Review

New Phytologist

Tansley insight

Horizontal and endosymbiotic gene transfer in early plastid evolution

Author for correspondence:
David Moreira
Tel: +33 69157608
Email: david.moreira@u-psud.fr

Received: *30 December 2018*
Accepted: *17 May 2019*

Rafael I. Ponce-Toledo, Purificación López-García and David Moreira

Unité d'Ecologie Systématique et Evolution, CNRS, Université Paris-Sud, AgroParisTech, Université Paris-Saclay, 91400, Orsay, France

Contents

New Phytologist (2019) **224:** 618–624
doi: 10.1111/nph.15965

Key words: Archaeplastida, Cyanobacteria, endosymbiosis, evolution, gene transfer, *Paulinella*, plastids.

Summary

Plastids evolved from a cyanobacterium that was engulfed by a heterotrophic eukaryotic host and became a stable organelle. Some of the resulting eukaryotic algae entered into a number of secondary endosymbioses with diverse eukaryotic hosts. These events had major consequences on the evolution and diversification of life on Earth. Although almost all plastid diversity derives from a single endosymbiotic event, the analysis of nuclear genomes of plastid-bearing lineages has revealed a mosaic origin of plastid-related genes. In addition to cyanobacterial genes, plastids recruited for their functioning eukaryotic proteins encoded by the host nucleus and also bacterial proteins of noncyanobacterial origin. Therefore, plastid proteins and plastid-localised metabolic pathways evolved by tinkering and using gene toolkits from different sources. This mixed heritage seems especially complex in secondary algae containing green plastids, the acquisition of which appears to have been facilitated by many previous acquisitions of red algal genes (the 'red carpet hypothesis').

I. Introduction

Oxygenic photosynthesis in eukaryotes appeared > 1 billion years ago (Eme *et al.*, 2014) via the endosymbiosis of a close relative of the deep-branching cyanobacterium *Gloeomargarita lithophora* within a phagotrophic eukaryotic host (Ponce-Toledo *et al.*, 2017). Subsequently, the cyanobacterium evolved into a permanent photosynthetic organelle called primary plastid. This endosymbiosis is at the origin of the Archaeplastida, a monophyletic supergroup composed of three primary plastid-bearing lineages: the green algae plus land plants, the red algae, and the glaucophytes (Fig. 1; Adl *et al.*, 2012). Several cercozoan amoebae of the genus *Paulinella* (*P. chromatophora*, *P. longichromatophora* and *P. micropora*) (Fig. 2) constitute the only known lineage apart from Archaeplastida in which an independent type of primary photosynthetic organelles evolved (Marin *et al.*, 2005). The *Paulinella* organelles, called 'chromatophores', derived from a cyanobacterium of the *Synechococcus/Prochlorococcus* group (α-cyanobacteria), in contrast with the *Gloeomargarita*-like plastids of the Archaeplastida. Whereas primary endosymbioses are extremely rare in biological history, with only the Archaeplastida and *Paulinella* cases known, green and red algae have participated as endosymbionts in numerous secondary and tertiary endosymbioses, resulting in a broad diversity of eukaryotic lineages with

complex plastids (Fig. 1; Moreira & Philippe, 2001; Keeling, 2010).

The reduced number of primary plastid acquisitions compared with that of eukaryotic lineages with complex plastids suggests that the enslavement of a photosynthetic cyanobacterial endosymbiont (or cyanobiont) is, for unknown reasons, more challenging from an evolutionary point of view and requires specific adaptations to stabilise the cyanobiont as a permanent organelle. Once the primary plastid was fully integrated within the host (through the evolution of metabolite export and protein import systems, transfer of plastid genes to the host nucleus and evolution of proteins involved in redox regulation) it was easier for algae to become endosymbionts of other eukaryotes. Here, we review the genetic and genomic changes that accompanied the evolution of primary plastids and explore the plastid proteome composition to propose a possible role for host-derived and exogenous genes in the establishment of these plastids. Likewise, we discuss the genetic mosaicism of nuclear genomes in complex plastid-harbouring lineages in the light of horizontal gene transfer (HGT) and putative cryptic endosymbioses.

II. Evolution of primary plastids in Archaeplastida

Plastids in Archaeplastida are derived from an endosymbiotic cyanobacterium that was fully integrated in a heterotrophic eukaryotic host and became an organelle. Nonetheless, plastids differ considerably from free-living cyanobacteria. One of the most remarkable differences is the drastic reduction of plastid genomes,

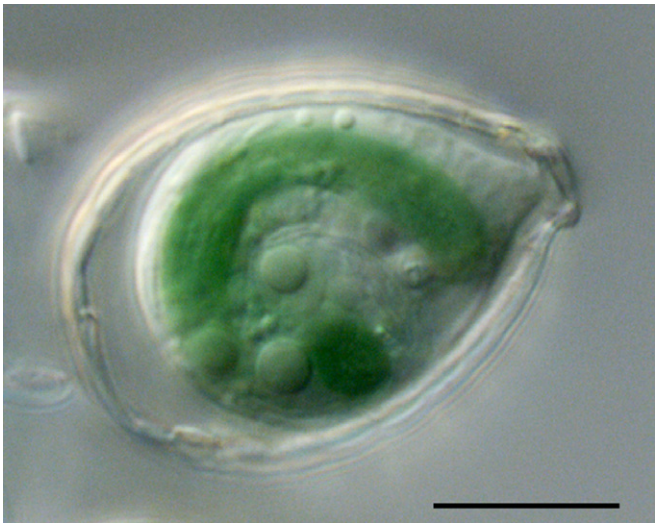

Fig. 2 Light microscopy image of *Paulinella chromatophora*. The pigmented *Synechococcus*-like primary plastids are easily visible within the cytoplasm. Bar, 10 μm. Image courtesy of Eva Nowack (Heinrich-Heine-Universität Düsseldorf, Germany).

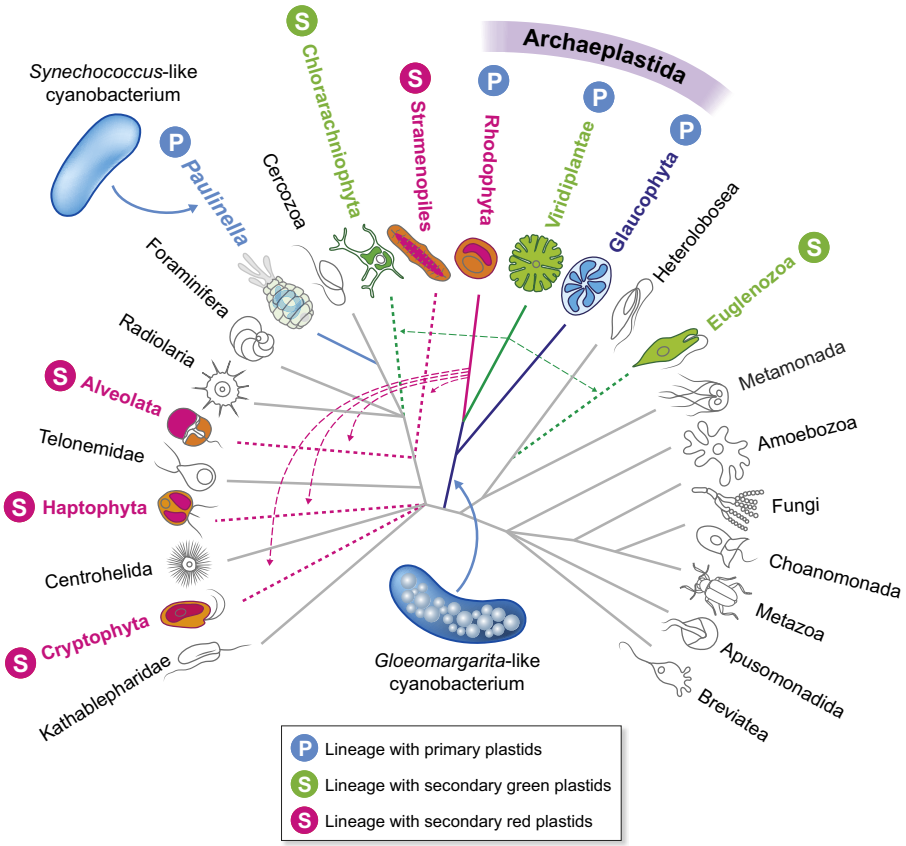

Fig. 1 The distribution of photosynthesis in global eukaryotic phylogeny. Coloured solid branches correspond to photosynthetic lineages endowed with primary plastids and coloured dashed branches to lineages with secondary plastids (green and red colours indicate the type of secondary endosymbiont, green or red algae, respectively). Blue arrows show the two known primary endosymbioses (in Archaeplastida and *Paulinella*) and green and red arrows indicate secondary endosymbioses involving green and red algal endosymbionts. Grey branches correspond to nonphotosynthetic eukaryotic phyla. The tree has been largely modified from Adl *et al*. (2012).

which encode < 5% of the genes found in typical free-living cyanobacteria (Green, 2011). The number of protein-coding genes in plastid genomes varies across Archaeplastida, with red algae harbouring the largest sets (160–235 proteins; Lee *et al.*, 2016), perhaps as a result of the massive gene loss in the nuclear genome likely to have been experienced by the rhodophyte ancestor, which exerted a selective pressure to keep plastid-encoded genes (Bhattacharya *et al.*, 2018).

Gene loss in the cyanobiont was a crucial event during early endosymbiosis stages and marked the transformation of the cyanobacterium into an obligatory endosymbiont. Glycogen synthase genes, absent in plastids, were possibly among the first genes to be lost and this event may have helped to achieve the early host−endosymbiont metabolic integration (Gavelis & Gile, 2018). Glycogen synthase-defective mutant cyanobacteria release organic carbon outside the cell as a homeostatic response to cope with excessive photosynthate production and the inability to store it as glycogen (Cano *et al.*, 2018). Therefore, it is possible that inefficient glycogen storage allowed the leakage of photosynthates that could be used by the host before the evolution of a specific and more efficient photosynthate export system.

It is commonly assumed that endosymbiotic gene transfer (EGT) was critical for plastid endosymbiosis because it allowed the host to gain control over the expression of plastid genes, therefore increasing the dependence of the cyanobiont upon the host gene expression machinery. Most plastid genes transferred to the host nucleus participate in photosynthesis-related functions (for example photosystem subunits, chlorophyll synthesis) and plastid metabolism and maintenance (for example amino acid synthesis and plastid division) (Reyes-Prieto *et al.*, 2006). Recent studies have suggested that EGTs also contributed to the expansion of the redox sensing capabilities of the host, sometimes through genetic tinkering that produced new chimeric proteins (Méheust *et al.*, 2016). This process was accompanied by the evolution of new redox sensing pathways that helped the host to cope with the increased reactive oxygen species concentration produced by the new organelle (Woehle *et al.*, 2017).

Even though plastid-encoded proteins plus plastid-targeted EGT products are essential for the correct functioning of these organelles, they represent less than half of the total plastid proteome in Archaeplastida (Qiu *et al.*, 2013). The fact that the plastid proteome contains a vast proportion of apparently noncyanobacterial proteins opens interesting questions about the mechanisms underlying plastid acquisition and suggests that plastid symbiogenesis was not as straightforward as it is commonly assumed (Larkum *et al.*, 2007). For instance, recent phylogenetic analyses suggest that the genetic machinery necessary to synthesize the galactolipids of plastid membranes − once thought to derive from the cyanobiont − do not have cyanobacterial origin, opening many questions about the nature of plastid membranes (Sato & Awai, 2017).

Host-derived proteins, evolved either from the retargeting of pre-existing proteins or from new gene innovations, participate in a wide range of plastid functions and represent the largest fraction of plastid-targeted proteins (Qiu *et al.*, 2013). For instance, metabolite transporters that originated from the retargeting of host membrane proteins are particularly overrepresented in the plastid envelope (58% of plastid transporters appear to derive from host membrane proteins in *Arabidopsis thaliana*; Tyra *et al.*, 2007). These findings suggest that the host drove the integration of the cyanobiont by providing the proteins necessary to connect the cyanobiont metabolism with the energy demands of the host ('host-centric' endosymbiosis model; Karkar *et al.*, 2015).

In addition to host-derived genes, noncyanobacterial bacterial genes were also crucial for plastid evolution and represent 7–15% of the plastid proteome (Qiu *et al.*, 2013). The phylogenetic origin of these noncyanobacterial prokaryotic genes in Archaeplastida seems to cover a wide range of bacterial phyla (Dagan *et al.*, 2013). After Cyanobacteria, Proteobacteria appear to be the most common contributors to the plastid proteome, particularly alphaproteobacterial proteins that may derive from the mitochondrial ancestor (Dagan *et al.*, 2013; Qiu *et al.*, 2013). Some of these noncyanobacterial bacterial genes were likely to have been present in the genome of the cyanobacterial plastid ancestor as a consequence of HGT events from various donors (Dagan *et al.*, 2013). Others may have been subsequently acquired by the Archaeplastida ancestor due to its presumed phagotrophic lifestyle and may have helped to compensate for the massive gene loss undergone by the cyanobiont genome.

An issue that has attracted much attention is the presence of a number of genes apparently transferred from Chlamydiales bacteria to the Archaeplastida (Huang & Gogarten, 2007; Becker *et al.*, 2008), and has even led to the proposal of a tripartite model of plastid origin (Facchinelli *et al.*, 2013). This model suggests that chlamydial cells infected the Archaeplastida ancestor. This infection would have helped in the early steps of plastid acquisition by protecting the cyanobiont from host defenses, supplying multiple enzymes to integrate photosynthates produced by the cyanobiont into the host carbohydrate metabolism (Facchinelli *et al.*, 2013) and/or allowing the cyanobiont to cope with ATP starvation resulting from the hypoxic environment of the host cytosol (Cenci *et al.*, 2018). However, this model encounters several problems. First, there is no report of any *Chlamydia* bacteria able to infect Archaeplastida, this situation suggests that the Archaeplastida ancestor was not a probable host for this genus of pathogenic bacteria. Second, and more importantly, phylogenetic re-analyses of putative *Chlamydia*-derived genes have detected various phylogenetic artefacts and reduced considerably the number of genes compatible with a putative chlamydial ancestry (Moreira & Deschamps, 2014; Domman *et al.*, 2015). Therefore, the tripartite model continues to be hotly debated.

III. Chromatophore evolution in *Paulinella*

Several cercozoan amoebae of the genus *Paulinella* (for example *P. chromatophora*; Fig. 2) adopted a primary photosynthetic organelle (called a chromatophore) by the endosymbiosis of a cyanobacterium from the *Synechococcus/Prochlorococcus* (Syn/Pro) clade (Marin *et al.*, 2005). *Paulinella* is a good model in which to study the early evolution of primary plastids as the divergence of the chromatophore from its Syn/Pro ancestor is relatively recent, only 90–140 Ma (Delaye *et al.*, 2016). Remarkably, there are important

Table 1. Comparison of plastid characteristics between Archaeplastida and *Paulinella chromatophora*.

Characteristic	Archaeplastida	*Paulinella*	References
Plastid genome size	100–200 kbp	1021 kbp	Nowack *et al*. (2008)
Number of plastid genes	80–250	911	Nowack *et al*. (2008)
EGTs in plastid proteome	70–390	> 70	Qiu *et al*. (2013); Singer *et al*. (2017); Nowack *et al*. (2016); Zhang *et al*. (2017)
Noncyanobacterial prokaryotic proteins in plastid proteome	40–240	> 170	Qiu *et al*. (2013); Singer *et al*. (2017); Nowack *et al*. (2016); Zhang *et al*. (2017)
Host-derived and proteins of uncertain origin in plastid proteome	320–900	> 390	Qiu *et al*. (2013); Singer *et al*. (2017); Nowack *et al*. (2016); Zhang *et al*. (2017)
Lineage age estimation	> 1000 Ma	90–140 Ma	Eme *et al*. (2014); Delaye *et al*. (2016)
Import system of nucleus-encoded proteins into the plastid	Translocons at the outer and inner plastid membranes (TOC/TIC complex)	Vesicles of the host endomembrane system fuse with the outer plastid membrane and proteins cross the inner membrane through a simplified TIC translocon	Mackiewicz *et al*. (2012)
Phagotrophic capacity	Lost in Rhodophyta, Glaucophyta and most Viridiplantae but preserved in some prasinophytes	Lost	Gagat & Mackiewicz (2017)
Peptidoglycan wall	Present in Glaucophyta and some Viridiplantae species	Present	Gagat & Mackiewicz (2017)

similarities between primary endosymbioses in Archaeplastida and *Paulinella*, likely to have been due to convergent evolution in the process of plastid acquisition (Table 1).

The chromatophore genome is highly reduced, encoding 867 proteins that represent about one-third of proteins of its free-living counterparts (Nowack *et al*., 2008). Similar to the EGTs found in Archaeplastida, *P. chromatophora* has relocated > 70 chromatophore genes into the nuclear genome (mostly involved in photosynthesis-related functions) (Nowack *et al*., 2016; Zhang *et al*., 2017). By contrast, these genes represent < 1% of the *Paulinella* nuclear genome, while in *A. thaliana* some reports have suggested that the genes of cyanobacterial origin can account for up to 18% of the nuclear genes (Martin *et al*., 2002). Nonetheless, chromatophore genome reduction is most likely to be ongoing and more genes will still be possibly transferred to the host nucleus. However, based on the reduced number of chromatophore-derived genes in the nuclear genome, it seems that EGT may have been less important in the establishment of the chromatophore in *P. chromatophora* than HGT from other bacteria, as *c*.170 genes of bacterial origin encode proteins likely to be targeted to the chromatophore (Nowack *et al*., 2016). Interestingly, the largest contribution of identified chromatophore-targeted proteins derive from the ancestral host genetic repertoire (Singer *et al*., 2017), a similar pattern to the one observed in Archaeplastida proteomes (Qiu *et al*., 2013), suggesting that in both primary endosymbioses the host played a crucial role in mediating cyanobiont integration.

IV. Complex plastids

Eukaryotic lineages with complex plastids have evolved by the engulfment of red or green algae by different hosts. This type of event, called secondary endosymbiosis, produced plastids with three or four membranes (in contrast with the two membranes of primary plastids) and introduced a high degree of reticulation within the eukaryotic global phylogeny that is not yet fully understood (Archibald, 2015). All, or a considerable fraction of, members from four ecologically diverse eukaryotic lineages have red alga-derived plastids: cryptophytes, alveolates, stramenopiles, and haptophytes (the 'CASH' assemblage). The makeup of plastids is particularly diverse in alveolates, including the three-membrane peridinin-containing dinoflagellate plastids and the four-membrane nonphotosynthetic ('apicoplasts') and photosynthetic plastids found in parasitic Apicomplexa and their relatives. While the phylogeny of plastid genes supports that a single red alga was at the origin of all complex red plastids (Yoon *et al*., 2002; Muñoz-Gómez *et al*., 2017), most phylogenetic analyses of nuclear genes have suggested that CASH lineages are not monophyletic. This has been interpreted as an indication that complex red plastids were acquired through an undetermined number of events that may have included serial tertiary endosymbioses and kleptoplastidy in different hosts (Bodył *et al*., 2009; Baurain *et al*., 2010; Petersen *et al*., 2014; Bodył, 2017). Conversely, the origin of complex green plastids is much clearer: euglenids and chlorarachniophytes acquired their plastids from two independent secondary endosymbioses involving two distantly related green algal endosymbionts (Jackson *et al*., 2018).

Similar to primary endosymbiosis, during secondary endosymbiotic events genes were transferred from the algal endosymbiont to the host nucleus (secondary EGTs). This transfer was probably accompanied by transfers from bacterial sources (for example Záhonová *et al*., 2018). Unexpectedly, the analysis of nuclear genes of algal origin in eukaryotic lineages

with complex plastids has revealed a wide range of putative algal donors that are different from the algal endosymbionts that exist today as secondary plastids (Curtis *et al.*, 2012). For instance, despite diatoms containing plastids clearly derived from red algae, phylogenetic analysis of diatom nuclear genomes has suggested that >1700 nucleus-encoded genes were apparently transferred from green algae (Moustafa *et al.*, 2009). Similarly, recent estimates have suggested that *c.* 25% of nucleus-encoded plastid-targeted proteins in the ancestor of ochrophytes (photosynthetic stramenopiles) have derived from green algae (Dorrell *et al.*, 2017). This apparent massive genetic mosaicism may be explained either as the result of a high frequency of eukaryote-to-eukaryote HGT or as the consequence of putative cryptic endosymbioses. Nevertheless, although it is possible that the 'green' signal observed in ochrophyte genomes, particularly in diatoms, might attest for a former endosymbiosis with a green alga (see Moustafa *et al.*, 2009), it seems more likely that the green contribution to diatom genomes has been largely overestimated because of several undetected tree reconstruction artefacts (Deschamps & Moreira, 2012).

V. The 'red carpet' hypothesis

For some secondary EGTs, phylogenetic analysis allows researchers to trace back the full sequence of endosymbiotic gene transfer from Cyanobacteria to their Archaeplastida ancestor and then from red or green algal endosymbionts to the secondary photosynthetic lineages. Phylogenetic inspection of these EGTs in the CASH lineages has shown that – as expected – the majority of these genes was transferred from the red algal secondary endosymbiont (Deschamps & Moreira, 2012; Ponce-Toledo *et al.*, 2018). By contrast, *c.* 30% and 50% of this type of cyanobacteria-derived secondary EGTs appear also to have red algal ancestry in euglenids and chlorarachniophytes, respectively (Ponce-Toledo *et al.*, 2018). Therefore, these two green plastid-harbouring lineages have an unexpected mix of red and green plastid-targeted proteins that creates a high mosaicism in their plastid metabolic pathways, including biosynthetic pathways, photosynthesis-related functions, and plastid biogenesis (Yang *et al.*, 2011). Interestingly, the dinoflagellate *Lepidodinium chlorophorum*, which is a clear example of recent replacement of an original red algal endosymbiont by a pedinophyte-like green algal one (Kamikawa *et al.*, 2015), harbours several nucleus-encoded genes transferred from the former red plastid that have been retargeted to the new green plastid. Therefore, the current green algal plastid functions with a mixture of red and green genes (Minge *et al.*, 2010; Matsuo & Inagaki, 2018).

However, for euglenids and chlorarachniophytes, it is difficult to know if the presence of this large number of genes of red algal origin is because these lineages also have experienced former endosymbioses with red algae. Cryptic endosymbiosis scenarios have to be considered cautiously when interpreting the chimerism observed in the nuclear genomes of photosynthetic eukaryotes (Deschamps & Moreira, 2012). Molecular clock analyses have suggested that chlorarachniophytes acquired their green plastid 578–318 Ma (Jackson *et al.*, 2018) while the chlorarachniophyte host lineage appears to have diverged from heterotrophic cercozoa *c.* 1000 Ma (Parfrey *et al.*, 2011). Therefore, there was a long period (>400 Myr) during which a putative secondary endosymbiosis with a red alga might have taken place before the acquisition of the present green algal plastid. What is certain is that red algae (or lineages containing red algal secondary plastids) have provided many plastid-related genes to both euglenids and chlorarachniophytes. These genes appear to be present in all known species of each of these two phyla, indicating that they correspond to ancient gene acquisitions before the diversification of these two phyla (Ponce-Toledo *et al.*, 2018). We introduce here the 'red carpet hypothesis' to propose that these red algal genes, transferred to the host nucleus before and/or during the early steps of endosymbiosis with green algae, provided important plastid-related functions and acted as a sort of 'red carpet' to facilitate the subsequent adoption of the new green algal endosymbionts.

VI. Conclusions

Phylogenetic analyses of nucleus-encoded plastid-targeted proteins have revealed the massive contribution of noncyanobacterial proteins to the plastid proteomes of Archaeplastida and *Paulinella*, the only two lineages known to harbour primary plastids (Qiu *et al.*, 2013; Nowack *et al.*, 2016). Gene transfers from bacteria other than cyanobionts and their retargeting to the early plastid seem to have been very frequent in both primary photosynthetic lineages, many possibly replacing genes that have been lost from the cyanobiont genome. Nonetheless, the eukaryotic host seems to have been the largest contributor of plastid-targeted proteins, with most of these involved in plastid maintenance and transport of metabolites, this situation supports the idea that the host drove the early steps of plastid endosymbiosis (Karkar *et al.*, 2015).

It is still not completely clear why secondary endosymbioses have occurred much more often than primary endosymbioses in the evolution of contemporary eukaryotic photosynthetic lineages. However, it seems that nucleus-encoded genes acquired from previous endosymbioses can be helpful in regaining a plastid, likely to be because the expression of these genes is already under the control of the host and because they carry targeting signals that can be more easily reused to target proteins towards a new endosymbiont (Matsuo & Inagaki, 2018), therefore speeding up host–endosymbiont integration. Although cryptic endosymbioses are difficult to prove, we postulate that plastid-targeted secondary EGTs are critical markers from which to test these evolutionary scenarios.

Acknowledgements

We thank Marc-André Selosse for the invitation to contribute this article, Eva Nowack for the *Paulinella chromatophora* picture, and Philippe Deschamps for discussion on plastid evolution topics. Our work was supported by ERC Advanced Grants 'Protistworld' and 'Plast-Evol' (322669 and 787904, respectively) and ANR grant ANCESSTRAM (ANR-15-CE32-0003).

ORCID

Purificación López-García https://orcid.org/0000-0002-0927-0651
David Moreira https://orcid.org/0000-0002-2064-5354
Rafael I. Ponce-Toledo https://orcid.org/0000-0002-6194-8845

References

Adl SM, Simpson AGB, Lane CE, Lukeš J, Bass D, Bowser SS, Brown MW, Burki F, Dunthorn M, Hampl V *et al.* 2012. The revised classification of eukaryotes. *Journal of Eukaryotic Microbiology* **59**: 429–493.

Archibald JM. 2015. Genomic perspectives on the birth and spread of plastids. *Proceedings of the National Academy of Sciences, USA* **112**: 10147–10153.

Baurain D, Brinkmann H, Petersen JJ, Rodríguez-Ezpeleta N, Stechmann A, Demoulin V, Roger AJ, Burger G, Lang BF, Philippe H. 2010. Phylogenomic evidence for separate acquisition of plastids in cryptophytes, haptophytes, and stramenopiles. *Molecular Biology and Evolution* **27**: 1698–1709.

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

Bhattacharya D, Qiu H, Lee JM, Su Yoon H, Weber APM, Price DC. 2018. When less is more: red algae as models for studying gene loss and genome evolution in eukaryotes. *Critical Reviews in Plant Sciences* **37**: 81–99.

Bodył A. 2017. How to lose the plasmalemma? Lessons from ciliates, dinoflagellates and euglenids. *BioEssays* **39**: 1700149.

Bodył A, Stiller JW, Mackiewicz P. 2009. Chromalveolate plastids: direct descent or multiple endosymbioses? *Trends in Ecology and Evolution* **24**: 119–121.

Cano M, Holland SC, Artier J, Burnap RL, Ghirardi M, Morgan JA, Yu J. 2018. Glycogen synthesis and metabolite overflow contribute to energy balancing in cyanobacteria. *Cell Reports* **23**: 667–672.

Cenci U, Qiu H, Pillonel T, Cardol P, Remacle C, Colleoni C, Kadouche D, Chabi M, Greub G, Bhattacharya D *et al.* 2018. Host-pathogen biotic interactions shaped vitamin K metabolism in Archaeplastida. *Scientific Reports* **8**: 15243.

Curtis BA, Tanifuji G, Burki F, Gruber A, Irimia M, Maruyama S, Arias MC, Ball SG, Gile GH, Hirakawa Y *et al.* 2012. Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs. *Nature* **492**: 59–65.

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

Delaye L, Valadez-cano C, Pérez-zamorano B. 2016. How really ancient is *Paulinella chromatophora*?. *PLoS Currents Tree of life* **8**: ecurrents.tol.e68a099364bb1a1e129a17b4e06b0c6b.

Deschamps P, Moreira D. 2012. Reevaluating the green contribution to diatom genomes. *Genome Biology and Evolution* **4**: 683–688.

Domman D, Horn M, Embley TM, Williams TA. 2015. Plastid establishment did not require a chlamydial partner. *Nature Communications* **6**: 6421.

Dorrell RG, Gile G, McCallum G, Méheust R, Bapteste EP, Klinger CM, Brillet-Guéguen L, Freeman KD, Richter DJ, Bowler C. 2017. Chimeric origins of ochrophytes and haptophytes revealed through an ancient plastid proteome. *eLife* **6**: e23717.

Eme L, Sharpe SC, Brown MW, Roger AJ. 2014. On the age of eukaryotes: evaluating evidence from fossils and molecular clocks. *Cold Spring Harbor Perspectives in Biology* **6**: a016139.

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

Gagat P, Mackiewicz P. 2017. *Cymbomonas tetramitiformis* - a peculiar prasinophyte with a taste for bacteria sheds light on plastid evolution. *Symbiosis* **71**: 1–7.

Gavelis GS, Gile GH. 2018. How did cyanobacteria first embark on the path to becoming plastids? Lessons from protist symbioses. *FEMS Microbiology Letters* **365**: fny209.

Green BR. 2011. Chloroplast genomes of photosynthetic eukaryotes. *The Plant Journal* **66**: 34–44.

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

Jackson C, Knoll AH, Chan CX, Verbruggen H. 2018. Plastid phylogenomics with broad taxon sampling further elucidates the distinct evolutionary origins and timing of secondary green plastids. *Scientific Reports* **8**: 1523.

Kamikawa R, Tanifuji G, Kawachi M, Miyashita H, Hashimoto T, Inagaki Y. 2015. Plastid genome-based phylogeny pinpointed the origin of the green-colored plastid in the dinoflagellate *Lepidodinium chlorophorum*. *Genome Biology and Evolution* **7**: 1133–1140.

Karkar S, Facchinelli F, Price DC, Weber APM, Bhattacharya D. 2015. Metabolic connectivity as a driver of host and endosymbiont integration. *Proceedings of the National Academy of Sciences, USA* **112**: 10208–10215.

Keeling PJ. 2010. The endosymbiotic origin, diversification and fate of plastids. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**: 729–748.

Larkum AW, Lockhart PJ, Howe CJ. 2007. Shopping for plastids. *Trends in Plant Science* **12**: 189–195.

Lee J, Cho CH, Park SI, Choi JW, Song HS, West JA, Bhattacharya D, Yoon HS. 2016. Parallel evolution of highly conserved plastid genome architecture in red seaweeds and seed plants. *BMC Biology* **14**: 75.

Mackiewicz P, Bodył A, Gagat P. 2012. Protein import into the photosynthetic organelles of *Paulinella chromatophora* and its implications for primary plastid endosymbiosis. *Symbiosis* **58**: 99–107.

Marin B, Nowack ECM, Melkonian M. 2005. A plastid in the making: evidence for a second primary endosymbiosis. *Protist* **156**: 425–432.

Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D. 2002. Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proceedings of the National Academy of Sciences, USA* **99**: 12246–12251.

Matsuo E, Inagaki Y. 2018. Patterns in evolutionary origins of heme, chlorophyll a and isopentenyl diphosphate biosynthetic pathways suggest non-photosynthetic periods prior to plastid replacements in dinoflagellates. *PeerJ* **6**: e5345.

Méheust R, Zelzion E, Bhattacharya D, Lopez P, Bapteste E. 2016. Protein networks identify novel symbiogenetic genes resulting from plastid endosymbiosis. *Proceedings of the National Academy of Sciences, USA* **113**: 3579–3584.

Minge MA, Shalchian-Tabrizi K, Tørresen OK, Takishita K, Probert I, Inagaki Y, Klaveness D, Jakobsen KS. 2010. A phylogenetic mosaic plastid proteome and unusual plastid-targeting signals in the green-colored dinoflagellate *Lepidodinium chlorophorum*. *BMC Evolutionary Biology* **10**: 191.

Moreira D, Deschamps P. 2014. What was the real contribution of endosymbionts to the eukaryotic nucleus? Insights from photosynthetic eukaryotes. *Cold Spring Harbor Perspectives in Biology* **6**: a016014.

Moreira D, Philippe H. 2001. Sure facts and open questions about the origin and evolution of photosynthetic plastids. *Research in Microbiology* **152**: 771–780.

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

Muñoz-Gómez SA, Mejía-Franco FG, Durnin K, Colp M, Grisdale CJ, Archibald JM, Slamovits CH. 2017. The new red algal subphylum Proteorhodophytina comprises the largest and most divergent plastid genomes known. *Current Biology* **27**: 1677–1684.

Nowack ECM, Melkonian M, Glöckner G. 2008. Chromatophore genome sequence of *Paulinella* sheds light on acquisition of photosynthesis by eukaryotes. *Current Biology* **18**: 410–418.

Nowack ECM, Price DC, Bhattacharya D, Singer A, Melkonian M, Grossman AR. 2016. Gene transfers from diverse bacteria compensate for reductive genome evolution in the chromatophore of *Paulinella chromatophora*. *Proceedings of the National Academy of Sciences, USA* **113**: 12214–12219.

Parfrey LW, Lahr DJG, Knoll AH, Katz LA. 2011. Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proceedings of the National Academy of Sciences, USA* **108**: 13624–13629.

Petersen J, Ludewig AK, Michael V, Bunk B, Jarek M, Baurain D, Brinkmann H. 2014. *Chromera velia*, endosymbioses and the rhodoplex hypothesis - plastid

evolution in cryptophytes, alveolates, stramenopiles, and haptophytes (CASH lineages). *Genome Biology and Evolution* **6**: 666–684.

Ponce-Toledo RI, Deschamps P, López-García P, Zivanovic Y, Benzerara K, Moreira D. 2017. An early-branching freshwater cyanobacterium at the origin of plastids. *Current Biology* **27**: 386–391.

Ponce-Toledo RI, Moreira D, López-García P, Deschamps P. 2018. Secondary plastids of euglenids and chlorarachniophytes function with a mix of genes of red and green algal ancestry. *Molecular Biology and Evolution* **35**: 2198–2204.

Qiu H, Price DC, Weber APM, Facchinelli F, Yoon HS, Bhattacharya D. 2013. Assessing the bacterial contribution to the plastid proteome. *Trends in Plant Science* **18**: 680–687.

Reyes-Prieto A, Hackett JD, Soares MB, Bonaldo MF, Bhattacharya D. 2006. Cyanobacterial contribution to algal nuclear genomes is primarily limited to plastid functions. *Current Biology* **16**: 2320–2325.

Sato N, Awai K. 2017. 'Prokaryotic pathway' is not prokaryotic: noncyanobacterial origin of the chloroplast lipid biosynthetic pathway revealed by comprehensive phylogenomic analysis. *Genome Biology and Evolution* **9**: 3162–3178.

Singer A, Poschmann G, Mühlich C, Valadez-Cano C, Hänsch S, Hüren V, Rensing SA, Stühler K, Nowack ECM. 2017. Massive protein import into the early-evolutionary-stage photosynthetic organelle of the amoeba *Paulinella chromatophora*. *Current Biology* **27**: 2763–2773.

Tyra HM, Linka M, Weber AP, Bhattacharya D. 2007. Host origin of plastid solute transporters in the first photosynthetic eukaryotes. *Genome Biology* **8**: R212.

Woehle C, Dagan T, Landan G, Vardi A, Rosenwasser S. 2017. Expansion of the redox-sensitive proteome coincides with the plastid endosymbiosis. *Nature Plants* **3**: 17066.

Yang Y, Maruyama S, Sekimoto H, Sakayama H, Nozaki H. 2011. An extended phylogenetic analysis reveals ancient origin of 'non-green' phosphoribulokinase genes from two lineages of 'green' secondary photosynthetic eukaryotes: Euglenophyta and Chlorarachniophyta. *BMC Research Notes* **4**: 330.

Yoon HS, Hackett JD, Pinto G, Bhattacharya D. 2002. The single, ancient origin of chromist plastids. *Proceedings of the National Academy of Sciences, USA* **99**: 15507–15512.

Záhonová K, Füssy Z, Birčák E, Novák Vanclová AMG, Klimeš V, Vesteg M, Krajčovič J, Oborník M, Eliáš M. 2018. Peculiar features of the plastids of the colourless alga *Euglena longa* and photosynthetic euglenophytes unveiled by transcriptome analyses. *Scientific Reports* **8**: 17012.

Zhang R, Nowack ECM, Price DC, Bhattacharya D, Grossman AR. 2017. Impact of light intensity and quality on chromatophore and nuclear gene expression in *Paulinella chromatophora*, an amoeba with nascent photosynthetic organelles. *The Plant Journal* **90**: 221–234.