Plastid Origin and Evolution: New Models Provide Insights into Old Problems[1]

Cheong Xin Chan, Jeferson Gross, Hwan Su Yoon, and Debashish Bhattacharya*

Department of Ecology, Evolution and Natural Resources, and Institute of Marine and Coastal Sciences, Rutgers University, New Brunswick, New Jersey 08901 (C.X.C., J.G., D.B.); Bigelow Laboratory for Ocean Sciences, West Boothbay Harbor, Maine 04575 (H.S.Y.); and Department of Biological Sciences, Sungkyunkwan University, Suwon 440–746, South Korea (H.S.Y.)

Algae are defined by their photosynthetic organelles (plastids) that have had multiple independent origins in different phyla. These instances of organelle transfer significantly complicate inference of the tree of life for eukaryotes because the intracellular gene transfer (endosymbiotic gene transfer [EGT]) associated with each round of endosymbiosis generates highly chimeric algal nuclear genomes. In this Update we review the current state in the field of endosymbiosis research with a focus on the use of the photosynthetic amoeba *Paulinella* to advance our knowledge of plastid evolution and current ideas about the origin of the plastid translocons. These research areas have been revolutionized by the advent of modern genomic approaches.

ALGAE AND PLANTS IN THE EVOLVING TREE OF LIFE

Plastids (e.g. chloroplasts in plants) are chlorophyll-containing, membrane-bound organelles that are the sites of photosynthesis in a cell. The first plastid is thought to have originated through primary endosymbiosis, in which a photosynthetic cyanobacterium was captured by a heterotrophic protist and eventually transformed into an intracellular organelle. Molecular clock analysis suggests this key event in eukaryote evolution occurred approximately 1.5 billion years ago (Yoon et al., 2004; see Douzery et al., 2004 for another perspective). This chimeric eukaryote was an evolutionary success story and gave rise to the three major photosynthetic lineages, the red, green (including land plants), and glaucophyte algae (Bhattacharya and Medlin, 1995; Delwiche et al., 1995; McFadden, 2001; Bhattacharya et al., 2004). The supergroup Plantae or Archaeplastida (Adl et al., 2005) was established to include these three primary plastid-containing groups that contain a double-membrane-bound photosynthetic organelle lying free in the cytosol (Cavalier-Smith, 1998). Plantae monophyly has been recovered when inferring trees with plastid genes (Yoon et al., 2002; Rodríguez-Ezpeleta et al., 2005; Kim and Archibald, 2010), nuclear genes that encode plastid-targeted proteins (Li et al., 2006; Nosenko et al., 2006; Deschamps and Moreira, 2009), and rarely, single nuclear genes (Moreira et al., 2000). The strongest support for Plantae monophyly comes from a 143-gene phylogeny with 34 taxa (Rodríguez-Ezpeleta et al., 2005), a 16-gene tree with 46 taxa that included all of the eukaryotic supergroups (Hackett et al., 2007), and trees inferred from >125 protein data sets (Burki et al., 2007, 2008). However, recent multigene studies based on genome data have often shown Plantae to be polyphyletic (Nozaki et al., 2009; Baurain et al., 2010). Some of these uncertainties likely stem from methodological limitations in phylogenetic inference (Stiller, 2007) and the impact of endosymbiotic or horizontal gene transfer (E/HGT) on the evolution of microbial eukaryote genomes (Keeling and Palmer, 2008). Nevertheless, this situation is also likely explained by the absence of broadly sampled red algal genes for phylogenomic analysis, and limited availability of EST data from the glaucophytes. All of the existing analyses have for example relied on complete genome data from the thermoacidophilic Cyanidiales red alga *Cyanidioschyzon merolae* (Matsuzaki et al., 2004) and partial genome data from its sister *Galdieria sulphuraria* (http://genomics.msu.edu/galdieria/). The Cyanidiales have reduced and specialized genomes and includes only seven described species. These highly derived taxa do not represent the taxonomic breadth (approximately 6,000 species) of Rhodophyta, most of which are mesophilic (Yoon et al., 2006a, 2010; Reeb and Bhattacharya, 2010).

In light of this issue, a recent analysis of genome data from two mesophilic red algae, *Porphyridium cruentum* (extensive ESTs) and *Calliarthron tuberculosum* (draft nuclear genome assembly) uncovered a strong phylogenetic signal for monophyly of the red + green lineages that is 4-fold greater than when only *C. merolae* is used to represent the rhodophytes (Chan et al., 2011). The addition of 60,000 novel genes from *P. cruentum* and *C. tuberculosum* in the phylogenomic analysis also showed that approximately 50% of red algal genes are shared with other eukaryotes and prokaryotes putatively via E/HGT (Fig. 1). This level

[1] This work was supported by the National Science Foundation (grant nos. EF 08–27023 and DEB 09–36884 [to H.S.Y. and D.B.] and MCB 09–46528 [to D.B.]).

* Corresponding author; e-mail bhattacharya@aesop.rutgers.edu.

www.plantphysiol.org/cgi/doi/10.1104/pp.111.173500

of gene sharing can potentially complicate resolution of the Plantae supergroup and of lineages that are recipients of Plantae genes (see below). These genome-wide data do however go a long way toward substantiating the Plantae hypothesis, at least for the red + green lineages (glaucophyte complete genome is currently being analyzed [D. Bhattacharya, unpublished data]). The evolution of many key synapomorphies shared by Plantae that are often associated with photosynthesis or other plastid functions, e.g. origin of Calvin cycle enzymes (Reyes-Prieto and Bhattacharya, 2007), plastid translocon components (Weber et al., 2006), solute transporters (Colleoni et al., 2010), and dozens of unique Chlamydiae gene transfers (Huang and Gogarten, 2007; Moustafa et al., 2008) can now be interpreted with some confidence under the parsimonious scenario of a single origin in the Plantae ancestor.

The difficulties faced in resolving the issue of Plantae monophyly pale however in comparison to the phylogeny of algal groups that harbor a red algal-derived plastid that contains chlorophyll *c*. This supergroup was united by Cavalier-Smith (1998) under the moniker Chromalveolata and originally contained the alveolate (ciliates, apicomplexans, dinoflagellates) and chromist (stramenopiles, cryptophytes, haptophytes) protist lineages. These taxa were hypothesized to share a single, ancient red algal secondary endosymbiosis that gave rise to the chromophyte plastid present in many photosynthetic chromalveolates such as diatoms and haptophytes and lost secondarily in groups like ciliates and the plastid-lacking cryptophyte lineage *Goniomonas*. Recently several other protist groups such as telonemids (Shalchian-Tabrizi et al., 2006), katablepharids, centrohelids (Okamoto and Inouye, 2005; Okamoto et al., 2009), picobiliphytes (Not et al., 2007), and rappemonads (Kim et al., 2011) have been suggested to share an affiliation with the chromalveolates. Undoubtedly the quickening pace of environmental genomics research will turn up other candidate taxa affiliated with the ever-growing assemblage of chromalveolates that no longer bears any resemblance to the taxonomic scheme that was originally envisaged by Cavalier-Smith (1998). The monophyly of this lineage has been tested by many groups with initial multigene analyses supporting the supergroup but more recently, a number of articles have provided evidence against its existence or specifically, the hypothesis of a single red algal secondary endosymbiosis among all chlorophyll *c*-containing chromalveolates (e.g. Baurain et al., 2010; Parfrey et al., 2010). The question of chromalveolate monophyly, the branching order of its constituent taxa, and the history of plastid endosymbiosis in its photosynthetic mem-

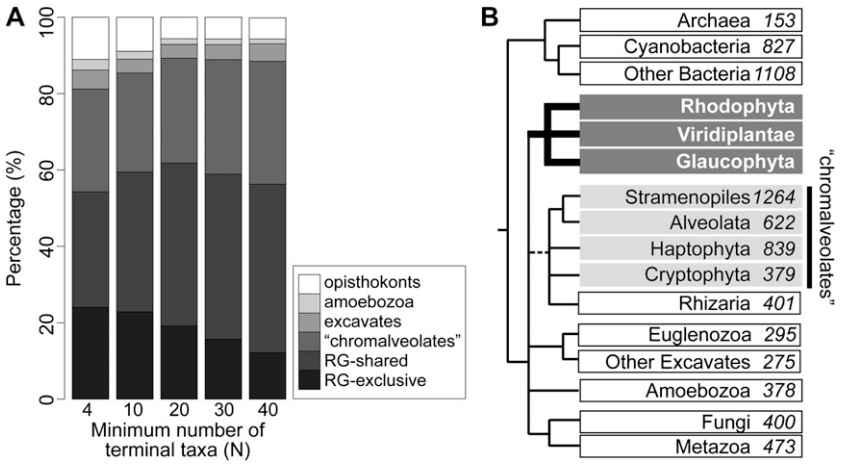

Figure 1. Evolutionary history of red algal genes as determined in Chan et al. (2011). A, Output from an automated phylogenetic pipeline was analyzed to identify the percentage of maximum likelihood (RAxML) trees that support different groupings of red and green (RG) algae with other eukaryotes at bootstrap support ≥90%. The impact of increasing the number of terminal taxa in each RAxML tree (N) on the percentage of shared genes is shown for taxa 4 → 10 → 20 → 30 → 40. The category RG-exclusive are trees in which RG algae and another phylum form an exclusive clade, whereas RG-shared are trees in which RG monophyly is recovered but other phyla are positioned within this clade (i.e. due to gene sharing explained by E/HGT). Surprisingly, nearly 50% of RG genes are shared with other eukaryotes due to E/HGT. Due to the expectation of gene transfer associated with algal endosymbiosis we do not interpret this result as evidence for monophyly of RG algae with groups such as alveolates. Rather, it is more likely that RG genes have been transferred to the nucleus of some alveolates and the signal of algal ancestry is recovered in the phylogenies. B, Schematic representation of the tree of life showing the extent of RG sharing with other eukaryotes and with prokaryotes indicated by the numbers to the right of the phylum names. The branch shown as a dashed line represents ambiguous relationships among the lineages to the right. As would be predicted, the chromalveolates (monophyly still in question) shown in the gray field are recipients of many RG genes due to endosymbiosis (see Moustafa et al., 2009), whereas the Euglenozoa share many algal genes due to their green alga-derived plastid and associated EGT (e.g. Durnford and Gray, 2006). Significant gene sharing however extends between these recipient lineages. See Chan et al. (2011) for more details.

bers (Baurain et al., 2010; Janouškovec et al., 2010) remain as challenging issues that are yet to be clarified. More important than getting the tree right is of course why it has proven so difficult in the first place when other similarly ancient groups such as opisthokonts and Amoebozoa appear monophyletic with strong bootstrap support in many multigene trees (e.g. Parfrey et al., 2006). The reason for this dichotomy may be explained by a complex history of reticulate gene evolution in chromalveolates that has not yet been fully understood (Stiller, 2007). E/HGT can produce a reticulate origin of gene fragments (in addition to whole genes) that would mislead inference of phylogenetic relationships (Chan et al., 2009). The Chan et al. (2011) analysis of mesophilic red algal genome data shows how widespread red genes are in the tree of life due to E/HGT. The role of red and green algal genomes as donors of genes means that unless carefully controlled for, many protist groups that have been recipients of Plantae genes may group with these taxa in phylogenies not due to vertical ancestry but rather E/HGT (Fig. 1), that may or may not be recognized in all single gene trees.

A striking example of unanticipated, large-scale E/HGT is provided by a recent phylogenomic analysis of predicted proteins in the completed genomes of the diatoms *Thalassiosira pseudonana* and *Phaeodactylum tricornutum* that showed hundreds of genes in these taxa to be derived from green algae (in particular prasinophytes; Moustafa et al., 2009). These data significantly complicate our understanding of chromalveolate evolution and may be explained by a cryptic green algal endosymbiosis that predated the canonical red algal capture. Under this scenario, the presence of a red alga-derived plastid in many chromalveolates conceals a past endosymbiosis, with the green E/HGTs acting as footprints of this ancient event (for discussion, see Elias and Archibald, 2009; Moustafa et al., 2009). Similarly, whereas the chlorarachniophyte *Bigelowiella natans* contains a large number of green alga-derived nuclear genes to support the function of its green plastid, a significant number of red algal-derived, plastid-targeted proteins also exist in this lineage (Archibald et al., 2003). A similar situation was recently described for the dinoflagellate *Lepidodinium chlorophanum* that contains a green alga-derived plastid but has collected nuclear genes that encode plastid-targeted proteins from multiple different algal lineages (Minge et al., 2010). These data suggest a more complex history of E/HGT in chlorarachniophytes and *L. chlorophanum* than would be expected based on plastid identity.

Another potential source of confusion wrought by endosymbiosis is plastid loss. The absence of a plastid in chromalveolates such as telonemids and katablepharids does not necessarily mean these genomes have been protected from EGT. As suggested for ciliates (Archibald, 2008; Reyes-Prieto et al., 2008), other heterotrophic lineages such as katablepharids and telonemids with phylogenetic affiliations to photosynthetic chromalveolates (Okamoto and Inouye, 2005; Okamoto et al., 2009) may have retained algal-derived genes that were introduced by EGT in their plastid-bearing ancestor, and survived due to a plastid-independent function. These unanticipated levels of genome complexity may partly explain the widely divergent views on Plantae and chromalveolate phylogeny that have regularly appeared in recent years (Burki et al., 2007; Nozaki et al., 2007; Patron et al., 2007; Kim and Graham, 2008; Baurain et al., 2010).

Rather than a continued focus on currently unresolved controversies regarding the structure of the algal and plant tree of life, we devote the remainder of the Update to recent advances in the understanding of plastid evolution that provide novel insights into the process of endosymbiosis. The two major areas to be covered are: (1) use of the photosynthetic amoeba *Paulinella* to understand primary plastid origin, and (2) models to explain the establishment of host control of plastid metabolism and evolution of the plastid machinery for protein import (the organelle translocons).

PAULINELLA AND PRIMARY PLASTID ORIGIN

Paulinella chromatophora is a thecate, filose amoeba (Bhattacharya et al., 1995) with blue-green chromatophores (Fig. 2) that was first described in 1895 by Robert Lauterborn (Lauterborn, 1895). With respect to Plantae, this species has undergone an independent primary (cyanobacterial) plastid acquisition (Kies, 1974; Marin et al., 2005; Melkonian and Mollenhauer, 2005; Rodríguez-Ezpeleta and Philippe, 2006; Reyes-Prieto et al., 2010) and consequently, *P. chromatophora* is an outstanding model for understanding plastid establishment. Sequence analysis of the nuclear small-subunit rDNA gene shows the photosynthetic amoeba to be closely related to species in the order Euglyphida (phylum Cercozoa, supergroup Rhizaria), whereas the *Paulinella* plastid rDNA is most closely related to *Synechococcus* WH5701 within a larger clade of *Prochlorococcus*- and *Synechococcus*-type cyanobacteria (Yoon et al., 2006b, 2009; Nowack et al., 2011), i.e. PS clade shown in Figure 2. The *P. chromatophora* plastid, also known as a cyanelle, retains cyanobacterial features such as a peptidoglycan wall and carboxysomes. Nonetheless several lines of evidence suggest the cyanelle is a bona fide endosymbiotic-derived organelle. The cyanelle is not bound by a vacuolar membrane and lies free in the cytoplasm, its number is regulated (i.e. two cyanelles per mature host cell) implying genetic integration, the cyanelle cannot be cultured without the host (Kies, 1974; Kies and Kremer, 1979; Johnson et al., 1988), and its genome is reduced in size to approximately 1 Mbp, that is about one-third of the size of the putative cyanobacterial donor genome (Nowack et al., 2008; Reyes-Prieto et al., 2010), demonstrating significant decay of genetic potential of the endosymbiont.

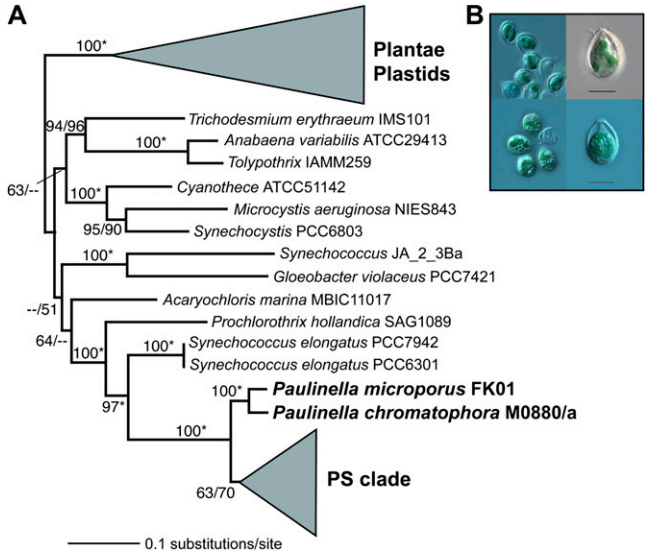

Figure 2. Evolution of photosynthetic *Paulinella* species. A, Schematic maximum likelihood (RAxML) phylogenetic tree of plastid-derived 16S rDNA from *P. chromatophora*, *P. microporus*, algal and plant (Plantae) plastids, and the cyanobacterial donors of these organelles. Note the clear independent origins of Plantae and photosynthetic *Paulinella* plastids. The numbers at the nodes show support values derived from a RAxML bootstrap analysis followed by those from a PhyML analysis. When same, this is marked with an asterisk and when a node is not resolved with a method than this is denoted with dashes. Only bootstrap values ≥50% are shown. Branch lengths are proportional to the number of substitutions per site (see scale bar). See Yoon et al. (2009) for details. B, Light micrograph images of *P. chromatophora* (top two images) and *P. microporus* (bottom two images). The scale bar indicates 5 μm.

The complete plastid genome of *P. chromatophora* M0880/a was recently published (Nowack et al., 2008). Following this, the plastid genome from another photosynthetic strain isolated in Japan, *Paulinella microporus* FK01 (Yoon et al., 2009) was determined by pyrosequencing whole-genome amplified DNA isolated via flow cytometric single-plastid sorting (Reyes-Prieto et al., 2010). Although greatly reduced in size, these plastid genomes show strong conservation in gene order when compared to the donor PS clade. The size of the *Paulinella* plastid genome however far exceeds the typical genome sizes for plastids (approximately 100–200 Kbp) in algae and plants, indicating an instance of work in progress with regard to gene loss. Based on the mode and tempo of plastid genome reduction, Nowack et al. (2008) postulated the minimum age of primary endosymbiosis in photosynthetic *Paulinella* to be 60 million years.

Alignment of *P. microporus* and *P. chromatophora* plastid DNA indicates strong conservation of gene order with only five inversions involving fragments of sizes 110.3 Kbp, 24.9 Kbp, 3.9 Kb, 2.1 Kbp, and 569 bp, as well as a single 9.2-Kbp translocation that distinguishes them. Comparison of gene inventories shows that 27 genes encoded in *P. microporus* are absent from *P. chromatophora*, whereas 39 genes in the latter are absent from *P. microporus*. These 66 genes are present in PS clade members, suggesting lineage-specific losses in these amoebae. The inspection of 681 pairwise alignments of plastid protein-encoding genes reveal K_a/K_s ratios <1 in the vast majority, consistent with strong purifying selection (Reyes-Prieto et al., 2010). These data highlight the evolutionary forces that impact organelle genomes and provide empirical evidence that differential gene losses, as well as selection to maintain protein function, are key characteristics of plastid genome evolution.

The *Paulinella* model also provides an excellent opportunity to study the phenomenon of EGT and to investigate its fundamental role in organellogenesis (see Gross and Bhattacharya, 2009a). The marked plastid genome reduction in *Paulinella* species, suggests that approximately 2 Mbp of cyanobacterial sequence has either been lost outright or transferred to the nuclear genome of the amoeba. Demonstration of differential gene loss suggests that both forces have clearly impacted this system. Evidence for EGT in *Paulinella* was initially reported for genes encoding PSI subunits, *psa*E and *psa*I (Nakayama and Ishida, 2009; Reyes-Prieto et al., 2010). Cyanobacterial *psa*I that encodes subunit VIII of PSI was silenced by two nonsense mutations in the *P. microporus* plastid genome, whereas an intact copy with a 198-nt spliceosomal intron was found in the amoeba nuclear genome. This suggests that the plastid copy of *psa*I likely become a pseudogene after activation of the transferred nuclear gene. In contrast, *P. chromatophora* retains *psa*I in the plastid genome. A more-recent analysis of 32,012 ESTs from *P. chromatophora* (Nowack et al., 2011) provide 32 examples of EGT (a majority of the implicated genes are involved in photosynthesis), and a recent example of intron insertion in different regions of *psa*E that distinguish *P. chromatophora* from *P. microporus*. These authors also reported two expressed genes (encoding *psa*K and high-light-inducible proteins) out of 19 that are plastid encoded in *P. microporus* but absent in the *P. chromatophora* organelle, suggesting differential EGT after the split of these taxa.

Given these exciting observations of EGT, a key area of future research with *Paulinella* will be to understand how nuclear-encoded proteins are targeted to the nascent plastid. One possibility is that delivery into and across the cyanelle outer membrane (OM) may involve vesicular transport via the endomembrane system (see Bhattacharya et al., 2007). This is observed for α-carbonic anhydrase in Arabidopsis (*Arabidopsis thaliana*; Villarejo et al., 2005) and is the standard pathway for targeting proteins to the complex plastids of euglenophytes, Apicomplexa, and peridinin-containing dinoflagellates (Bolte et al., 2009). Interestingly, the conspicuous absence in the reduced cyanelle genome of many transport components that control the flux of metabolites across the cyanobacterial cell membranes suggests that an irreversible, long-term metabolic and cell biological connection, between host and

plastid in *P. chromatophora* has already developed (Nowack et al., 2008). This could be primarily controlled by key host-derived metabolite transporters integrated into the OM of the endosymbiont, such as PfoTPT in *Plasmodium* (Mullin et al., 2006), or a wholesale reprogramming of the inner membrane (IM) permeome of the new organelle with host transporters (Weber et al., 2006; Tyra et al., 2007). Such a phenomenon is predicted by our recently postulated outsiders' hypothesis that posits an outside-to-inside host-guided establishment of protein-sorting components in an endosymbiotic-derived organelle; i.e. initially in the OM and then gradually inwards to the organelle lumen (Gross and Bhattacharya, 2009a). The model predicts that EGT of molecular components that act in the organelle interior is explained by an advanced stage in the evolution of organelle protein topogenesis. EGT has already occurred for key genes that control photosynthesis and cyanelle function in *Paulinella*, and determining the final inventory of transfers and gene activations (expression) in this system is currently under way using genome-wide approaches (H.S. Yoon and D. Bhattacharya, unpublished data).

A fascinating addition to the *Paulinella* story is provided by its closely related sister taxon *Paulinella ovalis*. This protist feeds actively on cyanobacteria and other prey that have been identified in food vacuoles in its cytoplasm (Johnson et al., 1988). Comparison of these amoebal genomes may allow the identification of genetic innovations (e.g. HGT events and gene duplication of loci potentially involved in plastid maintenance) that underlie the critical transition from heterotrophy to obligate autotrophy (Moustafa et al., 2008). Recent work in our lab has identified cyanobacterial DNA fragments that are putatively derived from prey ingested by *P. ovalis* cells captured in nature. Extensive sequencing of these single-cell derived genomic DNAs using next-generation sequencing has shown that many cyanobacterial fragments are derived from PS clade taxa and their cyanophages (D. Price, H.S. Yoon, and D. Bhattacharya, unpublished data). This result suggests a possible link between feeding behavior in a phagotroph and the source of a novel plastid in its photosynthetic sister.

EVOLUTION OF PLASTID PROTEIN SORTING

As an intracellular compartment in eukaryotes, plastids rely on thousands of nuclear-encoded proteins that are translated by cytosolic ribosomes and then moved and assembled into the organelle subcompartments (Gould et al., 2008; Gross and Bhattacharya, 2009b; Li and Chiu, 2010). However plastids derive from a cyanobacterial endosymbiont that was presumably free from nuclear control during the early stages of the intracellular association with its host. Therefore the progressive transformation of the endosymbiont into an organelle includes an increasing governance of the host nucleus over cyanobacterial functions. This must have been made possible by the evolution of a regulated system to transport (translocate) nuclear-encoded proteins into the endosymbiont-derived compartment (Gross and Bhattacharya, 2009a, 2009b). In extant algae and plants, this task is fulfilled by specialized multisubunit machines that constitute the translocons at the outer and inner envelope membranes of chloroplasts (Toc and Tic, respectively; Reumann et al., 1999; Gould et al., 2008; Kalanon and McFadden, 2008). Studies of the Toc and Tic protein import systems have primarily been done in land plants, yielding a complex picture in which more than 15 subunits serve as protein-conducting pores, receptors, chaperones, cochaperones, and regulatory subunits (Li and Chiu, 2010; see Fig. 3). This advanced level of complexity provides a challenge to evolutionary biologists interested in identifying the putative minimal set of protein-sorting subunits that supported early organelle evolution.

To improve our understanding of incipient plastid evolution we recently postulated a simple ancestral protein-sorting system in plastids (Gross and Bhattacharya, 2009a, 2009b). This system was composed solely of the Toc75 and Tic110 protein channels (at the OM and IM of the plastids, respectively), associated chaperones, and the Toc34 receptor at the plastid entrance (Fig. 3). The remainder of modern-day Toc and Tic subunits presumably accrued later to the complexes, during the evolution of green algae and land plants to improve or regulate translocon activity. Many of these subunits are members of protein families (mostly cyanobacterium derived) and likely arose via gene duplications (Kalanon and McFadden, 2008). In some cases the new paralog apparently retained the original preendosymbiotic properties (e.g. Tic55, Tic32, and Tic62 might be bona fide metabolic enzymes, and Tic20 and Tic21 are putative solute channels; Gray et al., 2004; Hörmann et al., 2004; Balsera et al., 2007; Duy et al., 2007; Gross and Bhattacharya, 2009b). This may mean that the new paralog was (1) recruited to the translocon complexes while still maintaining its previous function in a different system (multifunctionality), or alternatively (2) coopted to a new exclusive function at the translocon while still preserving its ancestral enzymatic property. Some of the Tic subunits that appear to be functional recruitments are traditionally interpreted under the view that they represent regulatory subunits of the translocon (e.g. Tic55, Tic32, and Tic62) that modulate protein import activities according to intracellular cues; e.g. redox status, calcium levels (Balsera et al., 2010). Under our view, the metabolic proteins Tic55, Tic32, and Tic62, as well as the putative channels Tic20 and Tic21, and chaperone Tic22, might act as integrative subunits (Gross and Bhattacharya, 2009b). As multifunctional proteins they could circumstantially interact with the Tic110 channel to regulate protein import according to the status of a biological process in which the protein is primarily associated. Reciprocally, their primary function might be regulated by protein import rates. For example, Tic55 is

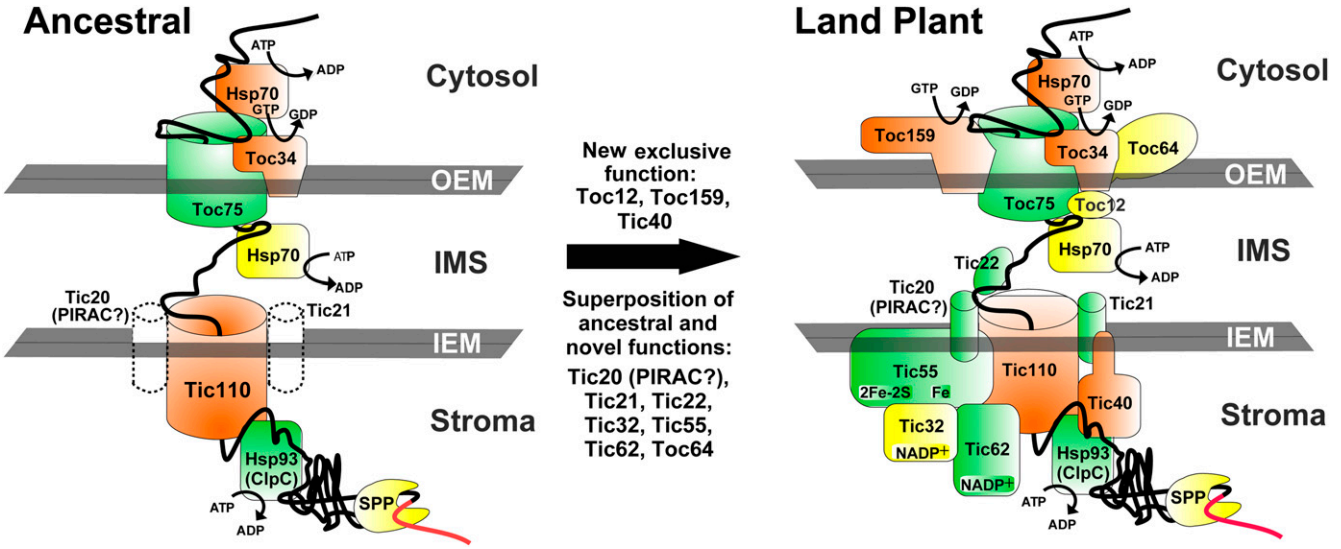

Figure 3. Schematic representation of the Toc/Tic translocons embedded in the chloroplast membranes of land plants (image on right) in the process of importing a preprotein from the cytosol (figure based on Gross and Bhattacharya, 2009b). The unfolded cytosolic protein contains a transit peptide (in red) that is initially recognized by the receptors Toc34 and Toc159 and then imported across the outer and inner envelope membranes (OEM and IEM, respectively) via Toc75 and Tic110. The energy-dependent translocation process is driven by heat-shock-type molecular chaperones (Hsp70 and Hsp93) at the cytosolic intermembrane space (IMS), and stromal sides of the organelle. The stromal processing peptidase (SPP) removes the transit peptide after translocation. The subunits of the translocon of cyanobacterial (endosymbiont) origin are shown in green, those of eukaryote origin are in red, whereas those in yellow are of unknown phylogenetic affiliation. The proposed ancestral state of the complex plant translocon is shown at the left and is composed of the Toc34, Toc75, and Tic110 subunits and the Hsp70 and Hsp93 chaperones. Tic20 and Tic21 are depicted by broken lines, indicating their possible participation in the ancestral translocon (discussed in Gross and Bhattacharya, 2009b). Evolution of the plant translocons required the addition of subunits to fulfill an exclusive new function in protein import (Toc12, Toc159, and Tic40) and subunits recruited to the translocon but maintaining ancestral properties (below the arrow). We tentatively classified Toc64 in the latter group because it shows sequence conservation with plant amidases (67% amino acid similarity over the full sequence of functional Amidase1 in Arabidopsis). The potential correspondence of protein import-related anion channel (PIRAC) to Tic20 is discussed in detail in Gross and Bhattacharya (2009b).

putatively involved in chlorophyll breakdown (Gray et al., 2004; Gross and Bhattacharya, 2009b). Hypothetically, its association with Tic110 might modulate protein import activity according to the status of chlorophyll catabolism in plastids. In contrast, Tic55 enzymatic turnover could be reciprocally fine tuned by information on protein import rates relayed by the Tic110 channel. Although speculative, this integrator hypothesis offers a novel perspective on the functions of proteins such as Tic55, Tic32, Tic62, Tic20, Tic21, and Tic22, which still lack a defined mechanistic role in protein translocation. These include (1) the possibility of multifunctionality (e.g. Teng et al., 2006; Duy et al., 2007); which implies (2) possible pleiotropic effects in gene knockout studies (e.g. Duy et al., 2007); (3) a family wide perspective to understand original and derived functions of the tested subunits (e.g. Balsera et al., 2007); (4) the use of cyanobacteria to investigate conserved ancestral properties of many translocon subunits (e.g. Lv et al., 2009); and (5) a system biology understanding of protein translocation in plastids.

Knowledge of the ancestral functions of the Toc and Tic subunits is also crucial to reconstruct the formative stages of plastid evolution. That the OM pore Toc75 is derived from a cyanobacterial Omp85 type of β-barrel assembly protein suggests the onset of plastid organellogenesis was marked by the host cell, thus gaining control over OM activities of the endosymbiont (Gross and Bhattacharya, 2009b). The Toc75 homolog OEP80 could still in modern plastids, retain this ancestral function in the biogenesis of β-barrel proteins (Patel et al., 2008). In addition, the development of plastid organellogenesis included the establishment of an IM pore to reprogram the cyanobacterial permeome predominantly with host-derived solute carriers (Tyra et al., 2007). To illuminate this crucial step in plastid evolution it is important to investigate whether special Tic subcomplexes (e.g. involving or not Tic20, Tic21, and/or Tic110) or a still-unknown pore-forming complex might control the IM permeome assembly in modern plastids (Firlej-Kwoka et al., 2008).

OUTLOOK

Developments and discoveries made in the fields of plastid origin and algal evolution have come rapidly in recent years. The foreseeable future of these fields is

rife with uncertainties and challenges, largely depending on the advancement of techniques in molecular and computational biology. There are many genome projects nearing completion; e.g. *Chondrus crispus* (florideophyte red alga), *C. tuberculosum* (coralline red alga), *Porphyra umbilicalis* (Bangiales red alga; see Blouin et al., 2011), *Guillardia theta* (cryptophyte), and *B. natans* (chlorarachniophyte). The rapid advances in sequencing, e.g. single-cell environmental genomics (Woyke et al., 2009) and functional genomics have opened up new windows into genome evolution and gene function in nonmodel taxa that were unimaginable only a decade ago. The ideas highlighted in this Update are therefore destined to change in the near future, as new data are analyzed and old hypotheses reinterpreted. In particular, work on *Paulinella* by different research groups will be crucial to test existing ideas about the evolution of organelle protein import and more generally the process of organellogenesis. We also expect the eukaryote tree of life to undergo some adjustments as investigators begin to understand better the dynamics of protist genome evolution and the limitations these forces place on phylogenetic approaches. Rather than being discouraged by this rapid pace of research we are excited to recognize that algal and plastid research have become center stage for both academic pursuits and for their key role in applied uses such as biofuel development.

ACKNOWLEDGMENTS

We appreciate the constructive comments made by two anonymous reviewers of this manuscript.

Received January 28, 2011; accepted February 19, 2010; published February 22, 2011.

LITERATURE CITED

Adl SM, Simpson AGB, Farmer MA, Andersen RA, Anderson OR, Barta JR, Bowser SS, Brugerolle G, Fensome RA, Fredericq S, et al (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. J Eukaryot Microbiol **52:** 399–451

Archibald JM (2008) Plastid evolution: remnant algal genes in ciliates. Curr Biol **18:** R663–R665

Archibald JM, Rogers MB, Toop M, Ishida K, Keeling PJ (2003) Lateral gene transfer and the evolution of plastid-targeted proteins in the secondary plastid-containing alga *Bigelowiella natans*. Proc Natl Acad Sci USA **100:** 7678–7683

Balsera M, Soll J, Buchanan BB (2010) Redox extends its regulatory reach to chloroplast protein import. Trends Plant Sci **15:** 515–521

Balsera M, Stengel A, Soll J, Bölter B (2007) Tic62: a protein family from metabolism to protein translocation. BMC Evol Biol **7:** 43

Baurain D, Brinkmann H, Petersen J, 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. Mol Biol Evol **27:** 1698–1709

Bhattacharya D, Archibald JM, Weber AP, Reyes-Prieto A (2007) How do endosymbionts become organelles? Understanding early events in plastid evolution. Bioessays **29:** 1239–1246

Bhattacharya D, Helmchen T, Melkonian M (1995) Molecular evolutionary analyses of nuclear-encoded small subunit ribosomal RNA identify an independent rhizopod lineage containing the Euglyphina and the Chlorarachniophyta. J Eukaryot Microbiol **42:** 65–69

Bhattacharya D, Medlin L (1995) The phylogeny of plastids: a review based on comparisons of small-subunit ribosomal RNA coding regions. J Phycol **31:** 489–498

Bhattacharya D, Yoon HS, Hackett JD (2004) Photosynthetic eukaryotes unite: endosymbiosis connects the dots. Bioessays **26:** 50–60

Blouin NA, Brodie JA, Grossman AC, Xu P, Brawley SH (2011) *Porphyra*: a marine crop shaped by stress. Trends Plant Sci **16:** 29–37

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

Burki F, Shalchian-Tabrizi K, Minge M, Skjaeveland Å, Nikolaev SI, Jakobsen KS, Pawlowski J (2007) Phylogenomics reshuffles the eukaryotic supergroups. PLoS ONE **2:** e790

Burki F, Shalchian-Tabrizi K, Pawlowski J (2008) Phylogenomics reveals a new 'megagroup' including most photosynthetic eukaryotes. Biol Lett **4:** 366–369

Cavalier-Smith T (1998) A revised six-kingdom system of life. Biol Rev Camb Philos Soc **73:** 203–266

Chan CX, Beiko RG, Darling AE, Ragan MA (2009) Lateral transfer of genes and gene fragments in prokaryotes. Genome Biol Evol **1:** 429–438

Chan CX, Yang EC, Banerjee T, Yoon HS, Martone PT, Estevez JM, Bhattacharya D (2011) Red and green algal monophyly and extensive gene sharing found in a rich repertoire of red algal genes. Curr Biol **21:** 328–333

Colleoni C, Linka M, Deschamps P, Handford MG, Dupree P, Weber APM, Ball SG (2010) Phylogenetic and biochemical evidence supports the recruitment of an ADP-glucose translocator for the export of photosynthate during plastid endosymbiosis. Mol Biol Evol **27:** 2691–2701

Delwiche CF, Kuhsel M, Palmer JD (1995) Phylogenetic analysis of *tuf*A sequences indicates a cyanobacterial origin of all plastids. Mol Phylogenet Evol **4:** 110–128

Deschamps P, Moreira D (2009) Signal conflicts in the phylogeny of the primary photosynthetic eukaryotes. Mol Biol Evol **26:** 2745–2753

Douzery EJP, Snell EA, Bapteste E, Delsuc F, Philippe H (2004) The timing of eukaryotic evolution: does a relaxed molecular clock reconcile proteins and fossils? Proc Natl Acad Sci USA **101:** 15386–15391

Durnford DG, Gray MW (2006) Analysis of *Euglena gracilis* plastid-targeted proteins reveals different classes of transit sequences. Eukaryot Cell **5:** 2079–2091

Duy D, Wanner G, Meda AR, von Wirén N, Soll J, Philippar K (2007) PIC1, an ancient permease in *Arabidopsis* chloroplasts, mediates iron transport. Plant Cell **19:** 986–1006

Elias M, Archibald JM (2009) Sizing up the genomic footprint of endosymbiosis. Bioessays **31:** 1273–1279

Firlej-Kwoka E, Strittmatter P, Soll J, Bölter B (2008) Import of preproteins into the chloroplast inner envelope membrane. Plant Mol Biol **68:** 505–519

Gould SB, Waller RF, McFadden GI (2008) Plastid evolution. Annu Rev Plant Biol **59:** 491–517

Gray J, Wardzala E, Yang M, Reinbothe S, Haller S, Pauli F (2004) A small family of LLS1-related non-heme oxygenases in plants with an origin amongst oxygenic photosynthesizers. Plant Mol Biol **54:** 39–54

Gross J, Bhattacharya D (2009a) Mitochondrial and plastid evolution in eukaryotes: an outsiders' perspective. Nat Rev Genet **10:** 495–505

Gross J, Bhattacharya D (2009b) Revaluating the evolution of the Toc and Tic protein translocons. Trends Plant Sci **14:** 13–20

Hackett JD, Yoon HS, Li S, Reyes-Prieto A, Rümmele SE, Bhattacharya D (2007) Phylogenomic analysis supports the monophyly of cryptophytes and haptophytes and the association of rhizaria with chromalveolates. Mol Biol Evol **24:** 1702–1713

Hörmann F, Küchler M, Sveshnikov D, Oppermann U, Li Y, Soll J (2004) Tic32, an essential component in chloroplast biogenesis. J Biol Chem **279:** 34756–34762

Huang JL, Gogarten JP (2007) Did an ancient chlamydial endosymbiosis facilitate the establishment of primary plastids? Genome Biol **8:** R99

Janouškovec J, Horák A, Oborník M, Lukes J, Keeling PJ (2010) A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. Proc Natl Acad Sci USA **107:** 10949–10954

Johnson PW, Hargraves PE, Sieburth JM (1988) Ultrastructure and ecology of *Calycomonas ovalis* Wulff, 1919, (Chrysophyceae) and its redescription as a testate Rhizopod, *Paulinella ovalis* N. Comb. (Filosea: Euglyphina). J Protozool **35:** 618–626

Kalanon M, McFadden GI (2008) The chloroplast protein translocation

complexes of *Chlamydomonas reinhardtii*: a bioinformatic comparison of Toc and Tic components in plants, green algae and red algae. Genetics **179:** 95–112

Keeling PJ, Palmer JD (2008) Horizontal gene transfer in eukaryotic evolution. Nat Rev Genet **9:** 605–618

Kies L (1974) [Electron microscopical investigations on Paulinella chromatophora Lauterborn, a thecamoeba containing blue-green endosymbionts (Cyanelles) (author's transl)]. Protoplasma **80:** 69–89

Kies L, Kremer BP (1979) Function of cyanelles in the thecamoeba *Paulinella chromatophora*. Naturwissenschaften **66:** 578–579

Kim E, Archibald JM (2010) Plastid evolution: gene transfer and the maintenance of 'stolen' organelles. BMC Biol **8:** 73

Kim E, Graham LE (2008) EEF2 analysis challenges the monophyly of Archaeplastida and Chromalveolata. PLoS ONE **3:** e2621

Kim E, Harrison JW, Sudek S, Jones MD, Wilcox HM, Richards TA, Worden AZ, Archibald JM (2011) Newly identified and diverse plastid-bearing branch on the eukaryotic tree of life. Proc Natl Acad Sci USA **108:** 1496–1500

Lauterborn R (1895) Protozoenstudien II. *Paulinella chromatophora* nov. gen., nov. spec., ein beschalter Rhizopode des Sußwassers mit blaugrunen chromatophorenartigen Einschlussen. Z Wiss Zool **59:** 537–544

Li HM, Chiu C-C (2010) Protein transport into chloroplasts. Annu Rev Plant Biol **61:** 157–180

Li S, Nosenko T, Hackett JD, Bhattacharya D (2006) Phylogenomic analysis identifies red algal genes of endosymbiotic origin in the chromalveolates. Mol Biol Evol **23:** 663–674

Lv HX, Guo GQ, Yang ZN (2009) Translocons on the inner and outer envelopes of chloroplasts share similar evolutionary origin in *Arabidopsis thaliana*. J Evol Biol **22:** 1418–1428

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

Matsuzaki M, Misumi O, Shin-I T, Maruyama S, Takahara M, Miyagishima SY, Mori T, Nishida K, Yagisawa F, Nishida K, et al (2004) Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. Nature **428:** 653–657

McFadden GI (2001) Chloroplast origin and integration. Plant Physiol **125:** 50–53

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

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 Evol Biol **10:** 191

Moreira D, Le Guyader H, Philippe H (2000) The origin of red algae and the evolution of chloroplasts. Nature **405:** 69–72

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

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

Mullin KA, Lim L, Ralph SA, Spurck TP, Handman E, McFadden GI (2006) Membrane transporters in the relict plastid of malaria parasites. Proc Natl Acad Sci USA **103:** 9572–9577

Nakayama T, Ishida K-i (2009) Another acquisition of a primary photosynthetic organelle is underway in *Paulinella chromatophora*. Curr Biol **19:** R284–R285

Nosenko T, Lidie KL, Van Dolah FM, Lindquist E, Cheng JF, Bhattacharya D (2006) Chimeric plastid proteome in the Florida "red tide" dinoflagellate *Karenia brevis*. Mol Biol Evol **23:** 2026–2038

Not F, Valentin K, Romari K, Lovejoy C, Massana R, Töbe K, Vaulot D, Medlin LK (2007) Picobiliphytes: a marine picoplanktonic algal group with unknown affinities to other eukaryotes. Science **315:** 253–255

Nowack EC, Melkonian M, Glöckner G (2008) Chromatophore genome sequence of *Paulinella* sheds light on acquisition of photosynthesis by eukaryotes. Curr Biol **18:** 410–418

Nowack EC, Vogel H, Groth M, Grossman AR, Melkonian M, Glöckner G (2011) Endosymbiotic gene transfer and transcriptional regulation of transferred genes in *Paulinella chromatophora*. Mol Biol Evol **28:** 407–422

Nozaki H, Iseki M, Hasegawa M, Misawa K, Nakada T, Sasaki N, Watanabe M (2007) Phylogeny of primary photosynthetic eukaryotes as deduced from slowly evolving nuclear genes. Mol Biol Evol **24:** 1592–1595

Nozaki H, Maruyama S, Matsuzaki M, Nakada T, Kato S, Misawa K (2009) Phylogenetic positions of Glaucophyta, green plants (Archaeplastida) and Haptophyta (Chromalveolata) as deduced from slowly evolving nuclear genes. Mol Phylogenet Evol **53:** 872–880

Okamoto N, Chantangsi C, Horák A, Leander BS, Keeling PJ (2009) Molecular phylogeny and description of the novel katablepharid *Roombia truncata* gen. et sp. nov., and establishment of the Hacrobia taxon nov. PLoS ONE **4:** e7080

Okamoto N, Inouye I (2005) The katablepharids are a distant sister group of the Cryptophyta: a proposal for Katablepharidophyta divisio nova/Kathablepharida phylum novum based on SSU rDNA and beta-tubulin phylogeny. Protist **156:** 163–179

Parfrey LW, Barbero E, Lasser E, Dunthorn M, Bhattacharya D, Patterson DJ, Katz LA (2006) Evaluating support for the current classification of eukaryotic diversity. PLoS Genet **2:** e220

Parfrey LW, Grant J, Tekle YI, Lasek-Nesselquist E, Morrison HG, Sogin ML, Patterson DJ, Katz LA (2010) Broadly sampled multigene analyses yield a well-resolved eukaryotic tree of life. Syst Biol **59:** 518–533

Patel R, Hsu S-C, Bédard J, Inoue K, Jarvis P (2008) The Omp85-related chloroplast outer envelope protein OEP80 is essential for viability in Arabidopsis. Plant Physiol **148:** 235–245

Patron NJ, Inagaki Y, Keeling PJ (2007) Multiple gene phylogenies support the monophyly of cryptomonad and haptophyte host lineages. Curr Biol **17:** 887–891

Reeb V, Bhattacharya D (2010) The thermo-acidophilic Cyanidiophyceae (Cyanidiales). *In* J Seckbach, D Chapman, eds, Red Algae in the Genomic Age, Vol 13, Cellular Origins, Life in Extreme Habitats and Astrobiology. Springer, New York, pp 409–426

Reumann S, Davila-Aponte J, Keegstra K (1999) The evolutionary origin of the protein-translocating channel of chloroplastic envelope membranes: identification of a cyanobacterial homolog. Proc Natl Acad Sci USA **96:** 784–789

Reyes-Prieto A, Bhattacharya D (2007) Phylogeny of Calvin cycle enzymes supports Plantae monophyly. Mol Phylogenet Evol **45:** 384–391

Reyes-Prieto A, Moustafa A, Bhattacharya D (2008) Multiple genes of apparent algal origin suggest ciliates may once have been photosynthetic. Curr Biol **18:** 956–962

Reyes-Prieto A, Yoon HS, Moustafa A, Yang EC, Andersen RA, Boo SM, Nakayama T, Ishida K, Bhattacharya D (2010) Differential gene retention in plastids of common recent origin. Mol Biol Evol **27:** 1530–1537

Rodríguez-Ezpeleta N, Brinkmann H, Burey SC, Roure B, Burger G, Löffelhardt W, Bohnert HJ, Philippe H, Lang BF (2005) Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. Curr Biol **15:** 1325–1330

Rodríguez-Ezpeleta N, Philippe H (2006) Plastid origin: replaying the tape. Curr Biol **16:** R53–R56

Shalchian-Tabrizi K, Eikrem W, Klaveness D, Vaulot D, Minge MA, Le Gall F, Romari K, Throndsen J, Botnen A, Massana R, et al (2006) Telonemia, a new protist phylum with affinity to chromist lineages. Proc Biol Sci **273:** 1833–1842

Stiller JW (2007) Plastid endosymbiosis, genome evolution and the origin of green plants. Trends Plant Sci **12:** 391–396

Teng YS, Su YS, Chen LJ, Lee YJ, Hwang I, Li HM (2006) Tic21 is an essential translocon component for protein translocation across the chloroplast inner envelope membrane. Plant Cell **18:** 2247–2257

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

Villarejo A, Burén S, Larsson S, Déjardin A, Monné M, Rudhe C, Karlsson J, Jansson S, Lerouge P, Rolland N, et al (2005) Evidence for a protein transported through the secretory pathway en route to the higher plant chloroplast. Nat Cell Biol **7:** 1224–1231

Weber AP, Linka M, Bhattacharya D (2006) Single, ancient origin of a plastid metabolite translocator family in Plantae from an endomembrane-derived ancestor. Eukaryot Cell **5:** 609–612

Woyke T, Xie G, Copeland A, González JM, Han C, Kiss H, Saw JH, Senin P, Yang C, Chatterji S, et al (2009) Assembling the marine metagenome, one cell at a time. PLoS ONE **4:** e5299

Yoon HS, Ciniglia C, Wu M, Comeron JM, Pinto G, Pollio A, Bhattacharya

D (2006a) Establishment of endolithic populations of extremophilic Cyanidiales (Rhodophyta). BMC Evol Biol **6:** 78

Yoon HS, Hackett JD, Ciniglia C, Pinto G, Bhattacharya D (2004) A molecular timeline for the origin of photosynthetic eukaryotes. Mol Biol Evol **21:** 809–818

Yoon HS, Hackett JD, Pinto G, Bhattacharya D (2002) The single, ancient origin of chromist plastids. Proc Natl Acad Sci USA **99:** 15507–15512

Yoon HS, Nakayama T, Reyes-Prieto A, Andersen RA, Boo SM, Ishida K, Bhattacharya D (2009) A single origin of the photosynthetic organelle in different *Paulinella* lineages. BMC Evol Biol **9:** 98

Yoon HS, Reyes-Prieto A, Melkonian M, Bhattacharya D (2006b) Minimal plastid genome evolution in the *Paulinella* endosymbiont. Curr Biol **16:** R670–R672

Yoon HS, Zuccarello GC, Bhattacharya D (2010) Evolutionary history and taxonomy of red algae. *In* J Seckbach, D Chapman, eds, Red Algae in the Genomic Age, Vol 13, Cellular Origins, Life in Extreme Habitats and Astrobiology. Springer, New York, pp 25–42