retention of active genes. Furthermore, the life cycles of these pathogens imply extracellular infection phases and transmission to novel hosts, making them less isolated and less subjected to the deleterious effects of Muller's ratchet than obligate non-infectious endosymbionts. Nevertheless, obligate intracellular pathogens have in common the loss of several metabolic steps, in particular biosynthetic pathways, suggesting that they import of a vaster array of metabolites than facultative, intracellular pathogens [40,41]. The most ancestral Rickettsiales lineages are now suggested to have a facultative intracellular or even an ectosymbiotic lifestyle [47]. Their genomes nevertheless testify to the presence of multiple opportunities for symbiotic or pathogenic biotic interactions with their hosts. This suggests a higher degree of connectivity to the host with an additional number of transporters, most of which remain to be characterized in Rickettsiales [48]. In addition, their obligate or facultative intracellular lifestyles implies the presence of active resistance mechanisms against the very aggressive antibacterial immunity of the host. Moreover, they share with facultative intracellular pathogens the secretion into the host cytosol of many and probably more effector proteins through sophisticated secretion machineries that facilitate their active multiplication or maintenance [48].

Finally, and most importantly, as do temperate phages, they often display both a "productive" pathogenic cycle leading to the massive production of infectious bacteria through lysis and a temperate cycle where the bacteria coordinate their replication and maintenance with that of their host to various degrees. This temperate phase can be accompanied or not by the production of infectious bacteria through budding. This state, also called "bacterial persistence", is reversible and the pathogen cycle can resume when the environmental conditions allow it [49]. Prophages can also be considered to be in a stage of persistence. Prophages are rarely neutral: in the few cases where this has been studied, they seem to provide a significant selective advantage to the bacterial clone hosting them. The best described fitness advantage is the resistance to infection by both virulent (productive) phages of the same and different types (reviewed in Bondy-Denomy and Davidson [50]). Curiously, a similar situation has been recently described for the infection of *Acanthamoeba castellanii* by facultative intracellular bacteria such as *Legionella pneumophila* [51]. Pre-infection of the amoeba by persistent forms of the Chlamydiales obligatory intracellular pathogens *Neochlamydia* or *Protochlamydia amoebophila* protects the host against superinfection by pathogenic (lytic) forms of Legionella. The notions of pathogen/symbiont persistent or temperate phases are particularly blurred in these organisms, and a specific bacterium can easily switch from one status to the other, depending on the host's nature. Temperate obligate intracellular bacteria have often evolved sophisticated means by which they can spread both horizontally and vertically in their hosts. For instance, Wolbachia, a Rickettsiales symbiont with over a thousand genes, has evolved several means by which they can spread in the germ line of insect populations through sophisticated manipulations of their sexual life cycle (for review see Landmann [52]). They can induce feminization, cytoplasmic incompatibility, parthenogenesis, male killing, and various types of other manipulations, thanks to which they have achieved the status of the most abundant symbiont on our planet [52]. Their presence can lead to both parasitic

or facultative and obligate mutualistic interactions depending on the benefits for the host of such manipulations. Transition from parasitic symbiosis to obligatory mutualism is also common and some Wolbachia strains have been shown to supply essential vitamins to their host [53]. Conversely, transition from symbiosis to a full-fledged pathogen is equally observed and can be reproduced experimentally [54]. At the time of mitochondrial endosymbiosis, a highly connected parasitic persistent Rickettsiales-like symbiont would have scored very high on the preadaptation scale with respect to the yet to be developed mitochondrial endosymbiosis.

The list of preadaptations of Rickettsiales to primary endosymbiosis is particularly long and has been reviewed by Wang and Wu [55]. Suffice it to say here that they are able to resist to the host antibacterial immunity and yet replicate in a persistent fashion, are able to rapidly invade populations of hosts vertically and horizontally, have a rich suite of transporters although many remain to be characterized, and are perfectly adapted to the intracellular environment. The anciently diverged Rickettsiales are also sufficiently gene rich to offer many distinct routes to reductive evolution that can lead to the evolution of a true novel organelle. They are also somewhat generally resistant to Muller's ratchet, provided they do not go too far in mutualistic interactions and keep on producing infectious particles. These advantages were also shared with Chlamydiales, but unlike Rickettsiales the latter do not harbour the complete set of genes needed for oxidative phosphorylation, which was quite obviously at the core of mitochondrial endosymbiosis.

7. The onset of mitochondrial endosymbiosis: lessons from Legionella and Neorickettsia

Eukaryogenesis, the process by which the first eukaryotes emerged from the archaea, is a highly controversial issue, and multiple conflicting hypotheses have been proposed to describe the evolution of the first eukaryotes [56]. Nevertheless, a consensus is emerging for a rather late occurrence of mitochondrial endosymbiosis on the way to LECA (Last Common Eukaryotic Ancestor) [57].

With such unknowns, proposing the nature of the primordial symbiosis biochemistry uniting the future mitochondrial and archaeal relatives, and deducing the implementation of this pathway during eukaryogenesis, is a speculative and perilous exercise. We may not have to do this if we simply hypothesize that the ancestor of the mitochondrion was an ancestral Rickettsiales symbiont/pathogen that had already managed to spread in or at the surface of the host ancestor populations, through both parasitic and facultative mutualistic interactions. We may not even have to necessarily imagine that this organism was an obligate intracellular symbiont: it might very well have been a facultative endosymbiont or a tightly associated ectosymbiont displaying many critical biotic interactions with the eukaryote ancestors. This is indeed suggested by the recent study of several ancestrally diverged Rickettsiales lineages, which suggest the presence of multiple transitions to the obligate intracellular lifestyle [47,58]. We suggest that a transition to an obligate intracellular life was selected on ATP transport relying on MCFs (Mitochondrial Carrier Family) exchange of host cytosolic ADP for bacterial endosymbiont generated ATP rather than the reverse. Indeed, the ATP import protein common to most Chlamydiales and Rickettsiales catalyses the unidirectional import of ATP from the host, and would have driven the bacteria into parasitism rather than symbiosis. If the host was still an oxygen-sensitive anaerobe, one can propose protection from oxygen by its respiratory symbiotic consumption as a possible initial symbiotic interaction. The loss of glycolysis-mediated fermentation by many ancestral Rickettsiales is in line with such biotic interactions.

As hypothesized by FA Wollman [25], the evolution of resistance to the host antimicrobial peptides (HA-RAMPs) by the progenitor of mitochondria would have further preadapted this lineage to primary endosymbiosis. Other bacteria or the host would have evolved distinct antibacterial peptides or proteins to tame or compete with this symbiont. Among such proteins, the MCF ADP/ATP transporter (hereafter called MCF for simplicity) can be viewed as antibacterial proteins responsible for ATP leakage of potentially pathogenic intracellular bacteria. The evolution of such transporters is often believed to have signalled the onset of mitochondrial endosymbiosis (for review, see Haferkamp and Schmitz-Esser [59]); they would have driven or consolidated the ancestor of mitochondria into a symbiotic relationship. This led to the assumption that MCF was necessarily of host origin, which

immediately raised the question of protein targeting. Novel findings bring this view into question. Quite recently, *Legionella pneumophila* was demonstrated to harbour a *bona fide* MCF effector protein secreted in an unfolded fashion by the type IV secretion system, affording for its secretion through three Legionella-controlled membranes [60]. The highly hydrophobic protein is thought to interact in the cytosol with the chaperones responsible for protein targeting to the mitochondrial inner membrane. Expressed in yeast, the Legionella MCF interfered with normal mitochondrial ATP export in particular genetic backgrounds involving mutations of the yeast MCF ATP transport complement. In addition, the MCF was active for ATP transport on its own. The rationale for the evolution of such a function in Legionella remains unknown. Interestingly, two out of the five candidate bacterial MCFs identified in the Legionella study were reported in Neorikettsia, a *bona fide* Rickettsiales [60]. One could argue that specific ATP export is very unlikely to have originated in the ancestor of the mitochondria prior to endosymbiosis. Nevertheless, toxins that induce specific ATP leakage are known, for instance, to be encoded by the *E. coli* chromosome and not only on plasmids [61]. It has been recently demonstrated that induction of the synthesis of such a toxin occurs under specific stress conditions, resulting in attenuated growth and multidrug resistance of pathogens normally sensitive to antibiotics. Evidently, the toxin is cytotoxic for *E. coli* but allows for the recovery of viable "persisters" in the treated population. This gene, however, does not encode an MCF and, most importantly, affects respiration. Hence, these novel findings concerning intracellular bacteria may suggest that, after all, the mitochondrial carriers could have originated in bacteria, either in the form of effectors designed to affect competitors in an intracellular environment, or in the form of a sophisticated chromosomal toxin–antitoxin system designed to prompt attenuated growth. In any case, whatever the origin of MCF, permanent, yet highly specific and controlled leakage of ATP signals, in our view, the onset of mitochondrial endosymbiosis. This can indeed be viewed as the biochemical innovation that separated the future organelle from its infectious relatives. It is indeed hard to envision that such an ATP-leaking symbiont would remain infectious, thereby signalling an irreversible switch from a parasitic symbiont to an obligate mutualistic symbiont. Therefore, this switch implies that the former pathogen had now become completely isolated in the intracellular environment of the eukaryote ancestors, with little ability to recombine with its original population, making it thus fully sensitive to reductive genome evolution through Muller's ratchet. This evidently boosted eukaryogenesis to generate LECA and allowed for the evolution of many free-energy demanding properties.

8. Obligate intracellular bacteria as helper genomes of plastid endosymbiosis: proposal for the carbohydrate metabolism symbiotic fluxes

Cyanobacteria are by no means preadapted to survive in the intracellular environment of eukaryotic phagotrophs. As is also the case for present-day plant cells, and by comparison to intracellular bacteria or free-living heterotrophs, their connectivity, albeit significant [64,65], is comparatively low, while these photoautotrophs most often require only, in addition to light and water, a few major minerals and trace elements. Yet, there is no question that an ancient cyanobacterial endosymbiont has achieved the status of a true eukaryotic organelle. But how did these organisms escape the deadly antimicrobial response of its phagotrophic host, and how did they establish the required connectivity while maintaining their membrane embedded electron transport chains fully functional? Possible answers to this question came from the finding of an unexpected phylogenetic signal in Archaeplastida genomes [41,66–73]. Archaeplastida comprise the three lineages (red algae, green algae and plants, and glaucophytes) that are direct descendants of the eukaryotic lineage that experienced primary plastid endosymbiosis. During the process of metabolic integration of the plastid, genes from the cyanobiont were transferred to the host nucleus. This protected the prokaryotic genes from Muller's ratchet, as discussed above, and generated a strong EGT signal in phylogenies amounting to 600/700 cyanobacterial genes that were transferred to the host nuclear genome(s) [74]. However, using the same phylogenomic techniques, another signal was found, uniting Archaeplastida and obligate intracellular bacterial pathogens in the order Chlamydiales. In this

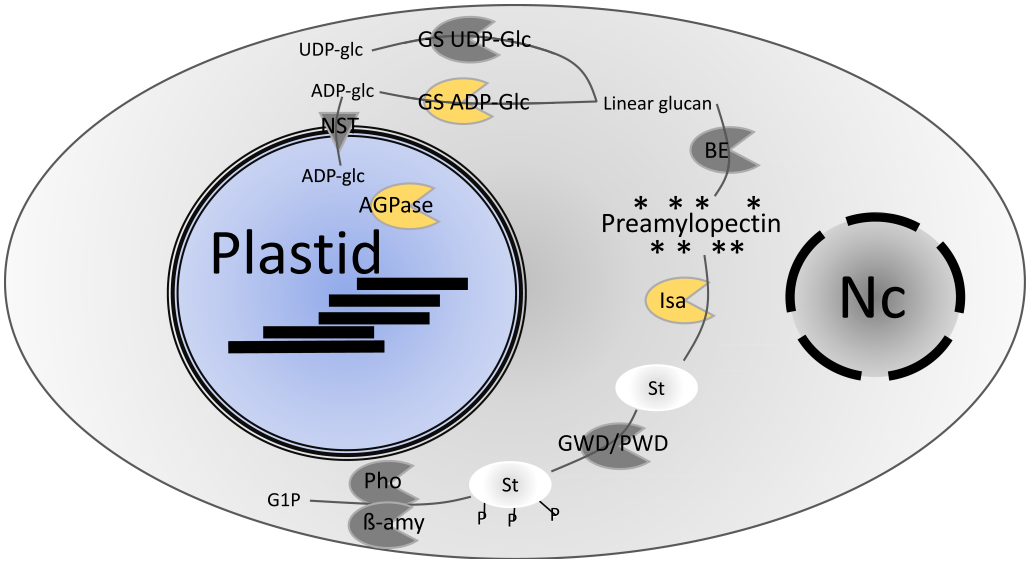

Figure 1. Simplified view of the proposed metabolic reconstruction of the function of starch metabolism in the last Archaeplastida common ancestor, based on [62,63]. Origins of the various enzymes are indicated with different colours, grey indicates a protein from host origin, while yellow indicates a protein from bacterial origin. Synthesis of the bacterial specific nucleotide-sugar ADP-glucose occurs in the cyanobiont. ADP-glucose is then exported from the plastid to the host cytosol through a host encoded NST transporter, where it is used by a bacterial glucan synthase using ADP-glucose. The latter consists of the ancestor of SSIII/SSIV starch synthase present in green plants and algae and in glaucophytes. This flux is combined to the cytosolic eukaryotic flux that relies solely on UDP-glucose through a UDP-glucose specific glucan synthase of host origin found in both glaucophytes and red algae. The linear glucans are then branched into preamylopectin. The step performed by isoamylase (Isa) consists in trimming the soluble preamylopectin precursor, thereby leading to crystallization and aggregation of the starch granule. Catabolism is first initiated by the two dikinases GWD and PWD, which destabilize the crystalline structure and allow degradation by β-amylase and glycogen/starch phosphorylase. Nc: Nucleus, BE: Branching enzyme, GS: Glucan synthase, AGPase: ADP-glucose pyrophosphorylase, Isa: Isoamylase, GWD/PWD: Glucan/Phosphoglucan Water Dikinase, Pho: Glycogen phosphorylase, β-amy: β-amylase, G1P: Glucose-1-phosphate, NST: nucleotide sugar transporter, ADP/UDP-glc: ADP/UDP glucose.

case, 50 to over 100 genes of chlamydial origin were identified in Archaeplastida genomes [41,73]. Several groups investigating this phylogenetic signal proposed that these pathogens played a role in endosymbiosis, possibly by supplying helper functions during metabolic integration of the nascent plastid [66,67,71,72]. A possible function was proposed by looking at the evolution of storage polysaccharide metabolism in Archaeplastida [62,66,75]. By comparing the corresponding extant enzyme networks, the simplest possible metabolic pathway for starch synthesis in the last common ancestor of Archaeplastida was proposed [62]. This hypothetical pathway displayed several interesting characteristics (Figure 1). First, it likely defined a major flux of reduced carbon from the plastid to the eukaryote host cytosol, where it was polymerized into starch or glycogen. Indeed, starch/glycogen, together with TAGs (lipid droplets), can be considered as the major form of carbon storage in the eukaryotic cytosol. Second, unlike eukaryotic glycogen metabolism, which relies exclusively on UDP-glucose as a synthesis substrate, this hypothetical pathway relied on the presence within the host cytosol of both ADP-glucose, the substrate

devoted to bacterial glycogen synthesis, and UDP-glucose (reviewed in Ball et al. [63,76,77]). The metabolic reconstruction had to account for the presence in the ancestors of the ADP-glucose-utilizing starch synthase SSIII/IV, found today in the green algae and in plant chloroplasts (Figure 1). This enzyme was initially thought to be encoded in the nucleus by an EGT from cyanobacteria, and thus to require ADP-glucose, at least initially. The proposed metabolic reconstructions clearly indicated that ADP-glucose pyrophosphorylase, the enzyme of ADP-glucose synthesis now found in the plastids of green algae and plants, was the only glycogen metabolism enzyme remaining inside the cyanobiont [62,63]. Interestingly, a recent phylogenetic analysis demonstrates a likely specific affiliation of ADP-glucose pyrophosphorylase to the *Gloemargarita lithophora* enzyme [78]. Because Gloeomargarita is presently thought to be the closest extant relative to the plastid donor, we believe that this affiliation strongly supports an EGT to the green lineage ancestors, rather than acquisition of the gene through LGT from other cyanobacteria post endosymbiosis. This further supports the idea that ADP-glucose pyrophosphorylase experienced a strong selection for its maintenance in the common ancestor of all Archaeplastida at the very onset of plastid endosymbiosis, despite all other enzymes of cyanobacterial glycogen metabolism having been lost! Indeed, enzymes of glycogen metabolism, in general, have been noted to be most often lost through the ratchet [79] as bacteria become facultative obligate intracellular pathogens or symbionts. In line with Henrissat's finding [79], the carbon-fixing Paulinella cyanobacterial chromatophore or the nitrogen-fixing cyanobacterial "spheroid bodies" mentioned above have had their genomes stripped clean of glycogen metabolism genes, while these are universally present in all their free-living relatives. Interestingly, ADP-glucose pyrophosphorylase is entirely and exclusively devoted to glycogen synthesis in cyanobacteria, where it defines the major rate-controlling step of the pathway (reviewed in Ball et al. [76]). This enzyme is finally tuned by photosynthesis through its substrates ATP and glucose-1-P and through its allosteric activator 3-PGA or through its Pi inhibitor. As a result, it generates a flux to storage optimized as a function of the photosynthetic activity and of the cyanobacterial energy charge, which justifies its maintenance within the plastid stroma. Unlike the ADP-glucose pyrophosphorylase, proposed to have remained in the plastids in the presence of its substrates and allosteric effectors, the ADP-glucose utilizing starch/glycogen synthase was, on the contrary, quite logically proposed to be located in the host cytosol. There, its glucan primer and polysaccharide products were selectively present, as is the case today for red algae and glaucophytes. For the flux to be operating, the ADP-glc synthesized within the plastid necessarily had to be transported from the cyanobiont to the cytosol. It was therefore proposed that an ancient host-derived nucleotide sugar transporter was responsible for this function [63]. At the time of the reconstruction proposal, solid phylogenetic evidence had been produced showing that most present plastidial carbohydrate transporters exporting photosynthate from the plastid to the cytosol in red and green algae or plants could be traced back to a unique ancestor that belonged to the nucleotide sugar translocator (NST) family of endomembrane eukaryotic transporters [80,81]. Of particular interest, some of the proximal eukaryotic endomembrane NSTs were GDP-mannose translocators, and such NSTs were subsequently shown to be very efficient ADP-Glc translocators, as demonstrated by uptake assays using recombinantly produced transporter protein reconstituted into proteoliposomes [75]. The reconstructed ancestral starch metabolism represents an optimal symbiotic flux to initiate a photosymbiosis based on the export of photosynthetic carbon to the cytosol of the eukaryotic host. Such a flux was challenging to establish due to the distinct and disconnected biochemical pathways of both partners. The cyanobacterium would provide carbon substrates during the day, independently of the host's demand, potentially flooding the exquisitely osmotically sensitive eukaryotic cytosol with unwanted and potentially cytotoxic osmotically active metabolites. On the other hand, the host may have strong carbon demands at night, at a time where the cyanobiont is unable to provide it. Linking the two disconnected biochemical networks through the osmotically inert carbon stores of the host neatly solves this asynchrony of demand and supply of carbon at the onset of plastid endosymbiosis, effectively buffering and synchronizing the biochemical fluxes of symbiosis. The export of ADP-glucose as photosynthate supply molecules that are unrecognized by eukaryotes

ensured that photosynthetic carbon would not affect the host's physiology and would be immediately incorporated into the host's glucan stores through the ADP-Glc specific glycogen/starch synthases. In addition, having the flux of photosynthate escape the cyanobiont through ADP-glucose is not expected to affect its central carbon metabolism in the light, as ADP-glucose synthesis only concerns the part of the cyanobiont's metabolism that was anyhow devoted to storage. However, NSTs exchange nucleotide sugars with AMP; consequently, export of NDP-sugars results in the net export of phosphate from the cyanobiont. If not balanced by an inverse flux of Pi, this NDP-sugar export will deplete the cyanobiont's phosphate pools and thereby affect, for example, ATP biosynthesis by photophosphorylation. This suggests the need for a Pi import mechanism coupled with carbon export in the light.

Mutants of *Synechocystis* defective for glycogen metabolism grow well photoautotrophically [82]. However, the same mutants display severely impaired growth and reduced viability when exposed to day and night cycles [82]. The absence of glucan stores in cyanobacteria is indeed logically expected to lead to severe ATP starvation in darkness in photoautotrophic cultures. However, when the symbiotic flux of storage polysaccharide metabolism was proposed [62], this problem was not addressed. Two key proteins were identified at the time as being required to operate metabolic symbiosis. These consisted of the NST transporter and the ADP-Glc utilizing starch/glycogen synthase (the ancestor of the green alga/plant SSIII/IV) [62]. This type of glucan synthase is normally never found in the eukaryote cytosol and displays a distribution restricted to prokaryotes. Its presence was deduced from the comparisons of red and green algae, due to its selective presence in the green lineage. The subsequent unexpected finding of the corresponding sequence in the first *Cyanophora paradoxa* glaucophyte genome highlighted the ancient presence of this enzyme for cytosolic starch synthesis [83]. This unexpected finding brought considerable support to the proposed reconstitution of ancestral starch metabolism.

However, while SSIII/IV was originally thought to be of cyanobacterial phylogeny [62], subsequent analysis rejected a cyanobacterial ancestry for this enzyme, while displaying a surprising, likely chlamydial origin [66]. Because the SSIII/IV enzyme seems to be at the core of the symbiotic connection, it was proposed that it was present in the cytosol of the host at the onset of the process. To achieve this, it was suggested that glycogen metabolism enzymes were, in fact, cytosolic effectors secreted by Chlamydiae in the cytosol of their hosts, to manipulate the host carbon fluxes at their advantage. Glycogen metabolism, prior to these studies, had never been viewed as a provider of effector enzymes. Glucose storage had mainly been considered as an intra-chlamydial housekeeping function, and therefore not concerned with highly specialized biotic interactions of the chlamydial growth cycle.

One important consequence of this proposal is that, if SSIII/IV was indeed a chlamydial effector that was targeted to the host cytosol at the onset of plastid endosymbiosis, it would *ipso facto* be at the centre of carbohydrate metabolism symbiotic fluxes. This would immediately tie the chlamydia pathogen/symbiont into the symbiotic relationship uniting not two but three genomes. This became known as the MATH (Ménage à Trois Hypothesis) where three rather than two genomes encoded essential symbiosis functions during plastid endosymbiosis. However if the gene encoding an essential function was transferred from the chlamydia to the host genome, ERGTs from this third genome would thus become the opposing force whereby the host replaced chlamydia as the genome encoding the chlamydial effectors. As long as the chlamydial ancestor either initially had, or shaped thereafter, novel favourable biotic interactions through secretion of chlamydia encoded effectors, its maintenance would remain necessary to the tripartite symbiosis because it housed indispensable genes. However when the supply of novel chlamydial effectors was exhausted and after the corresponding genes had all been secured by ERGT in the host nuclear genome, the pathogen/symbiont vanished. Indeed, its cellular compartment, at variance with the electron transport membranes carrying cyanobiont, was not hosting any indispensable biochemical reaction. The ancestor of the SSIII/IV starch synthase was not the only chlamydial effector transmitted through ERGT to the storage polysaccharide metabolism network of Archaeplastida. GlgX, encoding a "direct debranching enzyme" (also called isoamylase), was also transmitted to all three Archaeplastida lineages.

Isoamylase (ISA) is never found in eukaryotes, save Archaeplastida. Glycogen-accumulating eukaryotes use another, more complex enzyme named "indirect debranching enzyme (iDBE)" to hydrolyse the α-1,6 branch linkage through an altogether different mechanism (reviewed in Ball et al. [76,77]). Mutants of isoamylase in all plants tested revert to the accumulation of hydrosoluble glycogen. The latter was likely the ancestral form of hydrosoluble storage polysaccharide present in the cytosol of the host (reviewed in Cenci et al. [84]). It was therefore proposed that isoamylase was responsible for trimming out the branches that supported those chains during synthesis; otherwise, these branches would have prevented the proper alignment of chains required for the crystallization of glucans into huge solid and insoluble aggregates known as starch [85]. That ISA assumes an analogous function in red algae is suggested by the recent description of mutants of *Cyanidioschizon merolae* defective for ISA and starch synthesis [86]. However, we do not think that the initial function of the chlamydial effector was to generate starch from glycogen metabolism outright. GlgX ISA enzymes in bacteria have evolved to ensure glycogen breakdown through the release of small unbranched malto-oligosaccharides (MOS). Eukaryotes do not produce MOS in the cytosol and do not have MOS breakdown enzymes in this compartment. Hence, having GlgX together with other chlamydial effectors in the cytosol would not only induce glycogen synthesis through chlamydial effectors using host substrates and energy, but it would also give privileged and exclusive access to the Chlamydiae of the glycogen metabolism breakdown products. The function of ISA in starch crystallization likely evolved later and required the co-evolution of novel host-encoded glucan-water dikinases, which are indispensable to the breakdown of glucan crystallites in eukaryotes (reviewed in detail in Ball et al. [76], Cenci et al. [84]).

9. Obligate intracellular bacteria as helper genomes of plastid endosymbiosis: establishing the nitrogen metabolism symbiotic fluxes

Carbohydrate (starch) metabolism was not the only pathway impacted by multiple chlamydial ERGTs; tryptophan and isoprenoid metabolism were equally impacted. Of a total of seven protein subunits involved in the selective biosynthesis of tryptophan from chorismate, three are entirely cyanobacterial in all Archaeplastida, two are entirely chlamydial, one is entirely planctomycetal, and one is partly chlamydial (red algae) and partly cyanobacterial (glaucophytes and green algae) [87]. The Planctomycete ERGT of TrpE is of high significance, since this subunit controls and regulates the flux to tryptophan through its feedback sensitivity to the final product [87].

Of all 20 amino acids, tryptophan displays the most energy-consuming synthesis pathway and, as such, has led to the evolution of several documented biotic interactions, whereby chlamydia subverts its host to supply either tryptophan itself, tryptophan metabolites, or substitutes such as kynurenine or indole (reviewed in Bonner et al. [88]). According to the species and the strategy of tryptophan scavenging involved, chlamydia would have either lost or kept most of the genes for tryptophan synthesis. Chlamydiales, rather than Chlamydiaceae, are closer to the suspected donors of ERGTs to Archaeplastida. Among them, *Simkania negevensis* contains the full suite of tryptophan synthesis genes from chorismic acid (7 genes) on an operon hosting, in addition, a gene encoding 3-Deoxy-D-arabinoheptulosonate 7-phosphate synthase (DAHP synthase) [88]. DAHP synthase is a major flux-controlling enzyme in the shikimate pathway leading to chorismic acid, the substrate for aromatic amino acid biosynthesis, including tryptophan, and for the building of many other aromatic compounds. If such an 8-gene set was introduced into a eukaryote by LGT either to an organelle (mitochondrion or plastid) or to another intracellular pathogenic/symbiotic bacterium, then these compartments could be induced to massively overproduce tryptophan. However, for such a conjugative gene-transfer type of biotic interaction to be useful to an incipient Simkania-like Chlamydiales, the latter would have to also trigger the release of tryptophan to the cytosol from which it could be harnessed for subsequent use. It was thus proposed that an ancient Chlamydiales transferred both the genes of tryptophan synthesis and tryptophan export to the cyanobiont. The phylogeny of TyrP/Mtr, previously identified as the major tryptophan transporters of Chlamydiae [88] was thus investigated [87]. TyrP was thereby proven to define another well supported yet previously undetected ERGT from Chlamydiae to all

three Archaeplastida lineages, arguing for a very early implementation of this pathway during endosymbiosis [87]. It is very difficult to consider that ERGT of TyrP and those of several tryptophan biosynthetic genes from Chlamydiae to Archaeplastida could be coincidental. This observation thus lends considerable support to what could have been otherwise considered as far-fetched speculations. In addition, during these investigations, evidence was provided for a direct exchange of genes in the intracellular environment between Chlamydiales and other obligate or facultative intracellular bacteria (Legionella, Coxiella, Piscirickettsia), in both the phylogenies of the tryptophan synthesis genes and that of the TyrP transporters [87].

It must be emphasized that, before these analyses, some evidence had been obtained very early on concerning possible conjugative gene transfers of group I introns containing single LAGLIDADG Homing Endonuclease Genes (HEGs). These transfers involved the Chlamydiales *Simkania negevensis*, green alga plastids and protist mitochondria [69]. These intron transfers were reinvestigated a few years later by Brinkman et al. [68], and nearly a decade later by Haugen et al. [89], who added cyanobacteria to the previously reported list of organisms and organelles involved in these group I intron gene transfers. What was not previously noted in these studies is that the few thermophilic freshwater cyanobacteria reported to house these introns are phylogenetically close to what is now considered by many as the likeliest candidate relatives of the plastid donor [89–91]. However, because of the patchy distribution of the strains concerned by these exchanges, no firm conclusions can be made as to the source of these introns. The suspicion thus remains that they could reflect ancient gene transfers between Chlamydiales and the cyanobacterial ancestors of plastids.

We do believe, nevertheless, that a strong case exists for an initial interaction between chlamydia, cyanobiont and host that involved both the export of photosynthate through the action of chlamydial encoded cytosolic effector enzymes secreted through the chlamydial T3SS, and the induction of tryptophan synthesis within the evolving plastids through direct operon and gene transfers from chlamydial genes or plasmids. The presence of a robust planctomycete ERGT for TRPE (anthranilate synthase subunit I) in glaucophytes as well as green and red algae proves anyhow that the cyanobiont had lost very early on the cyanobacterial control of flux to tryptophan by anthranilate synthase feedback inhibition. The shikimate pathway, which provides the chorismate common substrate for aromatic amino acid biosynthesis, is absent in most eukaryotes except for fungi and Archaeplastida. While some evidence suggests the presence of multiple cyanobacterial EGTs in this prokaryotic pathway [92], the first rate-controlling step of this pathway (DAHP synthase) is suggestive of a non-cyanobacterial ERGT from a non-identified bacterial source [92], again arguing for early loss of cyanobacterial control of these fluxes.

10. Obligate intracellular bacteria as helper genomes of plastid endosymbiosis: protection from ATP starvation and anoxia

Under ATP starvation conditions in darkness, chlorophyll biosynthesis is impaired because magnesium chelatase cannot assemble correctly. This results in the accumulation of ROS-generating intermediates of chlorophyll synthesis during the night, and cell death when light returns (reviewed in Deschamps et al. [62]). It has been proposed that the redirection of the starch metabolism pathway to the plastid was selectively favoured in the rising green algae to prevent the photooxidative stresses due to the increased demand on chlorophyll synthesis in this lineage (reviewed in Deschamps et al. [62]). This increase resulted from the building of the novel chlorophyll b-containing light-harvesting complexes. Starch is thought to ensure ATP homeostasis in darkness in the stroma, thanks to the in situ availability of storage glucans. A very insightful study [93] proved that knockout mutants of the Arabidopsis ATP transporters (NTT), responsible for the unidirectional ATP import into plastids at night, experience photooxidative stresses. Furthermore, it proved that the severity of the stresses experienced depended chiefly on the amount of starch remaining in the chloroplasts at the end of the night phase, which is in line with the aforementioned proposal. It can be concluded that both plastidial starch accumulation and ATP import from the cytosol through the NTT transporters protect the plastid from otherwise lethal photooxidative stresses. In addition, a plastidial starch pool facilitates the recycling of organically bound phosphate, through the

release of Pi from sugar phosphates during starch biosynthesis, thereby increasing the robustness of photosynthetic metabolism in the light.

Cyanobacteria do not have ATP import proteins and thus entirely rely on their glycogen stores to ensure ATP homeostasis during the night, and prevent the accumulation of ROS-generating intermediates of chlorophyll synthesis. The slow growth and viability loss of glycogen metabolism mutants of *Synechocystis* sp. PCC 6803 [82] grown under alternating day and night cycles can be, at least partly, understood in this light. The MAT hypothesis posits that chlamydial effectors generate a flux of carbohydrates from the plastid to the host cytosol that is expected to substantially diminish or even wipe out glycogen accumulation in the cyanobiont. This is, of course, expected to generate severe ATP starvation during the night. It is therefore quite remarkable to report that all three Archaeplastida lineages (glaucophytes and red and green algae, and plants) have experienced a clear-cut chlamydial ERGT of a gene encoding the ATP-import protein, the hallmark of "energy parasitism" in Chlamydiales and Rickettsiales. It was proposed that the gene encoding this transporter was transfered by conjugation to the cyanobiont, together with other critical transporters, thereby obviating the need for an as yet inexistent organelle protein targeting machinery. Hence, not only did chlamydia initiate the symbiotic fluxes, but it also provided the tools to survive the ensuing ATP starvation generated from such fluxes.

Two other pathways (namely those of isoprenoid (MEP) and 2-methyl-1,4-naphthoquinone (DHNA) syntheses) have also been targeted by multiple chlamydial LGTs. Interestingly, the naphtoquinone is coupled to polyprenyl chains, which are themselves synthesized through the polymerization of isoprenoids to generate both menaquinone (vitamin K2) and phylloquinone (vitamin K1). The prenylated quinone yields the physiologically active electron carrier menaquinone. Menaquinones shuttle electrons between different respiratory complexes in anaerobic respiration, or aerobic respiration of bacteria in a microaerophilic environment. Reduction of the prenyl side chain leads to phylloquinone, which, unlike the membrane-diffusing respiratory menaquinones, is a fixed electron carrier of PSI found in cyanobacteria and Archaeplastida, or their derivatives. Menaquinone and phylloquinone share a common pathway, and all Archaeplastida contain such quinones, which are essential for PSI function and, therefore, for photosynthesis. Two distinct pathways (Men and Futalosine) have been documented for the conversion of chorismate to menaquinone or phylloquinone in bacteria and archaea (for review see Zhi et al. [94]). All red and green algae use the so-called "Men" pathway involving the production of OSB (o-succinyl benzoate), which is transformed into DHNA in the peroxisome. Our present knowledge of menaquinone/Phylloquinone metabolism in glaucophytes does not allow the identification of early Men pathway enzymes [95], save for the presence of MenB in *Cyanoptyche gloeocystis*. MenB uses O-succinyl benzoyl CoA as a substrate. This therefore suggests the presence of an OSB-generating Men rather than a Futalosine pathway in glaucophytes. However, we must presume in this case that the glaucophyte enzymes have diverged beyond recognition. Interestingly, freshwater basal cyanobacteria, such as Gloeobacter and Gloeomargarita, seem to lack both the Men and Futalosine pathways. Yet Gloeobacter, at least, can produce menaquinone under axenic growth conditions and hosts menaquinone in PSI as do all organisms performing oxygenic photosynthesis, which host either menaquinone or phylloquinone [96]. Because Gloeomargarita-like cyanobacteria are the closest relatives to the plastid ancestor, it is thus entirely possible that glaucophytes contain cyanobacterial Men pathway genes that remain unrecognizable.

Mesophilic red algae and green algae (and plants) contain nuclear sequences that yield the production of various fusion protein arrangements in the Men pathway. They are collectively named PHYLLO. However, unicellular thermophilic red algae (the Cyanidiales), which harbour the most gene-rich plastid genomes among Archaeplastida, contain a seven-gene cluster encoding all the genes required for the production of the DHNA menaquinone core, in addition to MenA, which is required for its prenylation. Cyanidiales, unlike other red and green algae, do not contain nuclear PHYLLO genes. While initial studies suggested that this gene cluster may have been monophyletic with the nuclear PHYLLO proteins [97], more recent analyses question this result while not rejecting it [95]. The relationship between PHYLLO and this cluster remains, unfortunately, unresolved.

No extant cyanobacteria contain these genes, and the cluster must have entered the plastome at minimum before Cyanidiales diversification (if the PHYLLO fusions and the cluster genes are considered polyphyletic) and perhaps even before Archaeplastida diversification (if PHYLLO and plastidial cluster genes have at least in part a common origin). Single gene phylogenies of the small Men proteins are particularly problematic. Few trees of the Men cluster are indeed congruent, which is due to the intensive gene sharing among bacteria, thereby defeating phylogenetic signal increases by concatenating the analysis of these small proteins. Despite this, the largest Men protein (MenD) displays robust phylogenies and clearly shows a chlamydial origin [95]. Moreover, the smaller MenF gene shows enough congruence with MenD to allow for concatenation, further proving a chlamydial origin. Despite the unresolved phylogenies for the other Men genes of the cluster, all trees show Cyanidiales monophyly, strongly suggesting that the cluster came in one piece. Hence, we can safely conclude that the plastome acquired this cluster through LGT from Chlamydia [95]. We propose that the mechanism involved is identical to that hypothesized for tryptophan metabolism, and consisted of conjugative transfer. Indeed, those Chlamydiales suspected of having donated genes to the Archaeplastida ancestors are fully equipped to handle not only protein effector secretion through the T3SS in the host cytosol or elsewhere, but also conjugative DNA transfer through the T4SS [98]. That such organellar cluster transfers are indeed occurring between obligatory intracellular pathogens/symbionts and their hosts is exemplified by the relationship uniting Phycorickettsia and its secondary plastid endosymbiosis host ochrophyte lineages [99,100]. In this case, Phycorickettsia has also donated a 6 "ebo" gene cluster of unknown function to the plastome of several eustigmatophyte algae. The cluster conditions the biosynthesis of an unknown prenylated compound, which in other bacteria is likely involved in the transport of small molecules across membranes [101]. Interestingly, evidence of multiple losses from the plastome of genes encoding the plastidial ACP (acyl carrier protein) gene responsible for plastidial fatty acid biosynthesis have been recently reported [99]. These losses were allowed by the prior presence of a nuclear LGT in eustigmatophyte algae. This LGT consists of a phycorickettsial ACP gene whose product was targeted to the secondary plastids. This is reminiscent of the chlamydial ERGTs to the host of primary plastid endosymbiosis. However, there is no present indication that the LGT is correlated to the secondary plastid endosymbiosis that generated the ochrophytes. Nevertheless, the timing of this nuclear LGT seems to coincide with the base of the eustigmatophyte algae and could still very well define a eustigmatophyte-specific acquisition resulting from a bacterial ERGTs that occurred during secondary endosymbiosis of ochrophytes. Unlike Chlamydiales, which disappeared a long time ago from Archaeplastida, Phycorickettsia is still actively replicating in eustigmatophyte algae today. Interestingly, only those eustigmatophyte lineages that have conserved the naked zoospores life stage in the eustigmatophytes life cycle seem to be prone to Phycorickettsia infection [100]. Other lineages devoid of this stage might have closed the door to infection by building a permanent cell wall. This, again, could recapitulate what has happened with Chlamydiales after primary plastid endosymbiosis. As with the menaquinone chlamydial gene cluster in Cyanidiales, the plastidial ebo gene cluster is present in strains that are not currently infected by Phycorickettsia. We can thus presume that it provides an important, albeit unknown, selection advantage to the extant ochrophyte host, which results from the exceptionally high levels of expression of these genes afforded by plastids [102]. The selection advantage yielded by the final biosynthetic product of ebo gene activity might very well be one of several reasons for the maintenance of the Phycorickettsia symbiont in those lineages lacking the ebo plastidial cluster.

Does the plastidial menaquinone gene cluster yield a comparable selection advantage in Cyanidiales? We argue that this is indeed the case. Because warmer waters are notably impoverished in oxygen, expression of the menaquinone gene cluster to very high levels from plastidial genes would have increased significantly and flooded the small Cyanidiales cell volume with menaquinone with no need for specific transporters, as such molecules readily cross membranes. Menaquinone supplied in the diet is known to be used in animal mitochondria, where its substitutes for ubiquinone to ensure optimal ATP synthesis in a microaerophilic environment. This would selectively protect extant Cyanidiales from anoxic stress. We argue, nevertheless, despite the

unresolved nature of the relationship between PHYLLOs and the plastidial gene cluster, that this function might also have applied to the common ancestor of Archaeplastida. Overexpression of the Men pathway in the cyanobiont would have equally protected the cyanobiont from anoxic stresses. In addition, if the speculations concerning early induction of both tryptophan synthesis and the shikimate pathway that generates the chorismic acid substrate are correct, then the implementation of massive menaquinone synthesis at endosymbiosis would have been greatly facilitated because of chorismate overproduction. Green alga photosymbionts of salamander embryo cells have been recently documented to express a severe hypoxic stress response as they penetrate within the animal cell cytosol [103,104]. It is likely that free-living cyanobacteria that have been enslaved within eukaryotes in a chlamydial inclusion would likely experience similar hypoxic stresses. Free-living cyanobacteria contain high affinity terminal oxidases to manage such stresses in the extracellular environment. These would have benefited from an increase in menaquinone supply to further optimize respiratory ATP synthesis. Under these conditions, a large increase in menaquinone synthesis afforded by conjugative transfer of the Men genes would have been mutually beneficial not only for both Chlamydiae and cyanobacteria, but also for mitochondrial respiration by the host. Indeed, this would have further protected the cyanobiont from ATP starvation in darkness, in conjunction with the expression of the ATP import protein discussed above.

However, for menaquinone to assume this protective function, overproduction of the core quinone must be in line with the supply of its C20 prenyl tails. The latter would have depended on the synthesis of geranylgeranyl diphosphate from the isoprenoids DMAPP and IPP. DMAPP and IPP are generated from the bacterial MEP pathway, which is otherwise absent from eukaryotes unrelated to Archaeplastida. The MEP pathway seems to have been massively impacted by chlamydial ERGTs [95,105]. Of a total of seven genes, two are certainly of chlamydial origin, while a third is possibly of chlamydial (or PVC) origin in all three Archaeplastida lineages. In addition to this, a fourth gene displays a clear-cut chlamydial origin in green algae, while both glaucophyte and red algae have conserved the cyanobacterial copy of the gene [95,105]. Only two genes display a clear cyanobacterial ancestry in all three Archaeplastida lineages, while the first committed step of isoprenoid synthesis (DOXP synthase (Dxs)) that initially gated the flux of the cyanobacterial pathway is now of alphaproteobacterial origin. This again, as was the case for tryptophan synthesis, suggests that the cyanobacteria had lost control of the MEP pathway early on. As for tryptophan metabolism, and possibly menaquinone biosynthesis, we propose that the MEP genes were donated by conjugative transfer to the cyanobiont, resulting in a chimeric merozygote genome. We speculate that the high levels of synthesis of isoprenoids that possibly resulted from the biotic interaction between chlamydia and the cyanobiont were not desirable to maintain while forging a novel cellular organelle, thereby leading to the preference for another bacterial source for ERGT of the flux-limiting step. We indeed believe that flux-controlling enzyme genes of chlamydial origin in general were tailored initially to favour chlamydia replication. After the plastid protein-targeting machines evolved, these genes were counterselected when the gene encoding such enzymes was transferred. This happened to effectively match the flux required by the host, rather than the pathogen/symbiont. Likewise, expression of the menaquinone core genes from the host nuclear "PHYLLO" gene fusion instead of the plastidial cluster significantly decreased the flux to menaquinone.

11. Obligate intracellular bacteria as helper genomes of plastid endosymbiosis: protection from host immunity and the modified MATH

When the first glaucophyte genome sequence became available, the genome was screened for the presence of the triose phosphate/phosphate translocators (TPT) which are members of the plastidic phosphate translocator (pPTs) family of transporters. TPTs are known to export reduced carbon from plastids of both red and green algae, as well as land plants. pPTs have been shown to be evolutionarily derived from a unique host nucleotide-sugar transporter of the NST3 family. This NST3 family encodes several GDP-mannose transporters of the endomembrane system that are known to display transport activity for the bacterial-specific ADP-glucose substrate, a structural analogue of GDP-mannose. As detailed

above, this observation led to the proposal of the MATH, according to which carbon was originally exported from plastids as ADP-glucose, which was then polymerized into host glycogen by an ADP-Glc-specific glycogen synthase that was encoded and secreted by Chlamydia. The enzyme was secreted by the T3SS as an effector. This ensured full membership of Chlamydia in an exclusive tripartite photosymbiosis club.

However, no TPTs or even worse, no pPTs, were found in the Cyanophora genome sequence. Only six NSTs were found in the genome, none of which was phylogenetically related to the pPT family of transporters. These results were confirmed by the purification of the glaucophyte plastids (the muroplasts) and an experimental characterization of the muroplast proteome [106]. This analysis identified only 14 transporters, which is at least one order of magnitude lower than the number found by detailed proteomic analysis in the green lineage, as well as the number suggested by bioinformatic analysis of red alga genomes. Surprisingly, 25% of the transporter complement was found to be of chlamydial origin, including the major glucose-6-P transporter of Chlamydiales, UhpC. This bacterial transporter exchanges glucose-6-P for orthophosphate in a manner reminiscent of pPTs. Although this finding can be seen as a spectacular confirmation of the importance of Chlamydiales in early events of plastid endosymbiosis, it challenges the original version of the MATH, which was based on ADP-glucose rather than glucose-6-P export from the ancestral cyanobiont. Because the MATH, as originally proposed, was supported by many other observations, Facchinelli et al. [107] proposed an alternative MATH model in which the cyanobiont entered together with a Chlamydiales via phagocytosis (Figure 2). The phagocytic vacuole was transformed into a chlamydial inclusion, preventing its acidification and thus protecting the cyanobiont from phagocytosis. The cyanobiont thus temporarily evaded the highly aggressive antibacterial immunity of the phagotrophic host. The secluded inclusion environment also facilitated direct gene transfer from chlamydia to the cyanobiont, equipping it with genes encoding essential transporters, such as the ATP import protein, TyrP, and most importantly, UhpC. These gene transfers may also have included deregulated but redundant pathways, such as those of tryptophan, isoprenoid, and menaquinone core biosynthesis. Chlamydia would also secrete its usual set of glycogen metabolism enzyme effectors, both into the inclusion lumen, where only bacterial enzymes including the enzyme of ADP-glucose synthesis were active, and into the host cytosol, in the presence of host glycogen metabolism, as described in the original MATH. Facchinelli et al. [107] therefore proposed that UhpC was the initial translocator that exported photosynthate into the lumen of the inclusion, and that the host membrane-derived inclusion contained a NST3 transporter for ADP-glucose export from the lumen of the inclusion to the host cytosol. It should be emphasized that in this scheme, the host only receives the overflow of inclusion glycogen metabolism (Figure 2). Nevertheless, the paradox of the missing pPTs in the glaucophyte genomes has been neatly solved and reconciled with the MATH. Furthermore, counter exchange of glucose 6-phosphate with orthophosphate by UhpC allows for metabolic coupling via phosphate balancing, and it avoids depletion of the cyanobiont's phosphate pool, which would be detrimental for photosynthetic carbon assimilation.

12. Support for MATH from extant host-chlamydia interactions

Can we support such seemingly far-fetched speculations with hard experimental data? As for the discovery of the chlamydial LGT signal in plants [73] and the demonstration of the effector nature of glycogen metabolism enzymes both in vitro and in vivo [66,108], the answer once again came from microbiologists and cell biologists studying these pathogens. By far the most studied models are defined by the animal-specific infecting Chlamydiaceae pathogens. Among these, certain strains of *Chlamydia trachomatis* (an important human pathogen) accumulate glycogen to such high levels within their inclusions that it becomes easy to quantify it through cytological staining. It can thus be easily distinguished from cytosolic host glycogen. Gehre et al., in 2016, first demonstrated, using gene silencing techniques, that the eukaryotic genes encoding enzymes of host glycogen metabolism had little or no effect on inclusion glycogen accumulation [109]. They further demonstrated that inclusion glycogen accumulation was under the control of chlamydial enzymes, thereby

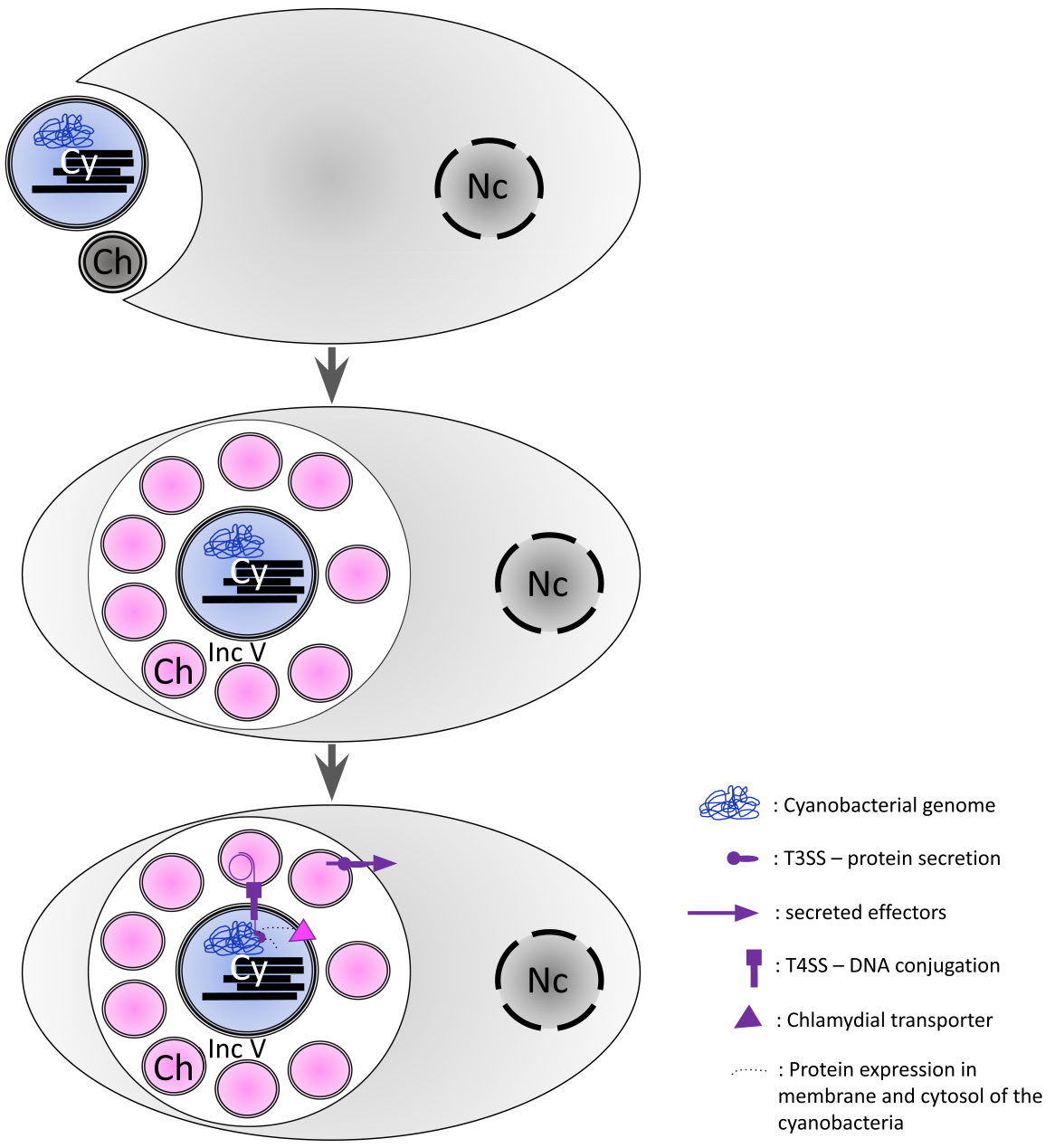

Figure 2. Simplified representation of the MATH (Ménage à Trois Hypothesis) suggested in [107]. Chlamydia and cyanobacteria were engulfed together leading to the localization of the cyanobacteria inside the chlamydial inclusion. This allowed two important types of transfers for plastid symbiosis: (i) Protein secretion in the host cytosol through the chlamydial type III secretion system (T3SS), which allowed to hijack host metabolism. (ii) Direct conjugation through the type 4 secretion system (T4SS) between chlamydia and cyanobacteria, allowing for direct gene transfer leading to the programmed enslavement of cyanobacterial metabolism, and transport of various metabolites inside the inclusion. This led to direct exchange between host, cyanobacteria, and chlamydia. Nc: Nucleus, Cy: Cyanobacteria, Inc V: Inclusion vesicule, Ch: Chlamydial cell.

confirming the results obtained by Nguyen and Valdivia through the isolation of glycogen-branching enzyme mutants of *Chlamydia trachomatis* [110]. A noticeable exception reported by Gehre et al [109] was the host UDP-glucose pyrophosphorylase, which greatly impacted inclusion glycogen synthesis when decreased through silencing. However, the Chlamydiaceae studied differed from the protist-infecting Chlamydiales that are thought to have affected plastid endosymbiosis by two criteria. First, ADP-glucose pyrophosphorylase (the enzyme of ADP-Glc synthesis) is the only enzyme of glycogen metabolism in *Chlamydia trachomatis* that has been shown not to be a T3SS-secreted effector, whereas all glycogen metabolism enzymes of protist-infecting Chlamydiales have been validated as T3SS effectors, including ADP-glucose pyrophosphorylase. Consistent with this observed difference, the recombinant chlamydial glycogen synthase of *C. trachomatis* is able to efficiently use both UDP-Glc and ADP-Glc, a highly unusual property for a glycogen synthase. However, the glycogen synthases tested so far in protist-infecting Chlamydiales were selective for ADP-Glc. The finding of a major impact of host UDP-Glc synthesis by the host UDP-Glc pyrophosphorylase on inclusion lumen glycogen accumulation in *C. trachomatis* led to the hypothesis that the pathogen was importing UDP-Glc across the inclusion membrane to fuel luminal glycogen synthesis; indeed, no ADP-glucose pyrophosphorylase can be secreted to the lumen of this organism. SLC35D2, one of the established human UDP-Glc translocators, was shown by Gehre et al. [109] to be recruited to the *Chlamydia trachomatis* inclusion membrane. Silencing of SLC35D2 had the strongest effect on inclusion glycogen accumulation, validating this NST as the major host transporter active for chlamydial inclusion glycogen metabolism in *C. trachomatis* (Figure 3). Thus, the modified MATH is supported by very solid experimental evidence demonstrating the presence of a critical NST for inclusion glycogen synthesis in Chlamydiaceae. Following these studies, Cenci et al. [111] very simply extrapolated the Chlamydiaceae results to protist-infecting Chlamydiales. The only differences taken into account were the additional presence of the chlamydial ADP-glucose pyrophosphorylase in both the host cytosol and in the inclusion lumen, and a substrate specificity of the chlamydial glycogen synthase restricted to ADP-glucose. In this context, ADP-Glc synthesis could occur in the cytosol from the host's ATP and Glucose-1-P suppies in a classical pathogenetic "energy parasite" fashion. The ADP-glc synthesized can only be used by the chlamydial glycogen synthase to feed into either cytosolic host or inclusion lumen glycogen pools. For such a pathogen flux to be functional in the inclusion, we need only to imagine that a host NST3 capable of translocating ADP-Glc replaces SLC35D2 on the inclusion membrane (Figures 3, 4) If we further imagine the presence within the lumen of the chlamydial inclusion of a cyanobacterium that has been reprogrammed by chlamydial DNA transfers of transporter genes and deregulated pathways, then we will have massive ADP-Glc and glycogen synthesis in the inclusion lumen during the day. This will, of course, immediately benefit the Chlamydia, but the host will also get its share through the reverse flux of ADP-Glc through the NST3 transporter and the polymerization of glucose into glycogen stores (Figure 4). However, the host would only get the overflow of photosynthate left over by Chlamydia. We do not consider this to be day 1 of plastid endosymbiosis, but rather day −1, when a cyanobacterium, otherwise rather unlikely to survive in the phagotroph's intracellular environment and ill-adapted to install sophisticated symbiotic fluxes, acquired the necessary tools to do so and, more importantly, to go all the way to become a true novel cellular organelle at variance with all currently known photosymbionts, with the noticeable exception of *Paulinella chromatophora*.

13. What was "Day 1" of plastid endosymbiosis?

In our view, the chlamydial biotic interactions described above do not signal the onset of plastid endosymbiosis. We believe that these interactions involve specific pathways that are reflected by the presence of multiple chlamydial ERGTs in Archaeplastida. Bacterial ERGTs are common and by no means limited to Chlamydiae, but multiple bacterial ERGTs from the same source in the same pathway are exceedingly rare and mostly limited to Chlamydiae. We believe that in such cases the original interaction consisted of overexpression of genes from whole pathways, which was achieved by donating deregulated chlamydial copies of these genes to increase flux. This was accompanied by the transfer of genes

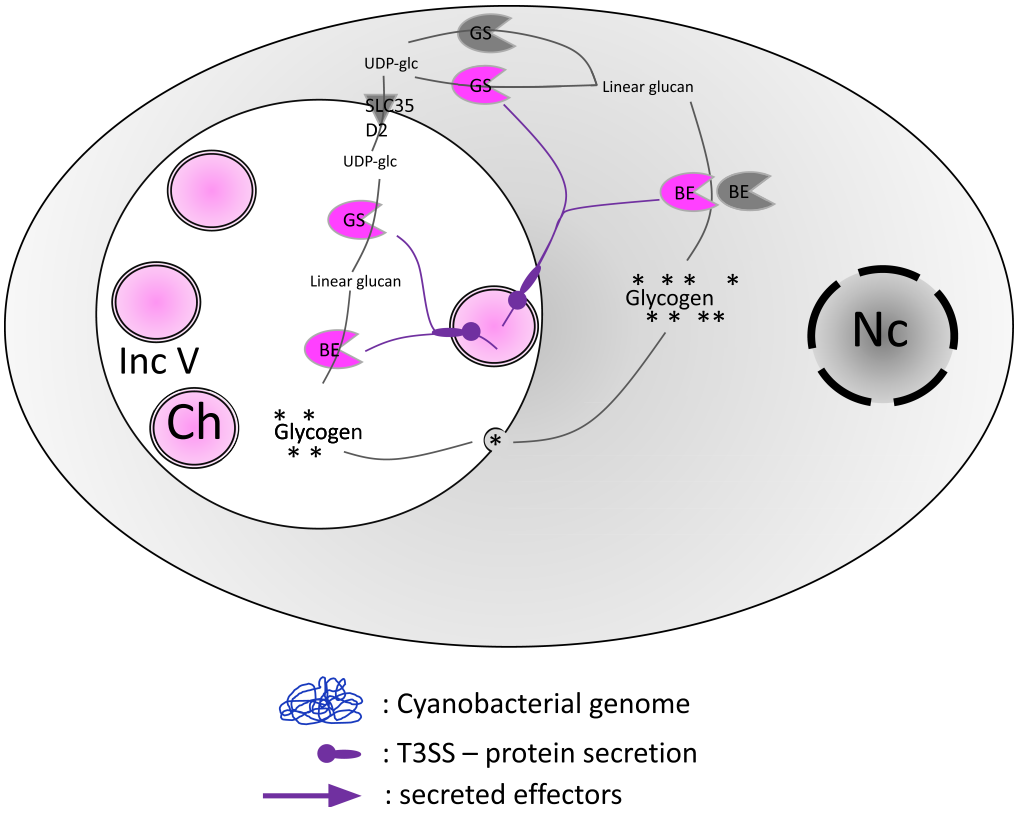

Figure 3. Simplified view of extant glycogen metabolism of *C. trachomatis* infected human cells, based on [109]. Chlamydia is found inside its inclusion, from where it secretes protein effectors in both the host cytosol and the lumen of the chlamydial inclusion. Chlamydial enzymes were demonstrated to control inclusion lumen glycogen accumulation provided that the pathogen are fed with host UDP-Glc substrate. This occurs thanks to the recruitment of the human UDP-glucose transporter SLC35D2 on the inclusion membrane. Besides this major pathway of inclusion lumen glycogen accumulation, some evidence has been provided for the presence of a minor pathway of cytosolic glycogen import within the inclusion. Nc: Nucleus, BE: Branching enzyme, GS: Glucan synthase, AGPase: ADP-glucose pyrophosphorylase, SLC35D2: nucleotide sugar transporter, UDP-glc: UDP glucose, Inc V: Inclusion vesicule, Ch: Chlamydial cell.

encoding future plastidial transporters, to ensure that whatever useful compound was being synthesized, the latter would be accessible, predominantly in the inclusion lumen. The host would get the leftovers. It must be emphasized that ERGTs of transporter genes are also exceedingly rare among bacterial ERGTs and mostly limited to Chlamydiae.

Nevertheless, the host leftovers were substantial and included most importantly photosynthetic O_2, which is of no use for Chlamydiales but extremely desirable for host mitochondrial respiration, as well as menaquinone for nocturnal respiration in a microaerophilic environment, some carbohydrates into its glycogen stores, and tryptophan, among other things. We suggest that the demands made on cyanobacterial metabolism were considerable, and that, despite the efforts of chlamydia to ensure ATP homeostasis in the enslaved cyanobionts through gene transfers, the association may have been transient. In the absence of extant representatives of the responsible chlamydial symbionts, we can only speculate about the number and type of genes involved and the modalities of such gene transfers. We propose that, at minimum, UhpC, the ATP

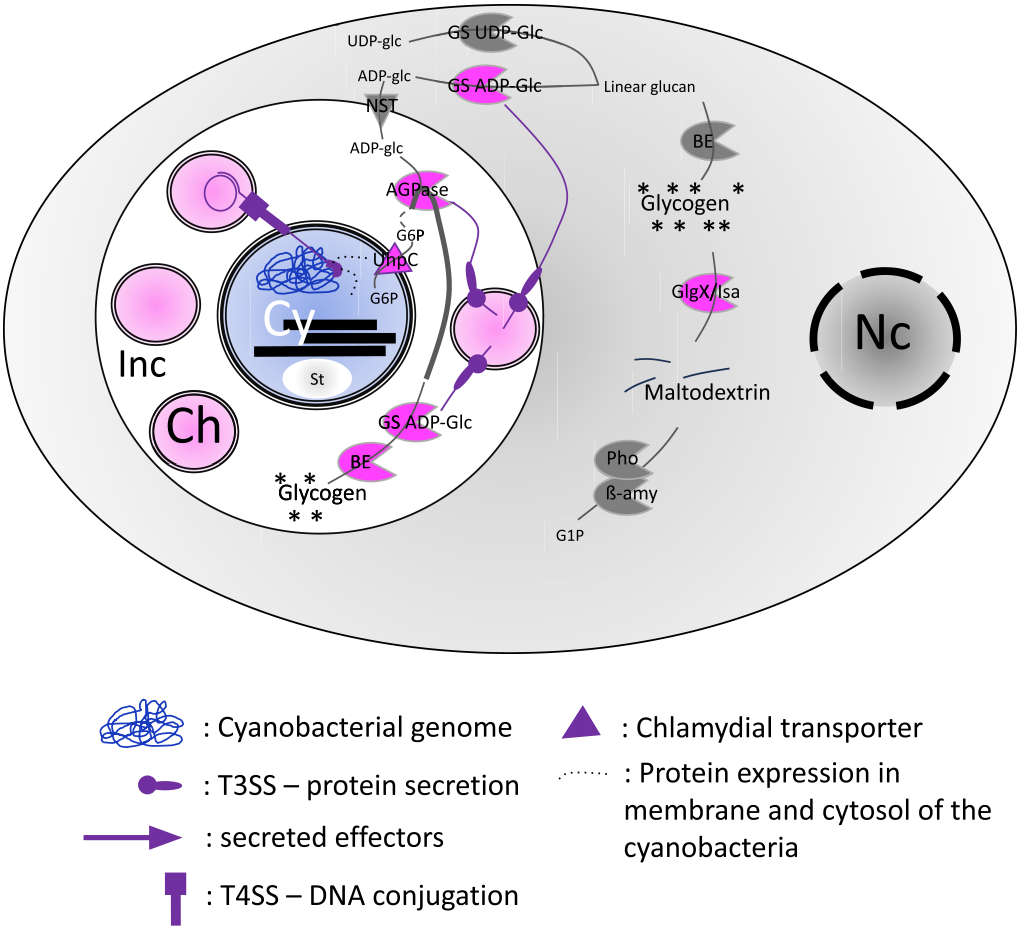

Figure 4. Proposed glycogen metabolism under the revised MATH following Facchinelli et al. [107] and Cenci et al. [111]. The cyanobiont entered together with chlamydia in the common ancestor of plastid endosymbiosis. It is seen inside the chlamydial inclusion where it received the chlamydial UhpC gene through T4SS-mediated conjugation leading to G6P export. As was demonstrated for *C. trachomis*-infected extant human cells (see Figure 3), the synthesis of glycogen was likely occuring in the inclusion thanks to the classical chlamydial enzymes effectors involved in alpha-glucan storage (AGPase, GS ADP-Glc, BE), and would have represented the main flux of glycogen accumulation (bold arrow). In addition, ADP-glucose, due to a high concentration of this metabolite in the inclusion, would have been exported to the host cytosol by an host encoded NST (nucleotide-sugar transporter). At this stage, the host would thus get the overflow of ADP-glucose. This bacterial-specific metabolite was incorporated in host glycogen thanks to the chlamydial ADP-glucose specific glucan synthase (GS-ADP-Glc) effector as detailed in Figure 1. Nc: Nucleus, BE: Branching enzyme, GS: Glucan synthase, AGPase: ADP-glucose pyrophosphorylase, G6P: glucose-6-phosphate, G1P: Glucose-1-phosphate Inc: Inclusion, Ch: Chlamydial cell, NST: nucleotide sugar transporter, ADP/UDP-glc: ADP/UDP glucose.

import protein (NTT), TyrP (the tryptophan transporter) the whole plastidial menaquinone cluster and the isoprenoid MEP and tryptophan pathways were donated by a chlamydial ancestor, possibly on a conjugative plasmid in a fashion similar to the Agrobacterium/plant biotic interaction, albeit in the intracellular environment of protists. The host was possibly a phagotroph, grazing mostly on

freshwater cyanobacteria. The introduction of photosymbiosis to alleviate the stress of anoxia may have been highly desirable in the environment, especially if the latter was warm and the available oxygen was limited. The reprogrammed cyanobacteria were a dead end as far as survival in the extracellular environment is concerned, because it was leaking tryptophan, glucose-6-P, quinones, and isoprenoid. We define Day 1 of plastid endosymbiosis as the first day following the escape from the chlamydial inclusion to the cytosol of such a Chlamydia-enslaved cyanobiont. This escape was not trivial because it possibly entailed a complex fusion of the prokaryotic cyanobacterial outer membrane with the eukaryotic-derived but chlamydia remodelled inclusion membrane, leading to novel outer cyanobiont membranes with both inner-layer prokaryotic and outer-layer mixed prokaryotic/eukaryotic features. We can only speculate about the properties displayed by the cyanobiont's novel outer membrane. We do not know what level of protection from host immunity is provided by this membrane, nor how much of the ancient chlamydial protein targeting to the inclusion has remained functional. This escape resulted in the presence of transporters not only of cyanobacterial or chlamydial origin, encoded in the cyanobacterial merozygote genome, but also of host origin that were recruited by Chlamydiales on their inclusion, thereby establishing early connectivity of the evolving plastid. Both ADP-Glc transport via the host NST and glucose-6-P transport via UhpC generated suitable carbohydrate fluxes cementing the endosymbiotic relationship. During the evolution of the glaucophyte lineage, the NST-mediated transport was lost; in contrast, during evolution of green and red algae, the NST that had evolved into a sugar phosphate transporter similar to UhpC replaced the latter. If, unlike the inclusion membrane, the novel rearranged cyanobiont membranes became sensitive to RAMPs, the cyanobiont may have needed to become resistant to such peptides. Several possible resistance mechanisms are documented. Among these, RAMP import into the cyanobacterium, followed by proteolytic degradation of the peptides, provided a scaffold on which the TIC–TOC machinery could evolve to ensure the targeting of proteins containing RAMP-like plastid targeting sequences, as discussed above, recapitulating the events that had occurred during mitochondrial endosymbiosis. It is plausible that the escaped cyanobiont received the critical genes from other intracellular bacteria present in the host, as suggested by recent studies [30]. At this stage (day 1) the cyanobiont still contained all the genes in its free-living relatives. However, it also contained a moderately chimeric genome, including the genes required to export photosynthate (UhpC) and tryptophan in an optimized fashion, as well as those required to obviate the oxidative, anoxic, and energy stresses generated by this export, and to survive in an intracellular environment. These are things that the cyanobacterial genomes were not tailored to achieve. Most importantly, it did all these things using specific protein transporters, without resorting to antimicrobial peptide tinkering of membrane permeability to generate symbiotic fluxes. Such tinkering is observed in the many extant symbioses studied in plants and insects (see above), including in Paulinella, which is devoid of transporters on the chromatophore membranes. It did this by leaving the photosynthetic membrane function intact. It is useful here to continue the comparison between Agrobacterium and the hypothetical chlamydia symbiont that could have primed plastid endosymbiosis to help grasp the amount of evolution required in these processes. As was the case for the Agrobacterium/plant mediated interactions, it probably took millions of years to evolve the chlamydial biotic interactions that resulted in high frequency conjugation of transporters and deregulated genes that prompted tryptophan and photosynthate export, as well as those genes required to resist ATP starvation and ensure prolonged survival in the chlamydial inclusion. From the cyanobacterium point of view, it took only hours, days and weeks to be engulfed by a protist together with chlamydia and to be reprogrammed by conjugation to become enslaved in exporting photosynthate, amino acids, quinones, or isoprenoids. This is similar to the growth of a tumor during the Agrobacterium/plant interactions. Reprogrammed plant tumor cells have no future on their own in the environment, and reprogrammed and metabolite leaking cyanobacteria would have met with very limited success if accidently released in the environment.

Successful escape from the inclusion and simultaneous acquisition of resistance to eukaryotic host cell immunity was a complex and rare event. The cyanobiont was now isolated from its free-living relatives and experienced Muller's ratchet, yielding

reductive genome evolution. The irreversible nature of the multiple possible patterns of gene losses began distinguishing several lineages that were selected in different environments and that experienced the forging of different types and numbers of new host genes. Finally, this process ended with the emergence of green algae, red algae and glaucophytes. This process of reductive genome evolution, leading to different outcomes of gene transfers to the host genome, together with the evolution of both common and different sets of novel host genes, is collectively referred to as metabolic integration of the protoplastid. The Chlamydia-primed plastid endosymbiosis can thus be considered as a highly successful process. As pointed out in this review, the precise nature of the symbiosis occurring at the very beginning of the reductive evolutionary process would have significantly influenced the outcome and success of the endosymbiosis. Indeed, the chlamydia-primed symbiosis offered an unparalleled level of preadaptation at the very beginning of the process, which may explain the rarity of primary plastid endosymbiosis compared to the plethora of extant diverse photosymbioses. Despite the relative speed of action of the ratchet, the entire metabolic integration may have taken up tens to hundreds of millions of years. Even after the escape of its enslaved endosymbiont, the first phases of this evolution still required the maintenance of the chlamydial symbiont, which secreted beneficial effector enzymes into the host cytosol. Indeed, the presence of chlamydial genes in the evolving chimeric plastome did not warrant the maintenance of the symbiont, since the plastid fully took over the production of these proteins; rather, it was the evolution of new useful cytosolic effectors for the tripartite symbiosis that have ensured further maintenance. In the end, the transfer of these useful genes to the nuclear genome by ERGT would have rendered chlamydia redundant and probably signalled its disappearance from Archaeplastida.

14. What next? Can we find the culprit of MATH?

Because of the apparent disappearance of Chlamydia pathogens from Archaeplastida, studies leading to the ménage à trois hypothesis have mainly relied on linking the phylogenetic imprint of chlamydial ERGTs to knowledge of biochemical pathways and their possible functions in both chlamydial and cyanobacterial physiology. While interesting and rich in useful predictions (see above), this approach alone is not sufficient to prove the MATH, let alone encourage young scientists to further investigate a topic that has been met with such strong and often poorly justified opposition [112–114]. Indeed, when looking at biochemical pathways and gene transfers in Archaeplastida genomes, we will always be looking at single gene trees of mostly enzymes. Depending on the nature of organisms investigated, the inference of evolutionary trajectories in some trees may be obscured by repeated gene sharing events, which are particularly difficult to interpret when occurring in deep time. In addition, these trees target enzymes that respond to local selection in lineages (e.g., subtle substrate changes, shifts in compartmental localization) rather than being highly conserved, slowly evolving functions such as ribosomal RNAs. As such, the trees will in most cases be poorly resolved, with either poor or no bootstrap support, as well as topological rearrangements that are challenging to reconcile. Everyone in the field of deep evolution is aware of these problems, and people who perform such analyses are accustomed to dealing with phylogenetic uncertainty. It is therefore logical to amend the deep-time toolkit with robust biochemistry-based arguments whenever possible. For instance, we believe that knowledge of gene distributions should be actively pursued to resolve questions about the directionalities of gene transfer [112,113]. For us, the shortcoming of phylogenetic approaches is not the inevitable uncertainties, but rather that even a well-established ERGT obviously does not tell us per se what was the underlying mechanism of its emergence. For instance, is a specific chlamydial ERGT a mere gene replacement of a cyanobacterial gene because the protein was poorly suited to the TOC–TIC targeting mechanism and because Chlamydiae were present in the environment? Or does it reflect a more sophisticated biotic interaction whereby the gene was first introduced into the cyanobiont? We cannot rely on phylogenetic trees to resolve these issues. One possibility would be to make progress by attempting endosymbiosis experimentally. Such approaches are ongoing, and will likely inform us on many issues dealing with the establishment of bacteria or eukaryotes within host phagotrophs and eukaryotes

in general. However, it is unlikely to yield information about the particular question dealt with in this review, which is addressing organelle-generating endosymbioses. Another possibility would be to find Chlamydiales that infect protists grazing chiefly on freshwater basal cyanobacteria related to those suspected to be the closest relatives of the plastid progenitor. Finding a chlamydia interacting with cyanobacteria in such a phagotroph would indeed be a substantial leap forward. The recent finding of the Phycorickettsia/ochrophyte interaction encourages us in this respect.

Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

Acknowledgements

SGB, UC, ML, and CC were supported by the CNRS, the Université de Lille CNRS, the ANR grants "Mathtest" (ANR-18-CE13-0027) "expendo" (ANR-14-CE11-0024) and "ménage à trois" (ANR-12-BSV2-0009) and the Région Hauts de France and European project Alibiotech. UC, SGB, CC acknowledge BILILLE to provide computational resources. IL and OV were supported by ANR grant "Mathtest" (ANR-18-CE13-0027).

References

[1] H. J. Muller, "The relation of recombination to mutational advance", *Mutat. Res.* **106** (1964), pp. 2–9.

[2] N. A. Moran, "Accelerated evolution and Muller's rachet in endosymbiotic bacteria", *Proc. Natl. Acad. Sci. USA* **93** (1996), no. 7, pp. 2873–2878.

[3] G. M. Bennett and N. A. Moran, "Small, smaller, smallest: the origins and evolution of ancient dual symbioses in a Phloem-feeding insect", *Genome Biol. Evol.* **5** (2013), no. 9, pp. 1675–1688.

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

[5] A. Maréchal and N. Brisson, "Recombination and the maintenance of plant organelle genome stability", *New Phytol.* **186** (2010), no. 2, pp. 299–317.

[6] M. Naito and T. E. Pawlowska, "Defying Muller's ratchet: ancient heritable endobacteria escape extinction through retention of recombination and genome plasticity", *mBio* **7** (2016), no. 3, article no. e02057-15.

[7] D. R. Speth, L. M. Zeller, J. S. Graf, W. A. Overholt, K. Küsel and J. Milucka, "Genetic potential for aerobic respiration and denitrification in globally distributed respiratory endosymbionts", *Nat. Commun.* **15** (2024), no. 1, article no. 9682.

[8] D. C. Bublitz, G. L. Chadwick, J. S. Magyar, et al., "Peptidoglycan production by an insect-bacterial mosaic", *Cell* **179** (2019), no. 3, article no. 703-712.e7.

[9] A. Nakabachi, K. Ishida, Y. Hongoh, M. Ohkuma and S.-Y. Miyagishima, "Aphid gene of bacterial origin encodes a protein transported to an obligate endosymbiont", *Curr. Biol.* **24** (2014), no. 14, R640–R641.

[10] T. H. Coale, V. Loconte, K. A. Turk-Kubo, et al., "Nitrogen-fixing organelle in a marine alga", *Science* **384** (2024), no. 6692, pp. 217–222.

[11] E. C. M. Nowack and A. R. Grossman, "Trafficking of protein into the recently established photosynthetic organelles of Paulinella chromatophora", *Proc. Natl. Acad. Sci. USA* **109** (2012), no. 14, pp. 5340–5345.

[12] J. M. Archibald, "Endosymbiosis and eukaryotic cell evolution", *Curr. Biol.* **25** (2015), no. 19, R911–R921.

[13] G. I. McFadden, "Origin and evolution of plastids and photosynthesis in eukaryotes", *Cold Spring Harb. Perspect. Biol.* **6** (2014), no. 4, article no. a016105.

[14] R. I. Ponce-Toledo, P. López-García and D. Moreira, "Horizontal and endosymbiotic gene transfer in early plastid evolution", *New Phytol.* **224** (2019), no. 2, pp. 618–624.

[15] T. G. Stephens, A. Gabr, V. Calatrava, A. R. Grossman and D. Bhattacharya, "Why is primary endosymbiosis so rare?", *New Phytol.* **231** (2021), no. 5, pp. 1693–1699.

[16] W. Martin and M. Müller, "The hydrogen hypothesis for the first eukaryote", *Nature* **392** (1998), no. 6671, pp. 37–41.

[17] T. Dagan, M. Roettger, K. Stucken, et al., "Genomes of stigonematalean cyanobacteria (Subsection V) and the evolution of oxygenic photosynthesis from Prokaryotes to Plastids", *Genome Biol. Evol.* **5** (2013), no. 1, pp. 31–44.

[18] C. Kneip, V. Christine, P. J. Lockhart and U. G. Maier, "The cyanobacterial endosymbiont of the unicellular algae Rhopalodia gibba shows reductive genome evolution", *BMC Evol. Biol.* **8** (2008), article no. 30.

[19] T. Nakayama and Y. Inagaki, "Genomic divergence within non-photosynthetic cyanobacterial endosymbionts in rhopalodiacean diatoms", *Sci. Rep.* **7** (2017), no. 1, article no. 13075.

[20] R. A. Foster and J. P. Zehr, "Diversity, genomics, and distribution of phytoplankton-cyanobacterium single-cell symbiotic associations", *Annu. Rev. Microbiol.* **73** (2019), pp. 435–456.

[21] T. H. Coale, V. Loconte, K. A. Turk-Kubo, et al., "Nitrogen-fixing organelle in a marine alga", *Science* **384** (2024), no. 6692, pp. 217–222. Epub April 11, 2024. PMID: 38603509.

[22] G. Ballabani, M. Forough, F. Kessler and V. Shanmugabalaji, "The journey of preproteins across the chloroplast membrane systems", *Front. Physiol.* **14** (2023), article no. 1213866.

[23] Ú. Flores-Pérez and P. Jarvis, "Molecular chaperone involvement in chloroplast protein import", *Biochim. Biophys. Acta* **1833** (2013), no. 2, pp. 332–340.

[24] Y. Wang, C. H. Chen, D. Hu, M. B. Ulmschneider and J. P. Ulmschneider, "Spontaneous formation of structurally diverse membrane channel architectures from a single antimicrobial peptide", *Nat. Commun.* **7** (2016), article no. 13535.

[25] F.-A. Wollman, "An antimicrobial origin of transit peptides accounts for early endosymbiotic events", *Traffic* **17** (2016), no. 12, pp. 1322–1328.

[26] E. A. Groisman, C. Parra-Lopez, M. Salcedo, C. J. Lipps and F. Heffron, "Resistance to host antimicrobial peptides is necessary for Salmonella virulence", *Proc. Natl. Acad. Sci. USA* **89** (1992), no. 24, pp. 11939–11943.

[27] I. Guefrachi, O. Pierre, T. Timchenko, et al., "Bradyrhizobium BclA is a peptide transporter required for bacterial differentiation in symbiosis with aeschynomene legumes", *Mol. Plant-Microbe Interact.* **28** (2015), no. 11, pp. 1155–1166.

[28] C. L. Shelton, F. K. Raffel, W. L. Beatty, S. M. Johnson and K. M. Mason, "Sap transporter mediated import and subsequent degradation of antimicrobial peptides in Haemophilus", *PLoS Pathog.* **7** (2011), no. 11, article no. e1002360.

[29] O. D. Caspari and I. Lafontaine, "The role of antimicrobial peptides in the evolution of endosymbiotic protein import", *PLoS Pathog.* **17** (2021), no. 4, article no. e1009466.

[30] C. Garrido, F.-A. Wollman and I. Lafontaine, "The evolutionary history of peptidases involved in the processing of organelle-targeting peptides", *Genome Biol. Evol.* **14** (2022), no. 7, article no. evac101.

[31] P. Mergaert, "Role of antimicrobial peptides in controlling symbiotic bacterial populations", *Nat. Prod. Rep.* **35** (2018), no. 4, pp. 336–356.

[32] L. Carro, P. Pujic, N. Alloisio, et al., "Alnus peptides modify membrane porosity and induce the release of nitrogen-rich metabolites from nitrogen-fixing Frankia", *ISME J.* **9** (2015), no. 8, pp. 1723–1733.

[33] H. Charles, S. Balmand, A. Lamelas, et al., "A genomic reappraisal of symbiotic function in the aphid/Buchnera symbiosis: reduced transporter sets and variable membrane organisations", *PloS One* **6** (2011), no. 12, article no. e29096.

[34] H. Feng, N. Edwards, C. M. H. Anderson, et al., "Trading amino acids at the aphid-Buchnera symbiotic interface", *Proc. Natl. Acad. Sci. USA* **116** (2019), no. 32, pp. 16003–16011.

[35] P. Mergaert, Y. Kikuchi, S. Shigenobu and E. C. M. Nowack, "Metabolic integration of bacterial endosymbionts through antimicrobial peptides", *Trends Microbiol.* **25** (2017), no. 9, pp. 703–712.

[36] A. Singer, G. Poschmann, C. Mühlich, et al., "Massive protein import into the early-evolutionary-stage photosynthetic organelle of the amoeba paulinella chromatophora", *Curr. Biol.* **27** (2017), no. 18, 2763–2773.e5.

[37] F. Masson, A. Zaidman-Rémy and A. Heddi, "Antimicrobial peptides and cell processes tracking endosymbiont dynamics", *Philos. Trans. R. Soc. Lond. Ser. B, Biol. Sci.* **371** (2016), no. 1695, article no. 20150298.

[38] S. G. Ball, D. Bhattacharya and A. P. M. Weber, "Evolution. Pathogen to powerhouse", *Science* **351** (2016), no. 6274, pp. 659–660.

[39] E. J. van Schaik, C. Chen, K. Mertens, M. M. Weber and J. E. Samuel, "Molecular pathogenesis of the obligate intracellular bacterium Coxiella burnetii", *Nat. Rev. Microbiol.* **11** (2013), no. 8, pp. 561–573.

[40] R. Seshadri, I. T. Paulsen and J. A. Eisen, "Complete genome sequence of the Q-fever pathogen Coxiella burnetii", *Proc. Natl. Acad. Sci. USA* **100** (2003), no. 9, pp. 5455–5460.

[41] A. Collingro, P. Tischler, T. Weinmaier, et al., "Unity in variety—the pan-genome of the Chlamydiae", *Mol. Biol. Evol.* **28** (2011), no. 12, pp. 3253–3270.

[42] S. A. Muñoz-Gómez, S. Hess, G. Burger, B. F. Lang, E. Susko, C. H. Slamovits and A. J. Roger, "An updated phylogeny of the Alphaproteobacteria reveals that the parasitic Rickettsiales and Holosporales have independent origins", *eLife* **8** (2019), article no. e42535.

[43] Z. Wang and M. Wu, "An integrated phylogenomic approach toward pinpointing the origin of mitochondria", *Sci. Rep.* **5** (2015), article no. 7949.

[44] A. J. Roger, S. A. Muñoz-Gómez and R. Kamikawa, "The origin and diversification of mitochondria", *Curr. Biol.* **27** (2017), no. 21, R1177–R1192.

[45] J. Martijn, J. Vosseberg, L. Guy, P. Offre and T. J. G. Ettema, "Deep mitochondrial origin outside the sampled alphaproteobacteria", *Nature* **557** (2018), no. 7703, pp. 101–105.

[46] S. A. Muñoz-Gómez, E. Susko, K. Williamson, L. Eme, C. H. Slamovits, D. Moreira, P. López-García and A. J. Roger, "Site-and-branch-heterogeneous analyses of an expanded dataset favour mitochondria as sister to known Alphaproteobacteria", *Nat. Ecol. Evol.* **6** (2022), no. 3, pp. 253–262.

[47] M. E. Schön, J. Martijn, J. Vosseberg, S. Köstlbacher and T. J. G. Ettema, "The evolutionary origin of host association in the Rickettsiales", *Nat. Microbiol.* **7** (2022), no. 8, pp. 1189–1199.

[48] T. P. Driscoll, V. I. Verhoeve, M. L. Guillotte, et al., "Wholly rickettsia! reconstructed metabolic profile of the quintessential bacterial parasite of eukaryotic cells", *mBio* **8** (2017), no. 5, article no. e00859-17.

[49] R. J. Hogan, S. A. Mathews, S. Mukhopadhyay, J. T. Summersgill and P. Timms, "Chlamydial persistence: beyond the biphasic paradigm", *Infect. Immun.* **72** (2004), no. 4, pp. 1843–1855.

[50] J. Bondy-Denomy and A. R. Davidson, "When a virus is not a parasite: the beneficial effects of prophages on bacterial fitness", *J. Microbiol. (Seoul, Korea)* **52** (2014), no. 3, pp. 235–242.

[51] L. König, C. Wentrup, F. Schulz, F. Wascher, S. Escola, M. S. Swanson, C. Buchrieser and M. Horn, "Symbiont-mediated defense against legionella pneumophila in amoebae", *mBio* **10** (2019), no. 3, article no. e00333-19.

[52] F. Landmann, "The Wolbachia endosymbionts", *Microbiol. Spectr.* **7** (2019), no. 2, article no. 72.

[53] N. Nikoh, T. Hosokawa, M. Moriyama, K. Oshima, M. Hattori and T. Fukatsu, "Evolutionary origin of insect-Wolbachia nutritional mutualism", *Proc. Natl. Acad. Sci. USA* **111** (2014), no. 28, pp. 10257–10262.

[54] W. Le Clec'h, J. Dittmer, M. Raimond, D. Bouchon and M. Sicard, "Phenotypic shift in Wolbachia virulence towards its native host across serial horizontal passages", *Proc. Biol. Sci.* **284** (2017), no. 1859, article no. 20171076.

[55] Z. Wang and M. Wu, "Phylogenomic reconstruction indicates mitochondrial ancestor was an energy parasite", *PloS One* **9** (2014), no. 10, article no. e110685.

[56] P. C. J. Donoghue, C. Kay, A. Spang, G. Szöllősi, A. Nenarokova, E. R. R. Moody, D. Pisani and T. A. Williams, "Defining eukaryotes to dissect eukaryogenesis", *Curr. Biol.* **33** (2023), no. 17, R919–R929.

[57] V. Tobiasson, J. Luo, Y. I. Wolf and E. V. Koonin, "Dominant contribution of Asgard archaea to eukaryogenesis", Preprint, bioRxiv, 2024. Online at https://doi.org/10.1101/2024.10.14.618318.

[58] M. Castelli, E. Sabaneyeva, O. Lanzoni, et al., "Deianiraea, an extracellular bacterium associated with the ciliate Paramecium, suggests an alternative scenario for the evolution of Rickettsiales", *ISME J.* **13** (2019), no. 9, pp. 2280–2294.

[59] I. Haferkamp and S. Schmitz-Esser, "The plant mitochondrial carrier family: functional and evolutionary aspects", *Front. Plant Sci.* **3** (2012), article no. 2.

[60] P. Dolezal, M. Aili, J. Tong, et al., "Legionella pneumophila secretes a mitochondrial carrier protein during infection", *PLoS Pathog.* **8** (2012), no. 1, article no. e1002459.

[61] D. Wilmaerts, M. Bayoumi, L. Dewachter, et al., "The persistence-inducing toxin HokB forms dynamic pores that cause ATP leakage", *mBio* **9** (2018), no. 4, article no. e007440-18.

[62] P. Deschamps, C. Colleoni, Y. Nakamura, et al., "Metabolic symbiosis and the birth of the plant kingdom", *Mol. Biol. Evol.* **25** (2008), no. 3, pp. 536–548.

[63] P. Deschamps, I. Haferkamp, C. d'Hulst, H. E. Neuhaus and S. G. Ball, "The relocation of starch metabolism to chloroplasts: when, why and how", *Trends Plant Sci.* **13** (2008), no. 11, pp. 574–582.

[64] S. Sengupta, D. Sahasrabuddhe and P. P. Wangikar, "Transporter engineering for the development of cyanobacteria as cell factories: a text analytics guided survey", *Biotechnol. Adv.* **54** (2022), article no. 107816.

[65] R. Stebegg, G. Schmetterer and A. Rompel, "Transport of organic substances through the cytoplasmic membrane of cyanobacteria", *Phytochemistry* **157** (2019), pp. 206–218.

[66] S. G. Ball, A. Subtil, D. Bhattacharya, et al., "Metabolic effectors secreted by bacterial pathogens: essential facilitators of plastid endosymbiosis?", *Plant Cell* **25** (2013), no. 1, pp. 7–21.

[67] B. Becker, K. Hoef-Emden and M. Melkonian, "Chlamydial genes shed light on the evolution of photoautotrophic eukaryotes", *BMC Evol. Biol.* **8** (2008), article no. 203.

[68] F. S. L. Brinkman, J. L. Blanchard, A. Cherkasov, et al., "Evidence that plant-like genes in chlamydia species reflect an ancestral relationship between Chlamydiaceae, Cyanobacteria, and the Chloroplast", *Genome Res.* **12** (2002), no. 8, pp. 1159–1167.

[69] K. D. Everett, S. Kahane, R. M. Bush and M. G. Friedman, "An unspliced group I intron in 23S rRNA links Chlamydiales, chloroplasts, and mitochondria", *J. Bacteriol.* **181** (1999), no. 16, pp. 4734–4740.

[70] J. Huang and J. P. Gogarten, "Concerted gene recruitment in early plant evolution", *Genome Biol.* **9** (2008), no. 7, article no. R109.

[71] J. Huang and J. P. Gogarten, "Did an ancient chlamydial endosymbiosis facilitate the establishment of primary plastids?", *Genome Biol.* **8** (2007), no. 6, article no. R99.

[72] A. Moustafa, A. Reyes-Prieto and D. Bhattacharya, "Chlamydiae has contributed at least 55 genes to plantae with predominantly plastid functions", *PLoS One* **3** (2008), no. 5, article no. e2205.

[73] R. S. Stephens, S. Kalman, C. Lammel, et al., "Genome sequence of an obligate intracellular pathogen of humans: chlamydia trachomatis", *Science* **282** (1998), no. 5389, pp. 754–759.

[74] H. Qiu, D. C. Price, A. P. M. Weber, F. Facchinelli, H. S. Yoon and D. Bhattacharya, "Assessing the bacterial contribution to the plastid proteome", *Trends Plant Sci.* **18** (2013), no. 12, pp. 680–687.

[75] C. Colleoni, M. Linka, P. Deschamps, M. G. Handford, P. Dupree, A. P. M. Weber and S. G. Ball, "Phylogenetic and biochemical evidence supports the recruitment of an ADP-glucose translocator for the export of photosynthate during plastid endosymbiosis", *Mol. Biol. Evol.* **27** (2010), no. 12, pp. 2691–2701.

[76] S. Ball, C. Colleoni and M. C. Arias, "The Transition from glycogen to starch metabolism in cyanobacteria and eukaryotes", in *Starch: Metabolism and Structure* (Y. Nakamura, ed.), Springer Japan: Tokyo, 2015, pp. 93–158.

[77] S. Ball, C. Colleoni, U. Cenci, J. N. Raj and C. Tirtiaux, "The evolution of glycogen and starch metabolism in eukaryotes gives molecular clues to understand the establishment of plastid endosymbiosis", *J. Exp. Bot.* **62** (2011), no. 6, pp. 1775–1801.

[78] P. Deschamps, S. G. Ball and D. Dauvillée, "The comparative, biochemistry, genetics, and evolution of starch metabolism in Chlamydomonas reinhardtii", *The Chlamydomonas Sourcebook*, Elsevier: Amsterdam, 2023, pp. 23–50.

[79] B. Henrissat, E. Deleury and P. M. Coutinho, "Glycogen metabolism loss: a common marker of parasitic behaviour in bacteria?", *Trends Genet.* **18** (2002), no. 9, pp. 437–440.

[80] H. M. Tyra, M. Linka, A. P. M. Weber and D. Bhattacharya, "Host origin of plastid solute transporters in the first photosynthetic eukaryotes", *Genome Biol.* **8** (2007), no. 10, article no. R212.

[81] A. P. M. Weber, M. Linka and D. Bhattacharya, "Single, ancient origin of a plastid metabolite translocator family in Plantae from an endomembrane-derived ancestor", *Eukaryot. Cell* **5** (2006), no. 3, pp. 609–612.

[82] M. Gründel, R. Scheunemann, W. Lockau and Y. Zilliges, "Impaired glycogen synthesis causes metabolic overflow

reactions and affects stress responses in the cyanobacterium *Synechocystis* sp. PCC 6803", *Microbiology* **158** (2012), no. Pt 12, pp. 3032–3043.

[83] D. C. Price, C. X. Chan, H. S. Yoon, et al., "Cyanophora paradoxa genome elucidates origin of photosynthesis in algae and plants", *Science* **335** (2012), no. 6070, pp. 843–847.

[84] U. Cenci, F. Nitschke, M. Steup, B. A. Minassian, C. Colleoni and S. G. Ball, "Transition from glycogen to starch metabolism in Archaeplastida", *Trends Plant Sci.* **19** (2014), no. 1, pp. 18–28.

[85] S. Ball, H.-P. Guan, M. James, et al., "From glycogen to amylopectin: a model for the biogenesis of the plant starch granule", *Cell* **86** (1996), no. 3, pp. 349–352.

[86] T. Maeno, Y. Yamakawa, Y. Takiyasu, et al., "One of the isoamylase isoforms, CMI294C, is required for semi-amylopectin synthesis in the rhodophyte Cyanidioschyzon merolae", *Front. Plant Sci.* **13** (2022), article no. 967165.

[87] U. Cenci, M. Ducatez, D. Kadouche, C. Colleoni and S. G. Ball, "Was the chlamydial adaptative strategy to tryptophan starvation an early determinant of plastid endosymbiosis?", *Front. Cell. Infect. Microbiol.* **6** (2016), article no. 67.

[88] C. A. Bonner, G. I. Byrne and R. A. Jensen, "Chlamydia exploit the mammalian tryptophan-depletion defense strategy as a counter-defensive cue to trigger a survival state of persistence", *Front. Cell. Infect. Microbiol.* **4** (2014), article no. 17.

[89] P. Haugen, D. Bhattacharya, J. D. Palmer, S. Turner, L. A. Lewis and K. M. Pryer, "Cyanobacterial ribosomal RNA genes with multiple, endonuclease-encoding group I introns", *BMC Evol. Biol.* **7** (2007), article no. 159.

[90] R. I. Ponce-Toledo, P. Deschamps, P. López-García, Y. Zivanovic, K. Benzerara and D. Moreira, "An early-branching freshwater cyanobacterium at the origin of plastids", *Curr. Biol.* **27** (2017), no. 3, pp. 386–391.

[91] N. C. Rockwell and J. C. Lagarias, "GUN4 appeared early in cyanobacterial evolution", *PNAS Nexus* **2** (2023), no. 5, article no. pgad131.

[92] T. A. Richards, J. B. Dacks, S. A. Campbell, et al., "Evolutionary origins of the eukaryotic shikimate pathway: gene fusions, horizontal gene transfer, and endosymbiotic replacements", *Eukaryot. Cell* **5** (2006), no. 9, pp. 1517–1531.

[93] T. Reinhold, A. Alawady, B. Grimm, et al., "Limitation of nocturnal import of ATP into Arabidopsis chloroplasts leads to photooxidative damage", *Plant J.* **50** (2007), no. 2, pp. 293–304.

[94] X.-Y. Zhi, J.-C. Yao, S.-K. Tang, Y. Huang, H.-W. Li and W.-J. Li, "The futalosine pathway played an important role in menaquinone biosynthesis during early prokaryote evolution", *Genome Biol. Evol.* **6** (2014), no. 1, pp. 149–160.

[95] U. Cenci, H. Qiu, T. Pillonel, et al., "Host-pathogen biotic interactions shaped vitamin K metabolism in Archaeplastida", *Sci. Rep.* **8** (2018), no. 1, article no. 15243.

[96] M. Mimuro, T. Tsuchiya, H. Inoue, et al., "The secondary electron acceptor of photosystem I in Gloeobacter violaceus PCC 7421 is menaquinone-4 that is synthesized by a unique but unknown pathway", *FEBS Lett.* **579** (2005), no. 17, pp. 3493–3496.

[97] J. Gross, J. Meurer and D. Bhattacharya, "Evidence of a chimeric genome in the cyanobacterial ancestor of plastids", *BMC Evol. Biol.* **8** (2008), article no. 117.

[98] M. Vouga, D. Baud and G. Greub, "Simkania negevensis, an insight into the biology and clinical importance of a novel member of the Chlamydiales order", *Crit. Rev. Microbiol.* **43** (2017), no. 1, pp. 62–80.

[99] T. Ševcíková, T. Yurchenko, K. P. Fawley, R. Amaral, H. Strnad, L. M. A. Santos, M. W. Fawley and M. Eliáš, "Plastid genomes and proteins illuminate the evolution of eustigmatophyte algae and their bacterial endosymbionts", *Genome Biol. Evol.* **11** (2019), no. 2, pp. 362–379.

[100] T. Yurchenko, T. Ševčíková, P. Přibyl, et al., "A gene transfer event suggests a long-term partnership between eustigmatophyte algae and a novel lineage of endosymbiotic bacteria", *ISME J.* **12** (2018), no. 9, pp. 2163–2175.

[101] K. Klicki, D. Ferreira, D. Hamill, B. Dirks, N. Mitchell and F. Garcia-Pichel, "The widely conserved ebo cluster is involved in precursor transport to the periplasm during scytonemin synthesis in nostoc punctiforme", *mBio* **9** (2018), no. 6, article no. e02266-18.

[102] M. Oey, M. Lohse, B. Kreikemeyer and R. Bock, "Exhaustion of the chloroplast protein synthesis capacity by massive expression of a highly stable protein antibiotic", *Plant J.* **57** (2009), no. 3, pp. 436–445.

[103] S. G. Ball and U. Cenci, "Gasping for air", *eLife* **6** (2017), article no. e27004.

[104] J. A. Burns, H. Zhang, E. Hill, E. Kim and R. Kerney, "Transcriptome analysis illuminates the nature of the intracellular interaction in a vertebrate-algal symbiosis", *eLife* **6** (2017), article no. e22054.

[105] L. Zeng and K. Dehesh, "The eukaryotic MEP-pathway genes are evolutionarily conserved and originated from Chlaymidia and cyanobacteria", *BMC Genom.* **22** (2021), no. 1, article no. 137.

[106] F. Facchinelli, M. Pribil, U. Oster, N. J. Ebert, D. Bhattacharya, D. Leister and A. P. M. Weber, "Proteomic analysis of the Cyanophora paradoxa muroplast provides clues on early events in plastid endosymbiosis", *Planta* **237** (2013), no. 2, pp. 637–651.

[107] F. Facchinelli, C. Colleoni, S. G. Ball and A. P. M. Weber, "Chlamydia, cyanobiont, or host: who was on top in the ménage à trois?", *Trends Plant Sci.* **18** (2013), no. 12, pp. 673–679.

[108] C. Lu, L. Lei, P. Bo, et al., "Chlamydia trachomatis GlgA is secreted into host cell cytoplasm", *PLOS One* **8** (2013), no. 7, article no. e68764.

[109] L. Gehre, O. Gorgette, S. Perrinet, et al., "Sequestration of host metabolism by an intracellular pathogen", *eLife* **5** (2016), article no. e12552.

[110] B. D. Nguyen and R. H. Valdivia, "Virulence determinants in the obligate intracellular pathogen Chlamydia trachomatis revealed by forward genetic approaches", *Proc. Natl. Acad. Sci. USA* **109** (2012), no. 4, pp. 1263–1268.

[111] U. Cenci, D. Bhattacharya, A. P. M. Weber, C. Colleoni, A. Subtil and S. G. Ball, "Biotic host-pathogen interactions

as major drivers of plastid endosymbiosis", *Trends Plant Sci.* **22** (2017), no. 4, pp. 316–328.

[112] S. G. Ball, D. Bhattacharya and A. P. M. Weber, "Infection and the first eukaryotes–response", *Science* **352** (2016), no. 6289, pp. 1065–1066.

[113] S. G. Ball, D. Bhattacharya, H. Qiu and A. P. M. Weber, "Commentary: plastid establishment did not require a chlamydial partner", *Front. Cell. Infect. Microbiol.* **6** (2016), article no. 43.

[114] D. Domman, M. Horn, T. M. Embley and T. A. Williams, "Plastid establishment did not require a chlamydial partner", *Nat. Commun.* **6** (2015), article no. 6421.