Acta Societatis Botanicorum Poloniae

INVITED ORIGINAL RESEARCH PAPER Acta Soc Bot Pol 83(4):291–302 DOI: 10.5586/asbp.2014.048
Received: 2014-11-17 Accepted: 2014-12-16 Published electronically: 2014-12-31

Primary endosymbiosis: have cyanobacteria and Chlamydiae ever been roommates?

Philippe Deschamps*
Unité d'Ecologie, Systématique et Evolution, CNRS UMR 8079, Université Paris-Sud, 91405 Orsay, France

Abstract

Eukaryotes acquired the ability to process photosynthesis by engulfing a cyanobacterium and transforming it into a genuine organelle called the plastid. This event, named primary endosymbiosis, occurred once more than a billion years ago, and allowed the emergence of the Archaeplastida, a monophyletic supergroup comprising the green algae and plants, the red algae and the glaucophytes. Of the other known cases of symbiosis between cyanobacteria and eukaryotes, none has achieved a comparable level of cell integration nor reached the same evolutionary and ecological success than primary endosymbiosis did. Reasons for this unique accomplishment are still unknown and difficult to comprehend. The exploration of plant genomes has revealed a considerable amount of genes closely related to homologs of Chlamydiae bacteria, and probably acquired by horizontal gene transfer. Several studies have proposed that these transferred genes, which are mostly involved in the functioning of the plastid, may have helped the settlement of primary endosymbiosis. Some of these studies propose that Chlamydiae and cyanobacterial symbionts coexisted in the eukaryotic host of the primary endosymbiosis, and that Chlamydiae provided solutions for the metabolic symbiosis between the cyanobacterium and the host, ensuring the success of primary endosymbiosis. In this review, I present a reevaluation of the contribution of Chlamydiae genes to the genome of Archaeplastida and discuss the strengths and weaknesses of this tripartite model for primary endosymbiosis.

Keywords: primary endosymbiosis; Archaeplastida; Chlamydiae; horizontal gene transfers

Introduction

Oxygenic photosynthesis in eukaryotes arose more than one billion years ago from the association of a cyanobacterial endosymbiont (or cyanobiont) with a heterotrophic host [1–3]. This event, called primary endosymbiosis, is considered to have happened with success only once during the evolution of eukaryotes. It allowed the emergence and diversification of a monophyletic supergroup, the Archaeplastida (or Plantae), composed of three lineages: Chloroplastida (green algae and land plants), Rhodophyta (red algae) and Glaucophyta (Fig. 1) [4–6]. From the original endosymbiotic relationship to the diversification of Archaeplastida, major metabolic and genomic rearrangements occurred. First, the genome of the cyanobiont was highly reduced and a significant part of it was transferred to the host nucleus [7,8]. Secondly, a specific targeting system evolved to export nucleus encoded proteins to the plastid, based on the combination of a N-terminus transit peptide addition to the protein to be exported, and on a specific multi-protein transporter located in the plastid membrane [9,10]. Finally, some regulatory and feedback pathways were implemented to regulate the organelle activity as well as protein import [11,12]. All these features are shared and processed by homologous genes in all three Archaeplastida lineages, indicating that all of them were tuned in their common ancestor before diversification. Many other eukaryotic phyla acquired the ability to perform photosynthesis by "recycling" the primary plastid by secondary endosymbioses, which involve a heterotrophic host and an Archaeplastida symbiont (either a green or a red alga). Red alga-like secondary plastids are found in cryptophytes, haptophytes, stramenopiles and alveolates while euglenids and chlorarachniophytes bear green alga-like plastids. There are also cases of tertiary endosymbiosis, where the symbiont is a photosynthetic species derived from secondary endosymbiosis; for instance, some dinoflagellates obtained their plastid from cryptophytes or haptophytes. The number of secondary and tertiary endosymbioses that occurred during the evolution of eukaryotes is still strongly debated (for a complete review, refer for example to [13]).

Thought it seems very likely that events similar to primary endosymbiosis happened during the evolution of eukaryotes, none have left any observable living lineage. The sole comparable case of endosymbiosis involving a heterotrophic eukaryote and a cyanobacterium is observed in the cercozoan genus *Paulinella* [14]. *Paulinella chromatophora* is a testate

* Corresponding author. Email: philippe.deschamps@u-psud.fr
Handling Editor: Andrzej Bodył

This is an Open Access article distributed under the terms of the Creative Commons Attribution 3.0 License (creativecommons.org/licenses/by/3.0/), which permits redistribution, commercial and non-commercial, provided that the article is properly cited. © The Author(s) 2014 Published by Polish Botanical Society

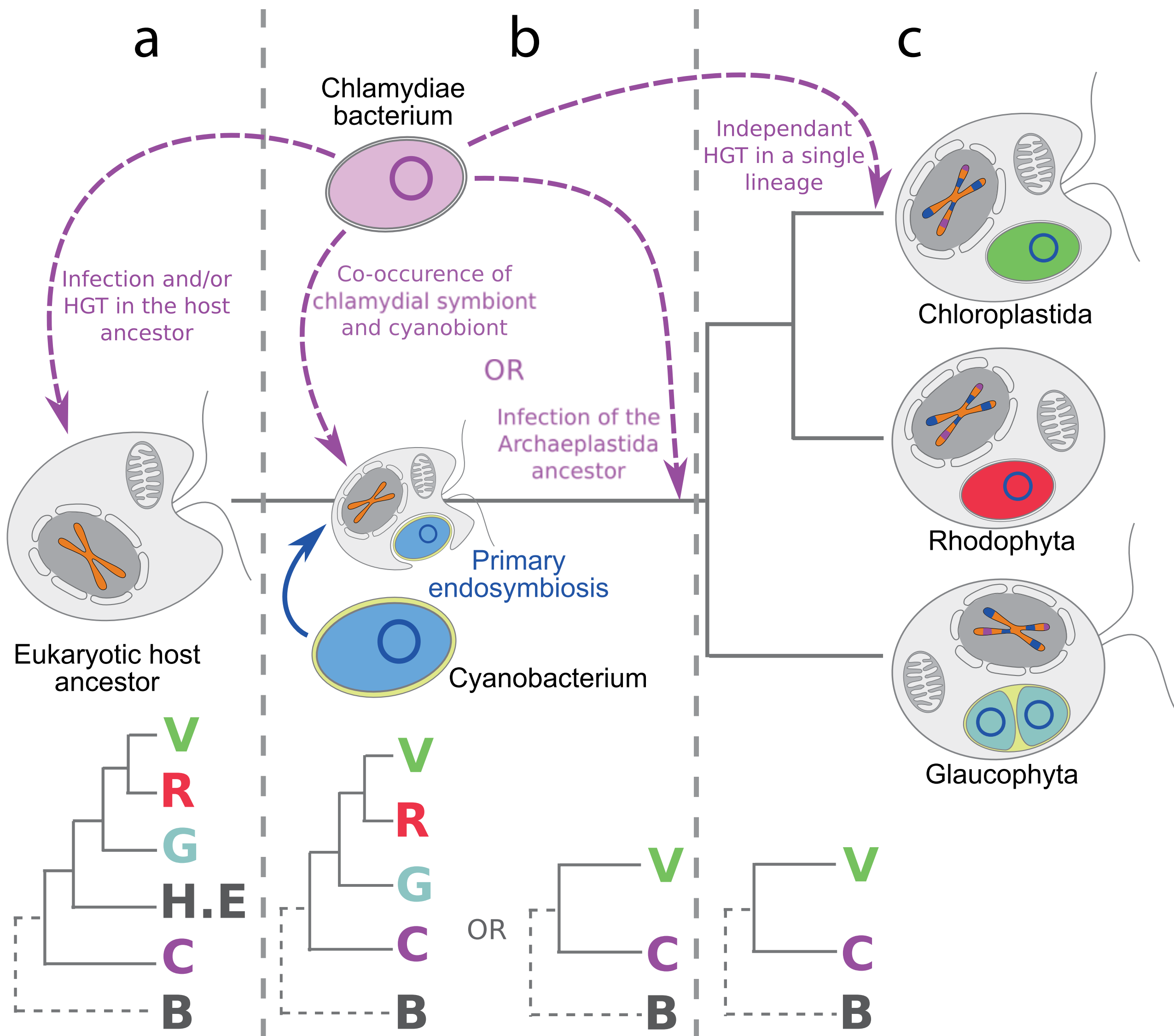

Fig. 1 Time-line of the evolution of Archaeplastida, from their common heterotrophic ancestor to extant Archaeplastida lineages. This line is split in three parts corresponding to three "moments" when HGT from Chlamydiae could have occurred: in the host ancestor before primary endosymbiosis (**a**); in the common ancestor of Archaeplastida, after the engulfment of the cyanobiont but before their diversification (**b**); in a single lineage after the diversification of Archaeplastida (**c**). For each of these moments, the corresponding theoretical phylogenetic tree supporting a HGT from Chlamydiae is displayed. V – Viridiplantae; R – Rhodophyta; G – Glaucophyta; H.E – Heterotrophic eukaryotes; C – Chlamydiae; B – Other bacteria.

amoeba-like organism carrying symbionts called "chromatophores" that are related to *Synechoccocus*/*Prochlorococcus* cyanobacteria and were probably acquired ~60 million years ago. *Paulinella chromatophora* cells display a lesser level of cytological integration compared to the plastids of Archaeplastida [15]. Indeed, very few endosymbiotic gene transfers could be detected in the nuclear genome of *Paulinella* (0.3% ~ 0.8% of the genome) [16] and no targeting system comparable to the one found in Archaeplastida seems to exist, even if sequences similar to targeting peptides have been detected [17]. Nowack and Grossman have speculated that proteins may use the Golgi secretory pathway for trafficking toward the chromatophores [18]. Finally, there is so far no description of regulation/feedback processes between the host and the chromatophores [16]. Thereby, primary endosymbiosis of the Archaeplastida plastid remains a unique event with respect to the conversion of the symbiont into a genuine organelle, and to the diversification and spreading of the descent phyla over the planet. Thus, one can wonder what were the specific circumstances that made this event more successful than any other.

Genome mosaicism in Archaeplastida

Like for any other eukaryote, Archaeplastida genomes are mosaics of genes of various prokaryotic origins [19,20]. Some are related to archaeal genes and support functions related to the maintenance, replication and expression of the genome [21,22]; some others are related to alphaproteobacterial genes and are mainly involved in operating the mitochondria [23]. These components are a testimony of the

involvement of the corresponding prokaryotes in shaping the genome, if not the cell of the ancestor of all eukaryotes [2,24–26]. Additionally, many enzymes involved in central metabolism are encoded by genes acquired from a wide range of bacterial sources [23,27,28] probably via horizontal gene transfers (HGT). Finally, a collection of genes is restricted to eukaryotes and represent specific innovations of this domain. In addition to that common set, Archaeplastida (and derived photosynthetic eukaryotes) carry genes of other prokaryotic origin. The majority of them are related to cyanobacterial genes and were acquired via endosymbiotic gene transfers (EGT) [7,8] during primary endosymbiosis. Previous studies have assayed the amount of cyanobacterial genes in Archaeplastida and ended up with results ranging from 10 to 20% of the nuclear genome [29–32]. Many of these genes encode plastid-located enzymes, but a significant part of them were alternatively relocated to other compartments of the cell, providing an incredible source of metabolic reshuffling and innovation. Whole genome surveys of genes having prokaryotic homologs as closest relatives in phylogenetic trees (a proxy of their evolutionary origin) [32] provide a mixed estimate of the Archaeplastida genome mosaic composition, which is composed of HGT potentially shared by all eukaryotes as well as H/EGT restricted to Archaeplastida species. To have a more precise view on the H/EGT that specifically contributed to Archaeplastida genomes, an interesting yet limited strategy is to apply the same survey procedures to plastid proteomes (excluding plastid-encoded proteins). This analysis, conducted several times independently, led to comparable results where cyanobacteria were the main contributors to plastid proteomes, followed by proteobacteria and Chlamydiae, as well as many other bacterial phyla [33–35]. While we start to understand how EGT from the cyanobiont have physically happened [36], it is still difficult to determine the chronology of their (probably continuous) occurrence in the interval between the engulfment of the cyanobiont and the diversification of Archaeplastida. Similarly, we know almost nothing about the mechanism and timing underlying HGT from other bacterial sources [27,37].

The intriguing phylogenetic link between plants and Chlamydiae

The Chlamydiae is a bacterial phylum composed of obligate intracellular specialized parasites and/or symbionts of eukaryotic cells. The whole group is considered to have diverged between one and two billion years ago [38,39]. The main order of the phylum, the Chlamydiales, is composed of the Chlamydiaceae (animal pathogens) and of several groups of symbionts of free-living Amoebozoa (for a complete review on Chlamydiae, refer to [40]). All Chlamydiae share a biphasic life cycle where one cell type is specialized in the infection process (the elementary body), and another one (the reticulate body) is dedicated to hijacking the host metabolism and producing daughter cells in an inclusion vesicle for future infection. Chlamydiae were believed to be "energy parasites" because of their ability to uptake ATP from their host. For this purpose, they use an ATP/ADP exchanger which is homologous to the one found in Rickettsiales (another group of intracellular parasitic bacteria), Microsporidia and photosynthetic eukaryotes [41]. However, this view has been challenged by the observation of a potentially active ATP synthesis complex encoded in the genome of some Chlamydiae species [40]. All Chlamydiae also possess a large repertoire of transporters able to import nucleotides [42], glucose-6-phosphate [43], NAD or amino acids [44]. Finally, all Chlamydiae have the ability to secret proteins into the cytoplasm of their host [40]. For this purpose, they use a Sec-independent protein secretion pathway, called the type III secretion system (TTSS), composed of about 20 components. The exported proteins, called effectors, are responsible for hijacking the central metabolism of the host cell in order to redirect certain metabolites toward the symbiont/parasite.

Starting from the 90's, several publications have reported the existence of close phylogenetic relationships between plant genes and chlamydial genes [45–48]. These observations were first interpreted as HGTs from plants towards infectious bacteria, because Chlamydiae were sister group to plants or were nested within plants in phylogenetic trees [49]. The latter topology was probably induced by low species sampling in gene trees (restricted to Chlamydiaceae parasites) and to the very long branches of their genes due to an acceleration of their evolutionary rate (like in many parasites). Brinkman et al. [48] in 2002 observed that many of these genes were also present in cyanobacterial genomes and/or encode plastidial functions. They proposed that these genes were a testimony of an ancestral relationship between Cyanobacteria, Chlamydiae and the ancestor of the chloroplast. New genome data, in particular from "environmental" Chlamydiae [39,50] contributed to detect additional genes related to plants and to enhance the resolution of the phylogenetic trees. Indeed, Chlamydiae symbionts of Amoebozoa have larger genomes than those of the species parasitizing animals, as well as genes displaying a smaller evolutionary rate [50]. Recent publications have reported between 20 and 55 genes relating plants and Chlamydiae [45–47,51]. Gene trees supporting this relationship recover plants and Chlamydiae as sister groups or, when the gene is present in other photosynthetic lineages, Archaeplastida is found as the closest group to Chlamydiae.

Many of these genes encode important enzymes located in the plastid (involved in lipid and amino-acid synthesis or tRNA synthesis) as well as many transporters, including the ADP/ATP transporter mentioned above, which is found in the plastid membrane of all photosynthetic eukaryotes. It has been demonstrated that, at least in plants, this transporter is involved in fueling enzymatic reactions in the dark, when the plastidial production of ATP is lower or disabled. This includes mechanisms of resorption of oxidative molecules from the chloroplast stroma [52] and mutants lacking this transporter suffer and die from oxidative stress. Given that several essential genes for the functioning of the plastid were acquired from Chlamydiae, many authors hypothesized that, maybe, primary endosymbiosis that led to Archaeplastida was successful thanks to a co-infection of the host by a Chlamydiae bacterium.

A little help from Chlamydiae to the success of primary endosymbiosis?

Huang and Gogarten reported in 2007 that some genes of putative chlamydial origin found in green plants were also shared among all Archaeplastida and that their phylogeny recovered the monophyly of the whole supergroup [46]. They concluded that HGT from Chlamydiae occurred in the common ancestor of Archaeplastida, before or after the recruitment of the cyanobiont. Moreover, they argued that the relatively high amount of HGT from Chlamydiae in genomes of Archaeplastida compared to other bacterial sources (excluding Cyanobacteria) was a testimony of a long period of close relationship between these two partners, probably of endosymbiotic nature considering the intracellular lifestyle of Chlamydiae. Additionally, they could not detect any of these genes in heterotrophic eukaryotes, indicating that this relationship did not or only shortly predated primary endosymbiosis. Using these observations, they hypothesized that primary endosymbiosis happened with a host that was already in an endosymbiotic relationship with an ancient Chlamydiae that resembled a modern *Protochlamydia*. The presence of a Chlamydiae symbiont would have allowed the recruitment of functions crucial to the long-term establishment of the cyanobiont, rather than its digestion. In the hypothesis of Huang and Gogarten, their view of the Chlamydiae symbiont is that of a genuine long lasting endosymbiont, providing a large amount of EGT, more than what is detectable in modern Archaeplastida genomes. For them, the loss of the symbiont provoked the loss of most of the related EGT, like what is observed in eukaryotes with a reduced mitochondria or reverting from phototrophy to heterotrophy.

Moustafa et al. contributed in 2008 to this issue with a re-evaluation of the number of HGT from Chlamydiae into Archaeplastida genomes using updated genome data [47]. Though they agreed with Huang and Gogarten on the pivotal involvement of Chlamydiae gene transfers in the success of primary endosymbiosis, they proposed an alternative interpretation of the phylogenies regarding the timing of these gene transfers. First, they observed that some HGTs from Chlamydiae were detected not in all Archaeplastida but only in red alga or in glaucophytes. This can be explained with almost the same likelihood, either by common gains in the common ancestor of Archaeplastida followed by lineage specific gene losses or by independent acquisitions in a single lineage. Moreover, they pointed out that many bacterivorous eukaryotes harbor prokaryotic genes of (probable) HGT origin in their genomes [53]. With that in mind, they considered that a scenario based on multiple HGT from Chlamydiae in a common ancestor of Archaeplastida couldn't be completely excluded. In their model, the phagotrophic ancestor of Archaeplastida captured a cyanobacterium and settled a "metabolic connection" allowing to switch to a mixotrophic lifestyle. After this event, and for a certain period, this new mixotrophic cell "continuously" acquired genes from various preys or parasites including Chlamydiae. At a specific point, the set of genes that were available in the nucleus became sufficient to start a new form of endosymbiotic relationship with the cyanobiont, leading to the evolution of the plastid.

The same year, Becker et al. published their own evaluation of Chlamydiae contribution to Archaeplastida genomes and discussed how gene phylogenies can help inferring if the presence of a Chlamydiae symbiont predated, co-occured of was subsequent to the engulfment of the cyanobiont during primary endosymbiosis [45]. Part of their discussion focused on genes having different origins between Archaeplastida lineages, for instance functions for which red algae and glaucophytes use a gene of cyanobacterial origin whereas green algae and plants use a homolog of Chlamydiae origin. They noted that there seems to be an exclusion rule applying on these HGT paralogs, because no cases where both paralogs were kept were observed. For these authors, the existence of such paralogy cases of HGT/EGT is a clue that the cyanobiont and Chlamydiae symbiont co-existed for a long period during the evolution of Archaeplastida.

In 2013, Ball et al. inspired by these previous publications, also provided their own evaluation of the contribution of Chlamydiae genes to Archaeplastida genomes and proposed a functional hypothesis for the interaction of the Chlamydiae symbiont, the cyanobiont and the host in a tripartite model of primary endosymbiosis [51]. This tripartite model has recently spread around several groups working on the subject and is used as a framework to infer additional evolutionary hypotheses on primary endosymbiosis.

Metabolic symbiosis and primary endosymbiosis

Given the high number of heterotrophic eukaryotes feeding by phagotrophy on cyanobacteria or using them as temporary symbionts, one can argue that opportunities to settle obligate symbiotic relationships are frequent. Shifting to an obligate relationship probably starts with a strong compromise where each partner "subcontracts" a metabolic service in exchange of one other to the point of being unable to process it and become definitely dependent. Thus, the limitation to start a dynamics of primary endosymbiosis resides on the kind of mutual services one cyanobacterium and one eukaryote can provide to each other, but also on the ability to settle an exchange of the specific metabolites upon which the symbiosis resides. The actual structure and localization of metabolic pathways common to all Archaeplastida, as well as the nature and phylogenetic origin of their components can provide hints to understand the nature of the metabolic symbiosis that could have definitely settled the primary endosymbiosis.

In 2006, Deschamps et al. proposed a hypothesis on the nature of the metabolic symbiosis that originally associated the cyanobacterial ancestor of the plastid and its host [54]. The model resides on the fact that one of the properties that emerged from primary endosymbiosis is the substitution of glycogen by starch to store glucose. Glycogen and starch share a similar molecular structure but have different physical properties. The chemical advantage of starch is that it can store more glucose in a single cell because, as opposed to soluble glycogen, insoluble starch granules do not account to osmotic pressure anymore. Every eukaryotic species known for synthesizing starch is either photosynthetic or derive from a photosynthetic ancestor, meaning that

starch metabolism is probably an innovation that occurred together with the acquisition of photosynthesis. All three lineages of Archaeplastida synthesize starch, either in the cytoplasm (Glaucophyta and Rhodophyta) or in the chloroplast (green algae and plants) using components found separately in cyanobacteria and eukaryotes [54]. The survey of all enzymes involved in starch metabolism in all three Archaeplastida lineages allows to infer the most probable composition of this pathway in the common ancestor of all Archaeplastida. Considering the actual cellular localization of the starch pathway in extant Archaeplastida, it was most probably located in the cytoplasm in their common ancestor. Phylogenetic analyses of the protein sequences of these enzymes showed that they were recruited from the glycogen synthesis pathways of both the cyanobiont and the host. It seems that primary endosymbiosis, and the accompanying EGT flow, have for the first time gathered all genes of both pathways in the same genome. These genes, expressed into the cytoplasm of the host, arranged a new hybrid pathway from which starch emerged, offering a high storage capacity for the carbon fixed by the new photosynthetic symbiont.

A second observation inspiring the hypothesis of Deschamps et al. [54] is the discovery of starch-like structures in unicellular group V diazotrophic cyanobacteria [55]. These starch-like molecules are synthesized using enzymes that are phylogenetically related to the ones involved in starch metabolism in Archaeplastida, and have not been found in any other group of Cyanobacteria so far. For this reason, Deschamps el al. proposed that unicellular diazotrophic cyanobacteria were possibly the ancestors of primary plastids. The fixation of nitrogen in diazotrophic cyanobacteria involves an oxygen-intolerant nitrogenase [56]. Filamentous cyanobacteria use specialized cells called heterocysts to physically isolate nitrogen fixation from oxygenic photosynthesis [57]. In unicellular diazotrophic cyanobacteria nitrogen fixation only occurs in the dark when photosynthesis is deactivated. For these unicellular cyanobacteria, becoming an endosymbiont offered an opportunity to acquire nitrogen directly from the host, thus eliminating diazothrophy and breaking the incompatibility with oxygenic photosynthesis. Thus, the first hypothesis of Deschamps and coworkers was that primary endosymbiosis was settled by a metabolic symbiosis based on the exchange of carbon against nitrogen [54]. The nature of the metabolite used for exporting glucose from the cyanobiont was proposed to be ADP-glucose, a specific precursor restricted to starch biosynthesis in the chloroplast of green algae and land plants and to glycogen synthesis in bacteria. Red algae and glaucophytes build starch from UDP-glucose, a widely spread substrate of many other enzymatic pathways requiring glucose. Starch synthesis in the cytoplasm of the common ancestor of Archaeplastida probably used both precursors, because if not, ADP-glucose specific enzymes acquired by EGT from the cyanobiont, encoded in the nucleus of green algae and plants and active in their chloroplast, would have never been positively selected until their diversification. Thus, the model of Deschamps et al. [54] is that ADP-glucose was synthesized in the cyanobiont using a cyanobacterial ADP-Glc pyrophosphorylase and exported toward the cytoplasm by an unknown transporter and used for starch accumulation. Glucose could then be mobilized again from this stock for other uses. The hypothesis on the reasons that could have subsequently selected for a relocation of starch metabolism in the chloroplast of Viridiplantae is described in references [58,59].

In 2010, Colleoni and coworkers proposed that a nucleotide sugar transporter (NST) already present in the eukaryotic ancestor of Archaeplastida, could have been involved in the export of ADP-glucose from the plastid [60]. This transporter is a putative ancestor of a modern protein, shared by all Archaeplastida, specific for GDP-mannose, but also able to transport ADP-glucose at lower rates and with a lower affinity. This observation provides additional likelihood for a metabolic symbiosis based on the export of ADP-glucose.

Ménage à trois

One weakness of the metabolic symbiosis hypothesis presented above is that the described metabolic link cannot be available immediately after the engulfment of the cyanobiont. First, it requires the efficient integration of a protein able to transport ADP-glucose through the double membrane of the cyanobiont. This assumes that the phagotrophic membrane that was probably originally present around the symbiont had already vanished. Secondly, the biochemical link requires at least one EGT event that would transfer a gene coding for an ADP-glucose specific glycogen synthase to the nucleus of the host and activate its expression. Given the stochasticity of the EGT process, this would require a certain time before such a gene gets indeed expressed into the cytoplasm of the host. Thus, how were the symbiont and the host interacting together before that? What would have prevented the digestion of the cyanobiont if a direct benefit for both partners was not instantly available?

In 2013, Ball et al. published a revision of the metabolic symbiosis model of primary endosymbiosis called the "ménage à trois" [51]. This new version assumes the simultaneous infection by a Chlamydiae symbiont together with the engulfment of the cyanobiont in the host of primary endosymbiosis. The advantage of this new model is that there is no need to wait for an EGT event to start consuming ADP-glucose in the host cytoplasm. Indeed, Ball et al. proposed that some effectors secreted by a Chlamydiae symbiont could have provided essential functions in the host cytoplasm to help priming the metabolic symbiosis with the cyanobiont. For instance, it has been recently shown that *Chlamydia trachomatis* exports enzymes of the glycogen synthesis pathway [61] provoking a change in the levels of glucose and glycogen production in the host. Ball et al. provide additional experimental evidence that many enzymes involved in glycogen metabolism in Chlamydiae are secreted in the host cytoplasm [51]. They also observe that two enzymatic components of starch metabolism in plants (an ADP-glucose specific starch synthase, GlgA, and an isoamylase, GlgX) are in fact derived from HGT of Chlamydiae genes. Given these new information, Ball et al. propose that Chlamydiae helped the settlement of metabolic symbiosis between the cyanobiont and its host by providing, directly through the secretion system, the necessary enzymes to reorganize the

host glycogen synthesis pathway, allowing the use of ADP-glucose and motivating its export from the cyanobiont. Later on, genes of the glycogen metabolism of Chlamydiae and cyanobacteria would have been horizontally transferred to the nucleus and used to start accumulating starch.

At the exact same period, Facchinelli et al. published a proteomic analysis of the muroplast of the glaucophyte algae *Cyanophora paradoxa* [62]. This analysis unveils a potential new route for exporting glucose from the cyanelle to the cytoplasm using a transporter of glucose-6-phosphate (G6P). In bacteria, the uhp operon encodes four genes involved in the "uptake of hexose phosphates" [63]. The UhpC protein usually serves as a sensor, except in Chlamydiae where it became able to translocate G6P [43]. A homolog of *uhpC* was found by Price et al. in the genome of *C. paradoxa* [64], and Facchinelli et al. describe the presence of the UhpC protein in the muroplast of this species [62]. The phylogeny of this gene, available in reference [65], shows that it is present in the genome of glaucophytes, red algae and green algae and that it was probably acquired by HGT from Chlamydiae in the common ancestor of Archaeplastida. All these observations, together with the absence of phosphate transporters of the NST family in *Cyanophora*, prompted Facchinelli et al. [62] to speculate that UphC may have been the first protein used to export carbon from the cyanobiont, making it directly available for glycogen synthesis in the host. Thus, the recruitment of an ADP-glucose transporter related to NST and the selection for an ADP-glucose based carbon storage metabolism would have only happened after the split of Glaucophyta from the other Archaeplastida lineages.

Proposing a scenario where there is no selection on ADP-glucose specific enzyme in the common ancestor of extant Archaeplastida questions the entire validity of the metabolic symbiosis hypothesis of Deschamps et al. [54]. This is also incompatible with the apparent presence of an ADP-glucose specific starch synthase in *Cyanophora paradoxa* [64] that must have been positively selected before the divergence of glaucophytes. Facchinelli and Ball published a joined paper to formulate a revision of the "ménage à trois" scenario compatible with all the recent data and restoring a putative selection on ADP-glucose based carbon storage metabolism [66]. This latest version assumes that the cyanobiont would have been integrated in the inclusion vesicle of the reticulate body of the Chlamydiae co-symbiont. According to Facchinelli et al. [66] G6P is exported out of the cyanobiont to the intravesicular space thanks to the UhpC protein encoded in the cyanobiont genome after a single HGT event from the Chlamydiae symbiont, possibly via the type IV secretion system (another kind of secretion system able to export proteins of protein-DNA complexes). G6P is then converted to ADP-glucose and stored into glycogen particles in the intravesicular space. Extra ADP-glucose is exported from the inclusion vesicle thanks to the NST-like transporter that would be installed there using the host Golgi secretion pathway. ADP-glucose is finally used by TTSS secreted enzymes of the chlamydial glycogen pathway, originally there for hijacking the host metabolism for parasitism purposes. Later on, after several EGT events, components of starch metabolism would have replaced the TTSS secreted chlamydial enzymes.

All these hypothetical models are based on very strong assumptions that will not be discussed in detail here. The main supposition is that Chlamydiae and cyanobacteria did interact in the same host at the same time during primary endosymbiosis. Are there any clues indicating that?

What is the actual genomic contribution of Chlamydiae to Archaeplastida?

At least four independent genome wide detections of potential HGT from Chlamydiae to Archaeplastida have been made available in the literature: Huang and Gogarten listed 21 putative HGTs [46], Becker et al. listed 39 putative HGTs [45], Moustafa et al. listed 55 putative HGTs [47] and Ball et al. listed 48 putative HGTs [51]. Even if the latest analysis dates back to 2013, a significant amount of new genomes/transcriptomes were released since then, for instance in reference [67]. To have an updated view of the contribution of Chlamydiae to Archaeplastida, I decided to compile all genes reported in these 4 publications and to reanalyze their phylogeny using the most complete available sample dataset. Merging the four lists ended up with a non-redundant set of 86 candidate genes possibly acquired by HGT. The procedure for this analysis was as follows: (***i***) the reference protein indicated for each reported HGT was used as a query for similarity search using BlastP [68] against a local custom protein database; (***ii***) the top Chlamydiae and Archaeplastida BlastP hits were extracted from the database and used as queries for Net-BlastP searches against the RefSeq protein database [69] on NCBI servers; (***iii***) for each HGT candidate, the 300 first hit sequences were extracted from the local BlastP as well as the 800 first hit sequences from the two Net-BlastP searches. All sequences were merged in a single file and duplicates automatically discarded; (***iv***) a first round of multiple alignments (MAFFT default parameters [70]), trimming (BMGE, default parameters [71]) and phylogenetic tree reconstructions (FastTree [72]) was processed. Every tree was manually inspected to extract the subset of sequences corresponding to the part of the tree centered on Chlamydiae; (***v***) a second round of tree reconstruction was done on this reduced dataset using the same software tools except for maximum likelihood (ML) tree inferences that were computed with TreeFinder [73] using the WAG+Γ substitution model. Final trees are available in Newick and PDF formats in the supporting material.

All 86 ML trees were inspected to determine if their topology was compatible with a HGT acquired from Chlamydiae and if the topology was consistent with a transfer toward the common ancestor of Archaeplastida. A summary of this analysis is available in Tab. 1. Ten gene phylogenies showing a clear vertical descent in eukaryotes were discarded because not compatible with a HGT event. Moreover, 7 trees were insufficiently sampled to allow any interpretation, and in 18 other trees the topology was compatible with a HGT but the closest sister group to Archaeplastida was not Chlamydiae but other bacterial groups. This later observation is in accordance with previous estimations of the composition of the protein repertoire of plastids [33–35].

Fig. 1 depicts a time-line of the evolution of Archaeplastida, from their common heterotrophic ancestor to their diversification in three lineages. On this figure are also indicated the kind of gene tree topologies that would support a HGT from Chlamydiae at certain points of this time-line. All topology compatible with a HGT event from Chlamydiae toward the common ancestor of Archaeplastida must recover the monophyly of Archaeplastida and show Chlamydiae as its closest bacterial sister group. If any non-photosynthetic eukaryotes (i.e. eukaryotes known for having no photosynthetic ancestors) branch between the Chlamydiae and the Archaeplastida, then one can infer that this gene was acquired before primary endosymbiosis (Fig. 1a). Eleven trees showed this topology, and in almost all cases the non-photosynthetic eukaryotes include Amoebozoa, a phylum known for often carrying Chlamydiae parasites. These trees can be interpreted in three different ways: either Amoebozoa and Archaeplastida acquired independently the same gene by HGT from Chlamydiae; or these genes are ancestral in eukaryotes and were transferred to the common ancestor of Chlamydiae; or, finally, the common ancestor of Amoebozoa and Archaeplastida acquired this gene from Chlamydiae, and it was subsequently lost in the whole descent except in these two groups. There is no way to clearly favor one these three scenarios, and thus these genes should not be counted as cases of direct chlamydial contribution to Archaeplastida genomes. As a matter of fact, proteins encoded by all these genes are reported to be located in the cytoplasm in *Arabidopsis thaliana* [74]. This is evidence that HGT of Chlamydiae origin can be positively selected in eukaryotes in the absence of a plastid. This fact was already mentioned by Thiergart et al. in 2012 [27].

Gene trees where Chlamydiae were closely related with primary photosynthetic eukaryotes but where Archaeplastida were not monophyletic were all compatible with a replacement event of the same function by genes of different origins (6 cases). There is no way to determine in what order these consecutive HGT occurred.

Trees compatible with HGT from Chlamydiae and where species of 2 or 3 Archaeplastida lineages were present (Fig. 1b) are the most consistent with an HGT during the time of endosymbiosis (24 cases). Nonetheless, these trees give no additional clues on when exactly each HGT happened during the possibly long lasting period between the initial engulfment of the cyanobiont and the diversification of Archaeplastida.

Finally, trees compatible with HGT from Chlamydiae and where only one Archaeplastida lineage was present can lead to two interpretations (10 cases). Indeed, there is an equal probability that these genes were acquired before the diversification of Archaeplastida and differentially lost (Fig. 1b), or that they were acquired independently in one lineage after diversification (Fig. 1c).

The above phylogenetic analyses suggest that a continuous flow of HGT from Chlamydiae took place during the evolution of eukaryotes. Some occurred before primary endosymbiosis, others happened in the time between the engulfment of the cyanobiont and the diversification of Archaeplastida, and some others probably took place in only one lineage after their diversification. This demonstrates that HGT from bacteria, including the ones from Chlamydiae should be seen as a "strafe" process, rather than a "one shot" event. Considering the lifestyle of extant Chlamydiae, we can speculate that most of these transfers are linked to periods of parasitism or symbiosis of the Chlamydiae donor with the respective eukaryotic recipient. This is nonetheless not mandatory, as HGT could have happened without internal contact. Indeed, many cases of HGT from diverse bacterial sources toward eukaryotes exist, in photosynthetic species [37] but also in heterotrophic ones [75,76], and do not necessarily imply endosymbiosis or parasitism. If we focus only on HGT events that seem to have happened during primary endosymbiosis, as we stated above, molecular phylogenies (if interpreted sensu stricto) cannot help us inferring when exactly these HGT occurred.

At this point, alternative and more mechanistic arguments can be invoked. For instance, one can argue that genes acquired by HGT are only kept if their function is positively selected at the time they start being expressed by the recipient organism. Because many of the described Chlamydiae HGTs encode plastidial functions, it was stated that their acquisition cannot predate primary endosymbiosis, but must be concomitant or subsequent to it. This is true only for genes that do not have any functional equivalent already encoded in the recipient genome. HGT replacing or duplicating existing functions can be positively selected during the evolution of a lineage, at least for a certain time. Many of the genes listed in Tab. 1, are involved in central or informational metabolisms that exist both in the cytoplasm and in the plastid. Thus, they do not necessarily need the co-occurrence of a cyanobiont/plastid to be positively selected but could have been selected for a cytoplasmic role in the non-photosynthetic ancestor of Archaeplastida and then converted to the same role in the plastid after primary endosymbiosis. Functional or structural constraints are also often invoked when trying to justify why a gene of a specific donor was selected after HGT in place of another one from a different origin. This is a particularly interesting question in a context where chlamydial and cyanobacterial endosymbionts could have provided a lot of homologous gene at the same time, or asynchronously, toward the same recipient genome. In my analysis, I detected 6 cases of gene replacement indicating that not every Archaeplastida encodes the same function with the same gene of a unique HGT origin. This observation illustrates that the selection applying on a specific function can be relaxed with respect to the origin of the gene encoding it. Thus arguments of functional constraints must always be used with caution.

Chlamydiae as helpers of the cyanobiont during primary endosymbiosis

There is no doubt that a set of genes acquired by HGT from Chlamydiae is mandatory to the functioning of plastids in extant Archaeplastida. There is also a strong indication that some of these HGTs probably occurred from a Chlamydiae symbiont in an ancestor of Archaeplastida. These two points do not imply that a Chlamydiae symbiont was necessarily present during the very first stage of primary endosymbiosis

Tab. 1 Summary of the phylogenetic reanalysis of genes reported as acquired by horizontal gene transfer from Chlamydiae to Archaeplastida by Huang and Gogarten [46], Ball et al. [51], Moustafa et al. [47] and Becker et al. [45].

	Huang 2007	Moustafa 2008	Becker 2008	Ball 2013	Ch	Ro	Gl	Topology interpretation	Putative function	Chlamydial ref. gene
1	•	•	•	•	•	•	-	**HGT from Chlamydiales**	**3-Oxoacyl-(acyl-carrier-protein) synthase**	**WP_006340838**
2	-	•	-	•	•	•	•	Eukaryotic gene	3-Phosphoadenosine 5-phosphosulfate 3- phosphatase	WP_021757506
3	-	•	-	-	•	-	-	Chlamydiales not closest	Mg^{2+}-dependent DNA exonuclease	WP_006340671.1
4	•	•	-	-	•	•	•	Chlamydiales not closest	Phosphoribosylanthranilate isomerase	WP_013943364.1
5	-	•	-	•	•	•	•	**HGT from Chlamydiales**	**Cytosine/adenosine deaminases**	**WP_011175198.1**
6	-	•	•	•	•	-	•	**HGT from Chlamydiales**	**Predicted sulfur transferase**	**WP_011174928.1**
7	•	•	-	-	•	•	•	**HGT from Chlamydiales**	**Sodium:hydrogen antiporter 1**	**WP_006340811.1**
8	-	•	•	•	•	•	•	N.P. Euka are present	tRNA d(2)-isopentenylpyrophosphate transferase	WP_011175793.1
9	•	•	•	•	•	•	-	**HGT from Chlamydiales**	**Malate/lactate dehydrogenases**	**WP_011176317.1**
10	-	•	•	-	•	-	-	HGT from Chlamydiales	F-box and associated interaction domains-containing protein	WP_011175702.1
11	•	•	•	-	•	•	-	Archaeplastida polyphyletic	Chloroplast polynucleotide phosphorylase	WP_011175193.1
12	-	•	•	•	•	•	-	**HGT from Chlamydiales**	**S-adenosylmethionine-dependent methyltransferase**	**WP_011174861.1**
13	-	•	•	•	•	-	-	Archaeplastida polyphyletic	Involved deoxyxylulose pathway of isoprenoid biosynthesis	WP_011175290.1
14	•	•	•	•	•	•	•	**HGT from Chlamydiales**	**ATP/ADP antiporter**	**YP_007249.1**
15	-	•	•	•	•	-	•	**HGT from Chlamydiales**	**Glycosyltransferase family A (GT-A)**	**WP_011174874.1**
16	•	•	•	-	•	•	•	**HGT from Chlamydiales**	**L,L-diaminopimelate aminotransferase**	**WP_011175235.1**
17	•	•	•	•	•	•	•	**HGT from Chlamydiales**	**isoamylase**	**WP_011175656.1**
18	•	•	-	•	•	•	•	**HGT from Chlamydiales**	**Cation transport ATPase**	**WP_006341401.1**
19	•	•	•	•	•	•	•	**HGT from Chlamydiales**	**4-Diphosphocytidyl-2C-methyl-D-erythritol 2-P synthase**	**WP_011176135.1**
20	-	•	•	•	•	•	•	**HGT from Chlamydiales**	**SAM-dependent methyltransferases**	**WP_011176541.1**
21	-	•	-	•	•	•	•	Eukaryotic gene	tRNA-dihydrouridine synthase	YP_008986.1
22	-	•	-	-	-	-	•	Chlamydiales not closest	Putative carbonic anhydrase	WP_011175608.1
23	-	•	-	•	•	•	•	**HGT from Chlamydiales**	**Anthranilate phosphoribosyltransferase**	**WP_013943362.1**
24	-	•	-	-	•	•	•	Eukaryotic gene	Inosine triphosphate pyrophosphatase family protein	WP_011175160.1
25	-	•	-	-	•	•	•	Eukaryotic gene	Glycine-rich RNA-binding protein 4	WP_011175443.1
26	-	•	-	-	-	-	-	Very poor sampling	Glycosyl transferase GT2	WP_011176127.1
27	-	•	-	-	•	•	-	Eukaryotic gene	tRNA guanine N7 methyltransferase	WP_011175834.1
28	•	•	-	•	-	•	-	HGT from Chlamydiales	Uncharacterized conserved protein	WP_011176253.1
29	•	•	•	•	-	•	•	**HGT from Chlamydiales**	**Oligoendopeptidase F**	**WP_011175728.1**
30	•	•	•	•	•	•	•	N.P. Euka are present	CMP-2-keto-3-deoxyoctulosonic acid synthetase	WP_011175375.1
31	-	•	-	-	•	•	-	Chlamydiales not closest	Glycosyl transferase	WP_011175669.1
32	•	•	•	•	•	•	•	**HGT from Chlamydiales**	**glycerol-3-phosphate acyltransferase**	**WP_011175867.1**
33	•	•	•	•	•	-	-	HGT from Chlamydiales	Phosphoglycerate mutase 1	WP_011174711.1
34	-	•	-	-	•	-	-	Chlamydiales not closest	Putative dimethyladenosine transferase	WP_011174945.1
35	-	•	-	-	•	•	•	Chlamydiales not closest	S-adenosyl-L-methionine-dependent methyltransferase	WP_011176166.1
36	-	•	-	-	•	-	-	Chlamydiales not closest	Tonoplast intrinsic protein	WP_011175346.1
37	-	•	-	-	•	-	•	**HGT from Chlamydiales**	**Starch synthase**	**WP_011176142.1**
38	-	•	-	-	•	•	•	Eukaryotic gene	Superoxide dismutase [Cu-Zn]	WP_011176304.1
39	-	•	-	-	•	•	-	Chlamydiales not closest	Leucine-rich repeat-containing protein	WP_011174886.1
40	-	•	•	-	•	•	•	N.P. Euka are present	Pyrophosphate-fructose 6-phosphate 1-phosphotransferase	WP_011175430.1
41	•	•	•	•	•	•	•	**HGT from Chlamydiales**	**2-C-Methyl-D-erythritol 4-phosphate cytidylyltransferase**	**WP_011174877.1**
42	-	•	-	-	•	•	•	N.P. Euka are present	UDP-glucuronate 4-epimerase 2	WP_011174629.1
43	-	•	•	•	•	•	•	**HGT from Chlamydiales**	**Uncharacterized conserved protein**	**WP_011175877.1**
44	•	•	-	•	•	-	-	HGT from Chlamydiales	16S rRNA uridine-516 pseudouridylate synthase	WP_011174710.1
45	•	•	-	•	•	-	-	HGT from Chlamydiales	FOG: CBS domain	WP_011176327.1
46	-	•	•	-	•	-	•	N.P. Euka are present	Biotin/lipoate A/B protein ligase family protein	WP_011176383.1
47	-	•	-	-	•	•	•	Eukaryotic gene	Superoxide dismutase [Cu-Zn]	WP_011174820.1
48	•	•	•	•	•	•	-	Archaeplastida polyphyletic	Tyrosyl-tRNA synthetase	WP_011175719.1
49	-	•	•	-	•	•	•	N.P. Euka are present	Queuine tRNA-ribosyltransferase	WP_011175344.1
50	-	•	-	-	•	-	-	Chlamydiales not closest	Ribosomal 5S rRNA binding domain-containing protein	WP_011176140.1
51	•	•	•	•	•	-	-	Archaeplastida polyphyletic	Phosphate/sulfate permeases	WP_011174649.1
52	-	•	-	-	•	-	-	HGT from Chlamydiales	D-alanine-D-alanine ligase family protein	WP_011174948.1
53	-	•	-	•	•	-	-	Eukaryotic gene	rRNA methylases	WP_011174691.1
54	-	-	-	•	•	-	-	Very poor sampling	Predicted ATPase of the PP-loop superfamily	WP_011175189.1
55	-	-	-	•	•	•	•	Archaeplastida polyphyletic	tRNA and rRNA cytosine-C5-methylases	WP_006341520.1
56	-	-	-	•	•	-	-	N.P. Euka are present	Uncharacterized conserved protein	WP_013181531.1
57	-	-	-	•	-	-	•	HGT from Chlamydiales	Aspartokinases	WP_011175315.1
58	-	-	-	•	-	•	-	HGT from Chlamydiales	FKBP-type peptidyl-prolyl cis-trans isomerase 1	WP_011174933.1
59	-	-	•	•	-	•	•	Chlamydiales not closest	Pseudouridylate synthases, 23S RNA-specific	WP_011174905.1
60	-	-	-	•	•	-	•	Archaeplastida polyphyletic	Phosphoenolpyruvate synthase/pyruvate phosphate dikinase	WP_013924500.1
61	-	-	-	•	-	•	•	**HGT from Chlamydiales**	**6-Pyruvoyl-tetrahydropterin synthase**	**WP_013181942.1**
62	-	-	-	•	•	-	-	HGT from Chlamydiales	UDP-N--acetylmuramate dehydrogenase	WP_011176170.1
63	-	-	-	•	•	•	-	Chlamydiales not closest	Predicted SAM-dependent methyltransferases	WP_011175593.1
64	-	-	-	•	•	•	-	Chlamydiales not closest	2-Succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase	WP_013943765.1

Tab. 1 *(continued)*

	Huang 2007	Moustafa 2008	Becker 2008	Ball 2013	Ch	Ro	Gl	Topology interpretation	Putative function	Chlamydial ref. gene
65	-	-	-	•	•	•	-	Archaeplastida polyphyletic	Isochorismate synthase	WP_013943764.1
66	-	-	-	•	-	-	-	Very poor sampling	D-Ala-D-Ala ligase and related ATP-grasp enzymes	WP_013943388.1
67	-	-	-	•	-	•	•	N.P. Euka are present	Membrane-associated lipoprotein involved in thiamine biosynthesis	WP_011176247.1
68	-	-	-	•	-	-	•	HGT from Chlamydiales	Uncharacterized conserved protein	WP_011176413.1
69	**-**	**-**	**-**	**-**	**•**	**•**	**•**	**HGT from Chlamydiales**	**UphC, Glucose-6-phosphate transporter**	**WP_011174937.1**
70	-	-	-	•	-	-	•	N.P. Euka are present	Predicted Rossmann fold nucleotide binding protein	WP_011175131.1
71	**-**	**-**	**-**	**•**	**•**	**•**	**-**	**HGT from Chlamydiales**	**Predicted metal-dependent hydrolase**	**WP_013944685.1**
72	-	-	-	•	-	•	-	Very poor sampling	Predicted O-linked N-acetylglucosamine transferase	WP_013925630.1
73	-	-	-	•	•	•	•	N.P. Euka are present	FOG: PPR repeat	WP_013943667.1
74	-	-	-	•	•	-	-	Very poor sampling	Putative uncharacterized protein	WP_011175444.1
75	-	-	-	•	-	•	-	Chlamydiales not closest	Protein involved in glycerolipid metabolism	WP_020966676.1
76	-	-	•	-	•	•	•	N.P. Euka are present	Asparaginyl-tRNA synthetase	WP_011174896
77	-	-	•	-	•	•	-	Eukaryotic gene	7-dehydrocholesterol reductase	WP_011175770
78	**-**	**-**	**•**	**-**	**•**	**•**	**•**	**HGT from Chlamydiales**	**tRNA-pseudouridine synthase I**	**YP_007681.1**
79	-	-	•	-	•	•	•	Eukaryotic gene	Lon protease 1	WP_011175012.1
80	-	-	•	-	-	-	-	Very poor sampling	DNA mismatch repair protein MutS	WP_011175771.1
81	-	-	•	-	•	•	•	Chlamydiales not closest	Ribosome recycling factor	WP_011176421.1
82	-	-	•	-	•	•	•	Chlamydiales not closest	Tyrosine transporter	WP_011174718.1
83	-	-	•	-	•	•	•	Chlamydiales not closest	50S rRNA methyltransferase	WP_011174889.1
84	-	-	•	-	-	•	•	Chlamydiales not closest	Cysteinyl-tRNA synthetase	WP_011175785.1
85	-	-	•	-	-	-	-	No Archaeplastida	Folylpolyglutamate synthase	WP_011176172.1
86	-	-	•	-	•	•	•	Chlamydiales not closest	Transketolase	WP_011176060.1

The citation of each gene in the three publications is indicated by "•". The presence of Chloroplastida (Ch), Rhodophyceae (Ro) and Glaucophyta (Gl) in each tree is indicated by "•". For each gene, an interpretation of the tree topology is provided, as well as its putative function and a chlamydial protein reference accession number. Bold lines are trees compatible with a HGT from Chlamydiae to the common ancestor of Archaeplastida. N.P. – non photosynthetic.

to help its settlement. There is no solid evidence for that but only speculation. In fact, the settlement of the metabolic relationship between the host and the cyanobiont could have predated the arrival of a Chlamydiae symbiont, and the subsequent HGTs would have only helped to convert the cyanobiont into its actual plastid form by providing a set of important or accessory functions. This alternative point of view is supported by the fact that HGT from Chlamydiae that were selected to function in the plastid can, for the most, only have occurred or been selected after the development of the plastid protein targeting system. This innovation took probably a certain time during which the cyanobiont and its host had to find a way to communicate without the help of chlamydial genes.

This timing problem affects the metabolic symbiosis hypothesis of primary endosymbiosis presented above, because it requires EGT events, either from the cyanobiont or from the Chlamydiae symbiont, to start exporting carbon from the cyanobiont in the form of ADP-Glc or G6P. The last version of the "ménage à trois" hypothesis was designed to prevent this flaw [66]. Proposing that the cyanobiont was included in the same vesicle as the Chlamydiae symbiont brings more opportunity for gene transfers between them. A single transfer of the *uhpC* gene toward the cyanobiont would be enough to make it export G6P in the intravesicular space. G6P will then be converted to ADP-glucose by the chlamydial ADP-glucose pyrophosphorylase and used internally for glycogen synthesis or exported into the cytoplasm. In this version of the model, neither EGT nor HGT toward the host nucleus are needed to settle a metabolic exchange of carbon, and the positive pressure for keeping ADP-glucose specific enzymes is maintained. Still, this very last version of the "ménage à trois" remains highly hypothetical. We should keep in mind that all is based on the assumptions made from the first metabolic symbiosis model, speculating that an ADP-glucose based metabolism was necessarily selected immediately in the common ancestor of extant Archaeplastida. This is compatible with phylogenetic data, but not at all demonstrated. There is no reason to exclude a scenario where the selection for ADP-glucose utilizing enzymes occurred relatively late after the engulfment of the cyanobiont, which is also compatible with phylogenetic data. The "finite" set of EGT that we detect emerged probably from a long and iterative process, with several rounds of transfers in an evolving selective environment. Even if it was repeatedly observed that genes related to carbon metabolism tend to be easily lost from genomes of parasitic species [77], nobody can evaluate the amount of time required to purge these genes definitely, preventing their transfer to the nucleus. With that in mind, why not considering the possibility that a metabolite different from ADP-glucose was first used to export carbon from the cyanobiont? Are we sure that the primitive exchange between cyanobiont and host has been kept intact in extant Archaeplastida? The time before their diversification was a succession of tunings and adjustments, what we infer today is the last state of their common ancestor, which is not necessarily the same as the first step of primary endosymbiosis.

In a recent review, Zimorski et al. [37] stress out, using data from reference [32], that Chlamydiae are not the only and major contributors to essential functions of the plastid by HGT. However, a lesser interest is put into explaining how and when these other genes of various bacterial origins were acquired by ancestors of Archaeplastida. What could have been their influence in helping the success of primary endosymbiosis? Zimorski et al., like Dagan et al. and Thiergart et al. before them [27,32], also point out that it is probably an error to infer the genome of the ancestor of the cyanobiont only from extant cyanobacteria. HGTs between prokaryotes are frequent, and were certainly already frequent in the past. The ancestor of the cyanobiont was a cyanobacterium with a core genome typical of cyanobacteria but also with a singular accessory genome of his own. If some of the genes constituting this accessory genome were subject to EGT toward the host nucleus, they cannot be detected today as cyanobacterial genes, but only as bacterial HGT. A scenario where a Chlamydiae symbiont would have transfered a gene to the cyanobiont while they were sharing the same host (like the one proposed for *uhpC* in the "Ménage à trois" hypothesis) can also be explained with equal likelihood by the presence of this gene in the pan genome of the future cyanobiont before endosymbiosis. What if the cyanobacteria about to become the cyanobiont contained many genes acquired from other bacterias, including Chlamydiae. This was proposed by Martin et al. [29] to explain why *A. thaliana* seems to have acquired EGT from so many bacterial sources without having actually shared an endosymbiotic relationship with all of them. This idea was refused by Brinkman et al. [48] with the argument that extant Cyanobacteria do not possess chlamydial genes in their genomes, at least not in comparable amount with what was transferred to Archaeplastida. As Zimorski et al. explain [37], cyanobacterial genomes have evolved an additional billion years since primary endosymbiosis, this is more than enough to add or eliminate a lot of genes from their accessory genome. Can't we imagine that the peculiarity of primary endosymbiosis is the precise content of the genome of the cyanobiont? A genome that may have contained the key to its special relationship with the host during endosymbiosis. A key that vanished since then.

Acknowledgments

The author would like to thank David Moreira for the numerous discussions on the topic, as well as the six anonymous reviewers of this manuscript for the time they put into writing extensive, precise and very constructive reviews. This work was not supported by any specific funding.

Competing interests

No competing interests have been declared.

Supplementary material

The supplementary material for this article is available online at http://pbsociety.org.pl/journals/index.php/asbp/rt/suppFiles/asbp.2014.048/0:

References

1. Mereschkowsky C. Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. Biol Cent. 1905;25(18):593–604.
2. Margulis L. Symbiosis and evolution. Sci Am. 1971;225(2):48–57.
3. Martin W, Kowallik K. Annotated English translation of Mereschkowsky's 1905 paper "Über Natur und Ursprung der Chromatophoren imPflanzenreiche." Eur J Phycol. 1999;34(3):287–295. http://dx.doi.org/10.1080/09670269910001736342
4. Adl SM, Simpson AGB, Lane CE, Lukeš J, Bass D, Bowser SS, et al. The revised classification of eukaryotes. J Eukaryot Microbiol. 2012;59(5):429–514. http://dx.doi.org/10.1111/j.1550-7408.2012.00644.x
5. Rodríguez-Ezpeleta N, Brinkmann H, Burey SC, Roure B, Burger G, Löffelhardt W, et al. Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. Curr Biol. 2005;15(14):1325–1330. http://dx.doi.org/10.1016/j.cub.2005.06.040
6. Moreira D, Le Guyader H, Philippe H. The origin of red algae and the evolution of chloroplasts. Nature. 2000;405(6782):69–72. http://dx.doi.org/10.1038/35011054
7. Martin W, Brinkmann H, Savonna C, Cerff R. Evidence for a chimeric nature of nuclear genomes: eubacterial origin of eukaryotic glyceraldehyde-3-phosphate dehydrogenase genes. Proc Natl Acad Sci USA. 1993;90(18):8692–8696. http://dx.doi.org/10.1073/pnas.90.18.8692
8. Martin W, Stoebe B, Goremykin V, Hapsmann S, Hasegawa M, Kowallik KV. Gene transfer to the nucleus and the evolution of chloroplasts. Nature. 1998;393(6681):162–165. http://dx.doi.org/10.1038/30234
9. Gutensohn M, Fan E, Frielingsdorf S, Hanner P, Hou B, Hust B, et al. Toc, Tic, Tat et al.: structure and function of protein transport machineries in chloroplasts. J Plant Physiol. 2006;163(3):333–347. http://dx.doi.org/10.1016/j.jplph.2005.11.009
10. Steiner JM, Löffelhardt W. Protein import into cyanelles. Trends Plant Sci. 2002;7(2):72–77. http://dx.doi.org/10.1016/S1360-1385(01)02179-3
11. Koussevitzky S, Nott A, Mockler TC, Hong F, Sachetto-Martins G, Surpin M, et al. Signals from chloroplasts converge to regulate nuclear gene expression. Science. 2007;316(5825):715–719. http://dx.doi.org/10.1126/science.1140516
12. Chan KX, Crisp PA, Estavillo GM, Pogson BJ. Chloroplast-to-nucleus communication: current knowledge, experimental strategies and relationship to drought stress signaling. Plant Signal Behav. 2010;5(12):1575–1582. http://dx.doi.org/10.4161/psb.5.12.13758
13. Keeling PJ. The endosymbiotic origin, diversification and fate of plastids. Phil Trans R Soc Lond B. 2010;365(1541):729–748. http://dx.doi.org/10.1098/rstb.2009.0103
14. Marin B, M. Nowack EC, Melkonian M. A plastid in the making: evidence for a second primary endosymbiosis. Protist. 2005;156(4):425–432. http://dx.doi.org/10.1016/j.protis.2005.09.001
15. Yoon HS, Nakayama T, Reyes-Prieto A, Andersen RA, Boo SM, Ishida K, et al. A single origin of the photosynthetic organelle in different Paulinella lineages. BMC Evol Biol. 2009;9(1):98. http://dx.doi.org/10.1186/1471-2148-9-98
16. Nowack ECM, Vogel H, Groth M, Grossman AR, Melkonian M, Glockner G. Endosymbiotic gene transfer and transcriptional regulation of transferred genes in *Paulinella chromatophora*. Mol Biol Evol. 2011;28(1):407–422. http://dx.doi.org/10.1093/molbev/msq209
17. Mackiewicz P, Bodył A, Gagat P. Possible import routes of proteins into the cyanobacterial endosymbionts/plastids of *Paulinella chromatophora*. Theory Biosci. 2012;131(1):1–18. http://dx.doi.org/10.1007/s12064-011-0147-7
18. Nowack ECM, Grossman AR. Trafficking of protein into the recently established photosynthetic organelles of *Paulinella chromatophora*. Proc Natl Acad Sci USA. 2012;109(14):5340–5345. http://dx.doi.org/10.1073/pnas.1118800109
19. Simonson AB, Servin JA, Skophammer RG, Herbold CW, Rivera MC, Lake JA. Decoding the genomic tree of life. Proc Natl Acad Sci USA. 2005;102(1 suppl):6608–6613. http://dx.doi.org/10.1073/pnas.0501996102

20. Ribeiro S, Golding GB. The mosaic nature of the eukaryotic nucleus. Mol Biol Evol. 1998;15(7):779–788.
21. Anantharaman V, Koonin EV, Aravind L. Comparative genomics and evolution of proteins involved in RNA metabolism. Nucl Acids Res. 2002;30(7):1427–1464.
22. Langer D, Hain J, Thuriaux P, Zillig W. Transcription in archaea: similarity to that in eucarya. Proc Natl Acad Sci USA. 1995;92(13):5768–5772.
23. Esser C. A genome phylogeny for mitochondria among alpha-proteobacteria and a predominantly eubacterial ancestry of yeast nuclear genes. Mol Biol Evol. 2004;21(9):1643–1660. http://dx.doi.org/10.1093/molbev/msh160
24. López-Garćia P, Moreira D. Metabolic symbiosis at the origin of eukaryotes. Trends Biochem Sci. 1999;24(3):88–93. http://dx.doi.org/10.1016/S0968-0004(98)01342-5
25. Martin W, Müller M. The hydrogen hypothesis for the first eukaryote. Nature. 1998;392(6671):37–41. http://dx.doi.org/10.1038/32096
26. Martin W. Archaebacteria (Archaea) and the origin of the eukaryotic nucleus. Curr Opin Microbiol. 2005;8(6):630–637. http://dx.doi.org/10.1016/j.mib.2005.10.004
27. Thiergart T, Landan G, Schenk M, Dagan T, Martin WF. An evolutionary network of genes present in the eukaryote common ancestor polls genomes on eukaryotic and mitochondrial origin. Genome Biol Evol. 2012;4(4):466–485. http://dx.doi.org/10.1093/gbe/evs018
28. Rochette NC, Brochier-Armanet C, Gouy M. Phylogenomic test of the hypotheses for the evolutionary origin of eukaryotes. Mol Biol Evol. 2014;31(4):832–845. http://dx.doi.org/10.1093/molbev/mst272
29. Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, et al. Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc Natl Acad Sci USA. 2002;99(19):12246–12251. http://dx.doi.org/10.1073/pnas.182432999
30. Reyes-Prieto A, Hackett JD, Soares MB, Bonaldo MF, Bhattacharya D. Cyanobacterial contribution to algal nuclear genomes is primarily limited to plastid functions. Curr Biol. 2006;16(23):2320–2325. http://dx.doi.org/10.1016/j.cub.2006.09.063
31. Deusch O, Landan G, Roettger M, Gruenheit N, Kowallik KV, Allen JF, et al. Genes of cyanobacterial origin in plant nuclear genomes point to a heterocyst-forming plastid ancestor. Mol Biol Evol. 2008;25(4):748–761. http://dx.doi.org/10.1093/molbev/msn022
32. Dagan T, Roettger M, Stucken K, Landan G, Koch R, Major P, et al. Genomes of Stigonematalean cyanobacteria (subsection V) and the evolution of oxygenic photosynthesis from prokaryotes to plastids. Genome Biol Evol. 2013;5(1):31–44. http://dx.doi.org/10.1093/gbe/evs117
33. Reyes-Prieto A, Moustafa A. Plastid-localized amino acid biosynthetic pathways of Plantae are predominantly composed of non-cyanobacterial enzymes. Sci Rep. 2012;2:955. http://dx.doi.org/10.1038/srep00955
34. Qiu H, Price DC, Weber APM, Facchinelli F, Yoon HS, Bhattacharya D. Assessing the bacterial contribution to the plastid proteome. Trends Plant Sci. 2013;18(12):680–687. http://dx.doi.org/10.1016/j.tplants.2013.09.007
35. Suzuki K, Miyagishima SY. Eukaryotic and eubacterial contributions to the establishment of plastid proteome estimated by large-scale phylogenetic analyses. Mol Biol Evol. 2010;27(3):581–590. http://dx.doi.org/10.1093/molbev/msp273
36. Stegemann S, Bock R. Experimental reconstruction of functional gene transfer from the tobacco plastid genome to the nucleus. Plant Cell. 2006;18(11):2869–2878. http://dx.doi.org/10.1105/tpc.106.046466
37. Zimorski V, Ku C, Martin WF, Gould SB. Endosymbiotic theory for organelle origins. Curr Opin Microbiol. 2014;22:38–48. http://dx.doi.org/10.1016/j.mib.2014.09.008
38. Kamneva OK, Knight SJ, Liberles DA, Ward NL. Analysis of genome content evolution in pvc bacterial super-phylum: assessment of candidate genes associated with cellular organization and lifestyle. Genome Biol Evol. 2012;4(12):1375–1390. http://dx.doi.org/10.1093/gbe/evs113
39. Horn M, Collingro A, Schmitz-Esser S, Beier CL, Purkhold U, Fartmann B, et al. Illuminating the evolutionary history of chlamydiae. Science. 2004;304(5671):728–730. http://dx.doi.org/10.1126/science.1096330
40. Horn M. Chlamydiae as symbionts in eukaryotes. Annu Rev Microbiol. 2008;62(1):113–131. http://dx.doi.org/10.1146/annurev.micro.62.081307.162818
41. Wolf YI, Aravind L, Koonin EV. Rickettsiae and Chlamydiae: evidence of horizontal gene transfer and gene exchange. Trends Genet. 1999;15(5):173–175. http://dx.doi.org/10.1016/S0168-9525(99)01704-7
42. Haferkamp I, Schmitz-Esser S, Wagner M, Neigel N, Horn M, Neuhaus HE. Tapping the nucleotide pool of the host: novel nucleotide carrier proteins of *Protochlamydia amoebophila*. Mol Microbiol. 2006;60(6):1534–1545. http://dx.doi.org/10.1111/j.1365-2958.2006.05193.x
43. Schwöppe C, Winkler HH, Neuhaus HE. Properties of the glucose-6-phosphate transporter from *Chlamydia pneumoniae* (HPTcp) and the glucose-6-phosphate sensor from *Escherichia coli* (UhpC). J Bacteriol. 2002;184(8):2108–2115. http://dx.doi.org/10.1128/JB.184.8.2108-2115.2002
44. Subtil A, Collingro A, Horn M. Tracing the primordial Chlamydiae: extinct parasites of plants? Trends Plant Sci. 2014;19(1):36–43. http://dx.doi.org/10.1016/j.tplants.2013.10.005
45. Becker B, Hoef-Emden K, Melkonian M. Chlamydial genes shed light on the evolution of photoautotrophic eukaryotes. BMC Evol Biol. 2008;8(1):203. http://dx.doi.org/10.1186/1471-2148-8-203
46. Huang J, Gogarten J. Did an ancient chlamydial endosymbiosis facilitate the establishment of primary plastids? Genome Biol. 2007;8(6):R99. http://dx.doi.org/10.1186/gb-2007-8-6-r99
47. Moustafa A, Reyes-Prieto A, Bhattacharya D. Chlamydiae has contributed at least 55 genes to Plantae with predominantly plastid functions. PLoS ONE. 2008;3(5):e2205. http://dx.doi.org/10.1371/journal.pone.0002205
48. Brinkman FSL, Blanchard JL, Cherkasov A, Av-Gay Y, Brunham RC, Fernandez RC, et al. Evidence that plant-like genes in *Chlamydia* species reflect an ancestral relationship between Chlamydiaceae, cyanobacteria, and the chloroplast. Genome Res. 2002;12(8):1159–1167. http://dx.doi.org/10.1101/gr.341802
49. Stephens RS. Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. Science. 1998;282(5389):754–759. http://dx.doi.org/10.1126/science.282.5389.754
50. Collingro A, Tischler P, Weinmaier T, Penz T, Heinz E, Brunham RC, et al. Unity in variety – the pan-genome of the Chlamydiae. Mol Biol Evol. 2011;28(12):3253–3270. http://dx.doi.org/10.1093/molbev/msr161
51. Ball SG, Subtil A, Bhattacharya D, Moustafa A, Weber APM, Gehre L, et al. Metabolic effectors secreted by bacterial pathogens: essential facilitators of plastid endosymbiosis? Plant Cell. 2013;25(1):7–21. http://dx.doi.org/10.1105/tpc.112.101329
52. Reinhold T, Alawady A, Grimm B, Beran KC, Jahns P, Conrath U, et al. Limitation of nocturnal import of ATP into *Arabidopsis* chloroplasts leads to photooxidative damage. Plant J. 2007;50(2):293–304. http://dx.doi.org/10.1111/j.1365-313X.2007.03049.x
53. Watkins RF, Gray MW. The frequency of eubacterium-to-eukaryote lateral gene transfers shows significant cross-taxa variation within amoebozoa. J Mol Evol. 2006;63(6):801–814. http://dx.doi.org/10.1007/s00239-006-0031-0
54. Deschamps P, Colleoni C, Nakamura Y, Suzuki E, Putaux JL, Buléon A, et al. Metabolic symbiosis and the birth of the plant kingdom. Mol Biol Evol. 2008;25(3):536–548. http://dx.doi.org/10.1093/molbev/msm280
55. Nakamura Y, Takahashi J, Sakurai A, Inaba Y, Suzuki E, Nihei S, et al. Some cyanobacteria synthesize semi-amylopectin type alpha-polyglucans instead of glycogen. Plant Cell Physiol. 2005;46(3):539–545. http://dx.doi.org/10.1093/pcp/pci045
56. Gallon JR. The oxygen sensitivity of nitrogenase: a problem for biochemists and micro-organisms. Trends Biochem Sci. 1981;6:19–23. http://dx.doi.org/10.1016/0968-0004(81)90008-6
57. Wolk CP, Ernst A, Elhai J. Heterocyst metabolism and development. In: Bryant DA, editor. The molecular biology of cyanobacteria. Dordrecht: Springer; 2004. p. 769–823. (Advances in photosynthesis and respiration; vol 1).
58. Deschamps P, Moreau H, Worden AZ, Dauvillée D, Ball SG. Early

gene duplication within chloroplastida and its correspondence with relocation of starch metabolism to chloroplasts. Genetics. 2008;178(4):2373–2387. http://dx.doi.org/10.1534/genetics.108.087205

59. Deschamps P, Haferkamp I, d' Hulst C, Neuhaus HE, Ball SG. The relocation of starch metabolism to chloroplasts: when, why and how. Trends Plant Sci. 2008;13(11):574–582. http://dx.doi.org/10.1016/j.tplants.2008.08.009
60. Colleoni C, Linka M, Deschamps P, Handford MG, Dupree P, Weber APM, et al. Phylogenetic and biochemical evidence supports the recruitment of an ADP-glucose translocator for the export of photosynthate during plastid endosymbiosis. Mol Biol Evol. 2010;27(12):2691–2701. http://dx.doi.org/10.1093/molbev/msq158
61. Lu C, Lei L, Peng B, Tang L, Ding H, Gong S, et al. *Chlamydia trachomatis* GlgA is secreted into host cell cytoplasm. PLoS ONE. 2013;8(7):e68764. http://dx.doi.org/10.1371/journal.pone.0068764
62. Facchinelli F, Pribil M, Oster U, Ebert NJ, Bhattacharya D, Leister D, et al. Proteomic analysis of the *Cyanophora paradoxa* muroplast provides clues on early events in plastid endosymbiosis. Planta. 2013;237(2):637–651. http://dx.doi.org/10.1007/s00425-012-1819-3
63. Shattuck-Eidens DM, Kadner RJ. Molecular cloning of the uhp region and evidence for a positive activator for expression of the hexose phosphate transport system of *Escherichia coli*. J Bacteriol. 1983;155(3):1062–1070.
64. Price DC, Chan CX, Yoon HS, Yang EC, Qiu H, Weber APM, et al. *Cyanophora paradoxa* genome elucidates origin of photosynthesis in algae and plants. Science. 2012;335(6070):843–847. http://dx.doi.org/10.1126/science.1213561
65. Rockwell NC, Lagarias JC, Bhattacharya D. Primary endosymbiosis and the evolution of light and oxygen sensing in photosynthetic eukaryotes. Front Ecol Evol. 2014;2:66. http://dx.doi.org/10.3389/fevo.2014.00066
66. Facchinelli F, Colleoni C, Ball SG, Weber APM. *Chlamydia*, cyanobiont, or host: who was on top in the ménage à trois? Trends Plant Sci. 2013;18(12):673–679. http://dx.doi.org/10.1016/j.tplants.2013.09.006
67. Keeling PJ, Burki F, Wilcox HM, Allam B, Allen EE, Amaral-Zettler LA, et al. The marine microbial eukaryote transcriptome sequencing project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. PLoS Biol. 2014;12(6):e1001889. http://dx.doi.org/10.1371/journal.pbio.1001889
68. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl Acids Res. 1997;25(17):3389–3402. http://dx.doi.org/10.1093/nar/25.17.3389
69. Pruitt KD, Tatusova T, Maglott DR. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucl Acids Res. 2007;35(database):D61–D65. http://dx.doi.org/10.1093/nar/gkl842
70. Katoh K, Toh H. Recent developments in the MAFFT multiple sequence alignment program. Brief Bioinform. 2008;9(4):286–298. http://dx.doi.org/10.1093/bib/bbn013
71. Criscuolo A, Gribaldo S. BMGE (block mapping and gathering with entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol Biol. 2010;10(1):210. http://dx.doi.org/10.1186/1471-2148-10-210
72. Price MN, Dehal PS, Arkin AP. FastTree 2 – approximately maximum-likelihood trees for large alignments. PLoS ONE. 2010;5(3):e9490. http://dx.doi.org/10.1371/journal.pone.0009490
73. Jobb G, von Haeseler A, Strimmer K. TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. BMC Evol Biol. 2004;4(1):18. http://dx.doi.org/10.1186/1471-2148-4-18
74. Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, et al. The *Arabidopsis* information resource (TAIR): improved gene annotation and new tools. Nucl Acids Res. 2012;40(D1):D1202–D1210. http://dx.doi.org/10.1093/nar/gkr1090
75. Keeling PJ, Palmer JD. Horizontal gene transfer in eukaryotic evolution. Nat Rev Genet. 2008;9(8):605–618. http://dx.doi.org/10.1038/nrg2386
76. Marcet-Houben M, Gabaldón T. Acquisition of prokaryotic genes by fungal genomes. Trends Genet. 2010;26(1):5–8. http://dx.doi.org/10.1016/j.tig.2009.11.007
77. Henrissat B, Deleury E, Coutinho PM. Glycogen metabolism loss: a common marker of parasitic behaviour in bacteria? Trends Genet. 2002;18(9):437–440. http://dx.doi.org/10.1016/S0168-9525(02)02734-8