Genomic Insights into Plastid Evolution

Shannon J. Sibbald[1,2] and John M. Archibald (iD)[1,2,*]

[1]Department of Biochemistry & Molecular Biology, Dalhousie University, Halifax, Nova Scotia, Canada
[2]Centre for Comparative Genomics and Evolutionary Bioinformatics, Dalhousie University, Halifax, Nova Scotia, Canada

*Corresponding author: E-mail: john.archibald@dal.ca.

Abstract

The origin of plastids (chloroplasts) by endosymbiosis stands as one of the most important events in the history of eukaryotic life. The genetic, biochemical, and cell biological integration of a cyanobacterial endosymbiont into a heterotrophic host eukaryote approximately a billion years ago paved the way for the evolution of diverse algal groups in a wide range of aquatic and, eventually, terrestrial environments. Plastids have on multiple occasions also moved horizontally from eukaryote to eukaryote by secondary and tertiary endosymbiotic events. The overall picture of extant photosynthetic diversity can best be described as "patchy": Plastid-bearing lineages are spread far and wide across the eukaryotic tree of life, nested within heterotrophic groups. The algae do not constitute a monophyletic entity, and understanding how, and how often, plastids have moved from branch to branch on the eukaryotic tree remains one of the most fundamental unsolved problems in the field of cell evolution. In this review, we provide an overview of recent advances in our understanding of the origin and spread of plastids from the perspective of comparative genomics. Recent years have seen significant improvements in genomic sampling from photosynthetic and nonphotosynthetic lineages, both of which have added important pieces to the puzzle of plastid evolution. Comparative genomics has also allowed us to better understand how endosymbionts become organelles.

Key words: plastids, chloroplasts, genomics, phylogenomics, organelles, algae, protists, evolution.

Introduction

Algae are diverse and ecologically important organisms found across the eukaryotic tree of life, and they all have at least one thing in common—they are photosynthetic. Although their precise evolutionary paths vary, all canonical plastids are believed to be derived from the same endosymbiotic cyanobacterium. How do we know this? Simply put, it is in their DNA. The genes that remain within modern-day plastid genomes, along with those that have relocated to the nuclear genome, provide important clues as to how plastids evolved across different branches of the eukaryotic tree. The first plant plastid genomes were sequenced in the 1980s (Ohyama et al. 1986; Shinozaki et al. 1986) and provided preliminary insight into their gene content and structure. Over 4,000 plastid genomes have now been sequenced from a wide diversity of photosynthetic and secondarily nonphotosynthetic eukaryotes. Together with nuclear genome sequences, these data have made it possible to investigate the "who, what, when, where, and how" of eukaryotic photosynthesis. Here, we re-

view recent genomics-based advances in our understanding of how plastids arose and spread. With the immense amount of information now coming from genome sequencing projects, the potential for discovery and insight into these fundamental questions is unparalleled. What is emerging is a more complete and nuanced view of plastid evolution, one that is nevertheless still lacking in important details.

The Origin of Plastids

A wealth of biochemical, molecular, and phylogenetic data support the notion that plastids evolved from endosymbiotic cyanobacteria on a single occasion (see Kim and Archibald 2009 and references therein for review). This landmark event is thought to have occurred at least ~900 Ma (e.g., Parfrey et al. 2011; Shih and Matzke 2013) in a common ancestor shared by land plants and green algae (Viridiplantae), red algae (Rhodophyta), and glaucophyte algae (Glaucophyta) (fig. 1), giving rise to so-called "primary" plastids surrounded
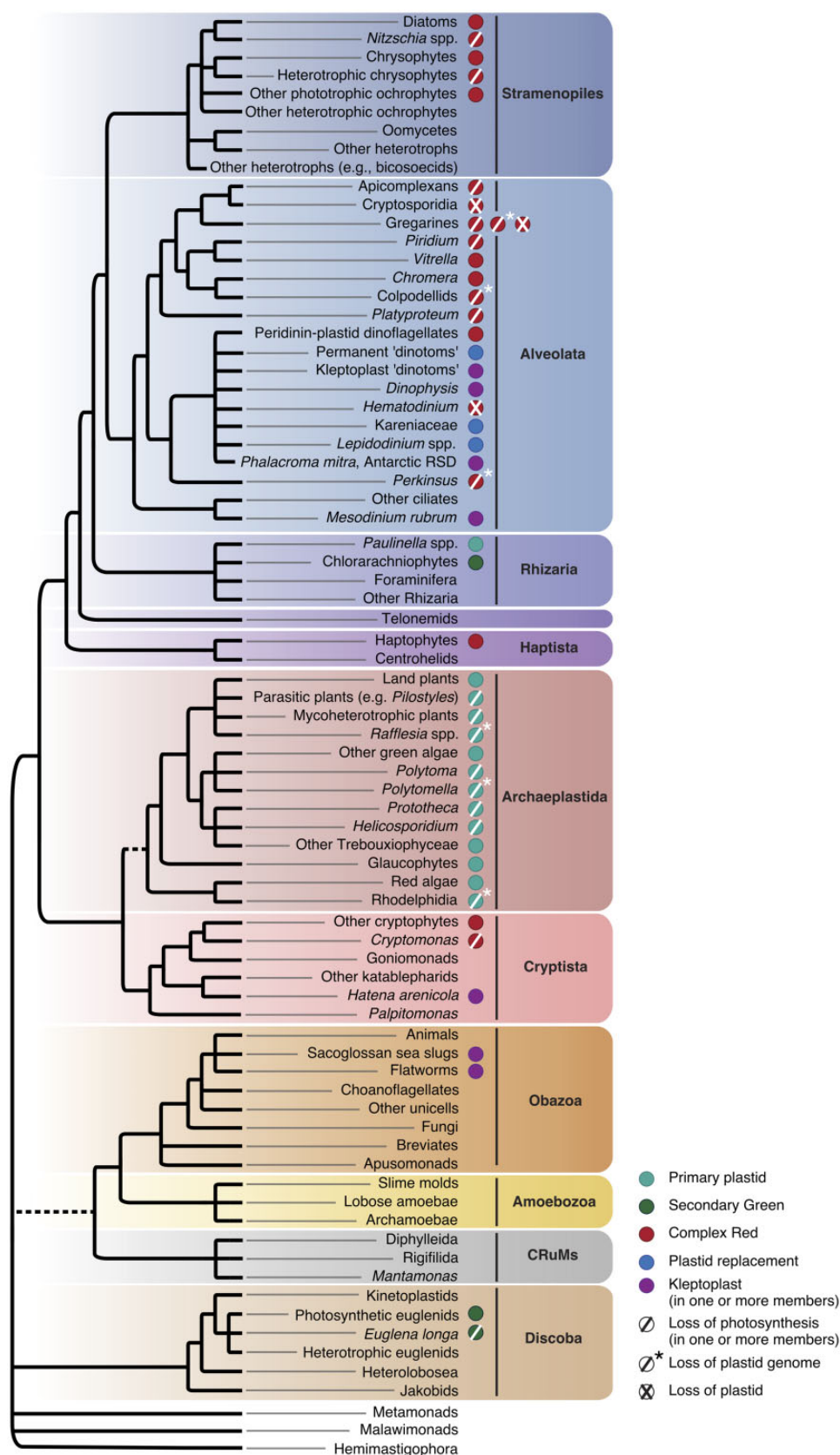
Fig. 1—Schematic of the eukaryotic tree of life with an emphasis on plastid-bearing lineages and their closest relatives. The tree topology is based on recent analyses and discussion in Strassert et al. (2019), Gawryluk et al. (2019), and Burki et al. (2020). The type of plastid (primary or complex) is indicated next to each lineage. Where known, specific complex events of kleptoplasty and plastid replacements (serial secondary or tertiary) are shown. Known instances of loss of photosynthesis are indicated with a line through the plastid circle; loss of photosynthesis with loss of the plastid genome is indicated by a line through the plastid circle and an asterisk. Complete loss of a plastid is indicated by two lines through their plastid circle. Dashed lines in the tree represent regions of uncertainty with respect to the phylogenetic placement of the corresponding lineages.

by two membranes. But although the phylogenetic information retained in plastid genomes clearly points to cyanobacteria as the source of the organelle, identifying the closest relative of plastids among present-day cyanobacteria has proven challenging. In attempting to address this question, researchers have analyzed the genes retained in plastid genomes as well as cyanobacterium-derived genes that migrated to the nuclear genome during and after the establishment of the organelle (see below). The picture has changed with the use of new analytical approaches and the discovery of new cyanobacterial lineages living in diverse environments.

A phylogenomic analysis performed by Deusch et al. (2008) suggested that plastids evolved from marine, filamentous nitrogen-fixing cyanobacteria related to "section IV" cyanobacteria (e.g., *Nostoc* and *Anabaena* spp.). In contrast, analysis of the genome of the biofilm-forming species *Gloeomargarita lithophora* suggested that plastids evolved from an early branching, *Gloeomargarita*-like cyanobacterium in a freshwater–terrestrial environment (Ponce-Toledo et al. 2017). Further evidence for a specific evolutionary connection between plastids and *Gloeomargarita* has recently come from Moore et al. (2019), who carried out an expanded phylogenetic analysis of plastid and cyanobacterial ribosomal proteins (including data from 20 newly sequenced cyanobacterial genomes). This is an active area of research, one in which the use of different data sets and methodologies can yield conflicting results. Interested readers are encouraged to refer to Moore et al. (2019), Ponce-Toledo et al. (2017), Shih et al. (2013), and references therein for diverse perspectives on the evolution of plastids relative to putatively deep-branching cyanobacteria such as *Gloeomargarita*, *Gloeobacter*, and *Pseudanabaena* sp.

Given the immense timescales involved, it is perhaps not surprising that we cannot (yet) determine whether plastids occupy a derived or deep position in the tree of extant cyanobacteria, or indeed the environmental conditions in which eukaryotic photosynthesis first evolved. In addition, although primary plastids are widely believed to have evolved only once, it should be noted that alternative scenarios of primary plastid origin have been proposed that involve independent endosymbiotic events in two or more archaeplastidal lineages, followed by convergent reductive evolution of plastid genome content (e.g., Stiller et al. 2003; Larkum et al. 2007). It has also been suggested that a *Chlamydia*-like pathogen was a critical third partner during the primary endosymbiotic origin of plastids, contributing genes to the eukaryotic host and impacting its glycogen metabolism (see Cenci et al. 2017 and references therein for an overview of the so-called ménage-à-trois hypothesis). As discussed below, a much more recent example of photosynthetic organelle evolution exists in an ameba called *Paulinella* (Marin et al. 2005; Nowack et al. 2008). Elucidation of the biology of *Paulinella* and its cyanobacterium-derived organelles has reinvigorated the debate over how endosymbionts and organelles are defined and the processes that underly the transition from one to the other.

The Endosymbiont-to-Organelle Transition

A critical initial step in the establishment of an obligate endosymbiotic relationship is metabolic crosstalk between host and endosymbiont. Genetic and cell biological integration, if they happen at all, come later, and serve to further link the metabolisms of the partner cells. In the fullness of time, genes that are redundant or no longer necessary for endosymbiont function are lost and genes are transferred from the endosymbiont to the host in a process referred to as endosymbiotic gene transfer (EGT) (Timmis et al. 2004). In order for the endosymbiont to utilize host-encoded proteins, the protein products must be targeted to and imported into the endosymbiont/organelle. It is the presence of a dedicated protein import system that is considered to be the defining feature of an endosymbiotically derived organelle (Cavalier-Smith and Lee 1985). In the case of primary plastids, nucleus-encoded, organelle-targeted proteins are flagged by the presence of amino (N)-terminal transit peptides that interact with membrane-anchored translocons (Gould et al. 2008). In more complex eukaryote–eukaryote endosymbioses (see below), additional targeting information comes in the form of a N-terminal signal peptide that helps guide the protein products across the additional membranes surrounding the organelle.

Nuclear genome sequences provide ample evidence for the important role of EGT in the early evolution of plastids. Less than 5% of the genes thought to have been present in the cyanobacterial progenitor of the plastid typically remain in the organellar genome, and hundreds of cyanobacterial genes can be found in the nuclear genomes of extant algae and plants (Dagan et al. 2013, Qiu, Yoon, et al. 2013). But are *all* of the proteins that make up the plastid proteome of cyanobacterial ancestry? No. The majority of nucleus-encoded, plastid-targeted proteins appear to be host derived; noncyanobacterial bacterial genes comprise 7–15% of plastid-targeted proteins in diverse lineages such as the model land plant *Arabidopsis*, the green alga *Chlamydomonas*, and the glaucophyte alga *Cyanophora* (Qiu, Price, et al. 2013). A fraction of this 7–15% are genes of α-proteobacterial ancestry and are likely repurposed EGTs from the mitochondrion. Others, however, appear to be lateral gene transfers (LGTs) into the nuclear genome of the eukaryotic host or LGTs from diverse bacteria into the genome of the cyanobacterial ancestor of plastids prior to endosymbiosis. There is presently little in the way of clarity on this point, as the inferred relative footprints of EGT and LGT vary from lineage to lineage and with differences in comparative genomic methodologies (e.g., Moustafa et al. 2009; Deschamps and Moreira 2012; Morozov and Galachyants 2019). We will revisit this issue below.

Although we can make educated guesses about how primary plastids transitioned from endosymbionts to organelles based on the properties of modern-day plastids and free-living cyanobacteria, we can also study more recently evolved photosynthetic lineages and their heterotrophic relatives to gain further insight into what might have occurred at the dawn of plastid evolution. This includes *Paulinella* and its primary plastid-like "chromatophore," examples of "recent" plastid replacements in dinoflagellate algae (some of which are temporary), and the newly discovered Rhodelphidia, a heterotrophic, phagotrophic protist lineage specifically related to red algae (Gawryluk et al. 2019). Comparative genomics is helping us sharpen the picture of how plastids have evolved throughout the eukaryotic tree of life.

The Chromatophores of *Paulinella*

Beyond the primary endosymbiotic event that led to plastid establishment in Archaeplastida, photosynthetic organelles are known to have evolved directly from cyanobacteria on at least one other occasion. First discovered more than 100 years ago (Lauterborn 1895), the freshwater thecate amoeba *Paulinella chromatophora* has green sausage-shaped chromatophores in its cytoplasm the origin of which has only recently become clear (Marin et al. 2005; Nowack et al. 2008). Initial sequence characterization of the chromatophore rDNA operon showed them to be of α-cyanobacterial ancestry (specifically the *Synechococcus/Prochloroccus* clade), in contrast to the β-cyanobacteria that canonical plastids are related to Marin et al. (2005). Recent molecular clock analyses estimate that the *Paulinella* chromatophore evolved a mere 90–140 Ma (Delaye et al. 2016). The 1.02-Mb chromatophore genome of *P. chromatophora* was nevertheless found to be substantially reduced compared to the ∼3-Mb genome of its closest free-living α-cyanobacterial relative, *Synechococcus* WH570, and has only a quarter of the protein coding capacity (Nowack et al. 2008). This includes loss of genes involved in essential amino acid and cofactor biosynthesis pathways, which suggests a significant degree of host–chromatophore dependency. The presence of many pseudogenes in the chromatophore genome suggests that genome reduction is ongoing (Nowack et al. 2008).

The sequencing of additional chromatophore genomes has provided further insight into the endosymbiont-to-organelle transition in this understudied lineage. Lhee et al. (2019) showed conservation of gene content and genome structure between the chromatophore genomes of a variety of *Paulinella* species and strains, suggesting that most of the genome reduction (∼65% of the protein coding genes) occurred in the common ancestor of studied *Paulinella* species before they diverged from one another. The differential loss of some genes between the studied chromatophore genomes indicates that genome reduction is still underway, but overall the chromatophore appears to be in a stabilizing stage (i.e., the rate of EGT has significantly decreased).

A draft nuclear genome for *P. chromatophora* was published in 2016 (Nowack et al. 2016) and combined with chromatophore proteomic data (Singer et al. 2017) has afforded a more detailed assessment of the role of EGT in this system. Complementary nuclear and chromatophore gene inventories were identified, consistent with the notion of a high degree of metabolic integration between host and organelle. Surprisingly, phylogenetic analyses revealed that only 17 of 433 nucleus-encoded, chromatophore-targeted proteins identified by proteomics appear to come from α-cyanobacteria. Twenty-six proteins are apparent LGTs from other bacteria and the remaining proteins are of unknown or eukaryotic (i.e., host) origin (Singer et al. 2017). The take-home message is that during the establishment of the chromatophore as a permanent intracellular entity, gaps in critical chromatophore biochemical pathways were filled by genes and proteins from sources other than the cyanobacterial progenitor of the organelle. As we shall see, such evolutionary mosaicism is apparent when one considers the nuclear genomes and plastid proteomes of other algae as well.

The Complexity of Complex Plastids

As important as it was, the primary endosymbiotic origin of plastids accounts for only a fraction of the diversity of plant and algal life on Earth. Multiple higher-order endosymbiotic events—mergers between two eukaryotic cells—have taken place. In some photosynthetic lineages, the details are reasonably well understood, whereas in others they are completely uncertain. So-called "secondary" plastids of green algal origin arose on two separate occasions, one in the chlorarachniophytes and the other in the euglenids. We can infer that these correspond to two separate secondary endosymbioses because the organisms belong to different eukaryotic supergroups, Rhizaria in the case of chlorarachniophytes and Discoba in the case of euglenids (fig. 1). Efforts to pinpoint the precise green algal source for each of these plastids have brought us closer to an answer as more genomic data have become available. For example, recent targeted sequencing of algal lineages closely related to potential plastid donors (based on previous sequencing and phylogenetic analysis) and taxon-rich phylogenetic analysis of their plastid genomes pinned down their sources to be a precursor of siphonous green algae (Bryopsidales) in the case of chlorarachniophytes, and a prasinophyte from the order pyramimonadales in euglenids (Jackson et al. 2018). Although progress has been made, Jackson et al. (2018) suggest that increased sampling of plastid genomes of secondary green plastid relatives may not resolve the question much further due to limitations in phylogenetic signal. Improved resolution may be obtained by increased nuclear genome sequencing of these taxa and investigation of endosymbiont-derived nuclear genes.

However, we must also acknowledge that in the case of chlorarachniophytes and/or euglenids, the plastid donor could have been an unknown lineage or a distant relative of extant taxa that has gone extinct.

Complex plastids of red algal origin (including those acquired from secondary, tertiary, and possibly even higher-order endosymbiotic events) are found across an even wider diversity of eukaryotes, including cryptophytes and haptophytes, as well as some alveolates and stramenopiles (including diatoms and brown/golden algae) (fig. 1). Despite decades of study, many uncertainties still surround the origin(s) and evolution of red algal complex plastids—how many times they were established and how often they were horizontally spread (and between whom) is unclear.

Why is the evolutionary history of such plastids so difficult to discern? Part of the problem lies in incongruences between phylogenies of plastid and nuclear genes. Plastid multigene phylogenies typically place complex red plastid-containing algae in a monophyletic clade that branches from within the red algae, albeit with internal tree topologies that are sensitive to phylogenetic method and taxon sampling (a trend that has nevertheless emerged is plastid trees that unite haptophytes and cryptophytes to the exclusion of all other complex algae; see, e.g., Janouškovec et al. 2010; Ševčíková et al. 2015; Kim et al. 2017 and references therein for discussion). Such phylogenies are consistent with the idea of a single secondary endosymbiotic origin of complex red alga-derived plastids, although it is important to note that, in isolation, they are also consistent with multiple independent endosymbioses involving closely related red algal endosymbionts. The precise nature of the algal donor is similarly unclear. At present, the data suggest that red alga-type complex plastids share more recent common ancestry with those of mesophilic red algae such as *Porphyra* and *Chondrus* than they do with the plastids of extremophiles such as *Cyanidioschyzon* and *Galdieria* (e.g., Ševčíková et al. 2015; Kim et al. 2017; Muñoz-Gómez et al. 2017).

Host nuclear gene phylogenies, however, generally do not resolve the various algae containing complex red plastids as a monophyletic group (e.g., Strassert et al. 2019); the organisms are found in several different eukaryotic supergroups, with intervening heterotrophic lineages in which there is no evidence of there ever having been a plastid (fig. 1). This discordance makes it unclear whether complex plastids of red algal origin were established once, vertically inherited, and differentially lost, or whether they were horizontally transferred from one lineage to another on multiple occasions.

One of the most prominent and controversial hypotheses of the past two decades of research in this area is the chromalveolate hypothesis (Cavalier-Smith 1999), which suggests that all red algal complex plastids are derived from a single secondary endosymbiosis with a red alga in the common ancestor of all taxa who possess them (i.e., some stramenopiles, some alveolates, haptophytes, and cryptophytes). Enthusiasm

for the chromalveolate hypothesis has diminished in recent years, in part due to the results of large-scale phylogenomic analyses. The hypothesis was initially founded on the premise that the number of inferred plastid establishments should be minimized due to the perceived difficulties associated with evolving an organelle, including the nucleus-to-nucleus transfer of hundreds to thousands of genes and the establishment of a functional protein import apparatus with each secondary endosymbiosis (Cavalier-Smith 1999). However, given our current understanding of the structure of the eukaryotic tree of life (e.g., Burki et al. 2016; Strassert et al. 2019), the chromalveolate hypothesis is unparsimonious in the sense that it requires extensive plastid loss in numerous, phylogenetically intervening, heterotrophic lineages (fig. 1; see below). Recent large-scale phylogenomic studies of nuclear genes place haptophytes and the heterotrophic centrohelids together (Haptista) as sister to the SAR assemblage (stramenopiles, alveolates, and rhizarians), whereas Cryptista (to which cryptophytes belong) branches completely separate in a highly supported relationship with Archaeplastida (e.g., Burki et al. 2016; Strassert et al. 2019). As Burki et al. (2016) pointed out, this branching pattern altogether rules out the chromalveolate hypothesis as red alga-derived secondary plastids would have had to originate before red algal plastids themselves even existed.

Although the genomes of red alga-derived complex plastids appear monophyletic (see above), the conflicting evolutionary histories of the plastid and nucleus have prompted alternative (and generally quite similar) hypotheses centered around scenarios involving a single secondary endosymbiosis followed by additional higher-order endosymbiotic events (e.g., Sanchez-Puerta and Delwiche 2008; Bodył et al. 2009; Petersen et al. 2014; Stiller et al. 2014; Ševčíková et al. 2015; Burki 2017; Bodył 2018). Stiller et al. (2014) proposed one such scenario, where the initial secondary endosymbiosis occurred in a heterotrophic ancestor of cryptophytes. The secondary plastid in cryptophytes is then suggested to have spread to an ancestor of photosynthetic stramenopiles via a tertiary endosymbiosis, and then into haptophytes by a quaternary endosymbiotic event with a stramenopile. This "cryptophyte first" model places the initial secondary event in the ancestor of cryptophytes on the basis of plastid gene phylogenies, statistical analyses of EGTs in ochrophytes, haptophytes, and cryptophytes, and the preservation of a relic of the primary red algal nucleus in cryptophytes (i.e., the nucleomorph; Stiller et al. 2014). Ševčíková et al. (2015) further expanded phylogenetic analysis to include the plastid genome of a chromerid (a representative of alveolates, which were not included in the analyses of Stiller et al. [2014]). In doing so, these authors proposed that the chromerid plastid, and perhaps other plastid-bearing alveolates, emerged from a tertiary or quaternary endosymbiotic event with an early ochrophyte,

with the exact type of endosymbiosis depending on the evolutionary origin of the ochrophyte plastid itself (see below for discussion).

As its name suggests, the "cryptophyte first" scenario places cryptophytes at the origin of red algal complex plastids, making these algae and the rest of Cryptista of particular interest for the study of how these secondary plastids evolved. Cryptista is generally thought to be an ancestrally nonphotosynthetic clade due to a lack of molecular evidence for a cryptic plastid or plastid-derived genes in the katablepharids (Burki et al. 2012), one of the early diverging plastid-lacking clades within this phylum. Until recently, it was not known if the closest heterotrophic lineage to the cryptophytes—the goniomonads—was ancestrally heterotrophic, or if it once had a red alga-derived plastid. In the absence of cytological evidence for the existence of a vestigial plastid in the goniomonad species *Goniomonas avonlea* (Kim and Archibald 2013), Cenci et al. (2018) searched for a genomic footprint of red algal endosymbiosis (i.e., EGTs) and past plastid ancestry. Consistent with a previous transcriptome-based survey of *Goniomonas pacifica* (Yabuki et al. 2014), the genomic survey of *G. avonlea* found no convincing evidence for the previous existence of a red alga-derived plastid in goniomonads, suggesting that the red algal plastid in cryptophytes was established after their divergence from goniomonads (Cenci et al. 2018). Intriguingly, Cenci et al. (2018) identified genes in *G. avonlea* suggesting that it grazes not just on bacteria but eukaryotes as well, including algae and specifically red algae (via the identification of an agarase gene encoding a protein with a signal peptide). It is thus not a stretch to imagine a scenario in which a red algal cell was phagocytosed by a *Goniomonas*-like ancestor but not digested, giving rise to the first photosynthetic cryptophytes.

Red–Green Mosaicism in Complex Algae

Although genome and transcriptome-based studies have provided some insight into the evolution of red algal complex plastids as a whole, numerous uncertainties remain. Difficulties associated with proving plastid loss (see below) and the extraction of ancient phylogenetic signal from molecular data, combined with complicating factors such as plastid replacements, all impact our ability to resolve the evolutionary trajectories of red alga-derived plastids. One of the most confounding factors has been the realization that the nuclear genomes of red complex plastid-bearing lineages harbor genes of both red and green algal origin. At present there is no consensus on how best to interpret these data.

A well-studied example of such red–green mosaicism is diatoms. Despite the fact that these ubiquitous algae harbor plastids of red algal ancestry, their nuclear genomes contain a substantial number of genes that appear to be of green algal origin. An early analysis of the relative contribution of red versus green algal genes to the diatom nuclear genome found that ~70% were green algal in nature (Moustafa et al. 2009)—a contribution deemed so substantial that it led the authors to propose the existence of a previous cryptic green algal secondary plastid in diatoms. However, a reanalysis of these diatom genomes using additional red algal genomic data and a stricter set of analytical criteria found that only ~13% of EGTs could be confidently traced from cyanobacteria to green algae to diatoms, with ~66% clearly traceable to red algae (Deschamps and Moreira 2012). A recent study by Morozov and Galachyants (2019) found the relative footprint of red and green algal EGTs in diatom nuclear genomes to be approximately equal, leading them to question the existence of a fully integrated green alga-derived plastid in a diatom ancestor. The authors suggest that the diatom "green" genes are more likely to be the legacy of transient endosymbioses involving at least two distinct green algal endosymbionts prior to fixation of the current red alga-derived plastid (Morozov and Galachyants 2019). This idea is consistent with the "shopping bag" model of plastid evolution put forth by Larkum et al. (2007) and Howe et al. (2008), which emphasizes genetic contributions from multiple endosymbionts over extended periods of time as playing a role in the establishment of a permanent photosynthetic organelle. In 2017, Dorrell et al. published an exhaustive analysis of the plastid proteomes of diatoms, pelagophytes and other stramenopiles, as well as diverse complex plastid-bearing algae. These authors concluded that "... the ancestral ochrophyte plastid proteome was an evolutionary chimera, with 25% of its phylogenetically tractable nucleus-encoded proteins deriving from green algae." They posit that the red algal-type plastid currently residing in extant ochrophytes is a "late" addition to stramenopiles, and that the ochrophyte "green" genes are a legacy of the presence of a green algal-type plastid in an ochrophyte ancestor (Dorrell et al. 2017).

The "red carpet" hypothesis of Ponce-Toledo et al. (2019) was recently put forth to explain the converse situation, that is, the presence of red genes in algae with secondary green algal plastids. These authors traced the evolutionary history of nuclear genes from cyanobacteria to red/green algae and on to the euglenids and chlorarachniophytes. They found that ~30% and 50% of these genes, respectively, were of apparent red algal origin rather than green, resulting in the existence of highly mosaic plastid metabolic pathways (Ponce-Toledo et al. 2018). It was suggested that the establishment of green alga-derived secondary plastids in euglenids and chlorarachniophytes was facilitated by the acquisition of genes from red algae prior to and/or during the early stages of secondary endosymbiosis, again akin to the shopping bag model of plastid origin.

The extent to which plastid replacements and/or shopping bag-type processes have given rise to red–green genome mosaicism in complex algae is presently unclear. But in attempting to make sense of these data, it is important to note that mosaic genomes can be seen in all such algae that have been

studied, not just diatoms and other photosynthetic stramenopiles (e.g., Dorrell et al. 2017), euglenids, and chlorarachniophytes but cryptophytes (Curtis et al. 2012) and haptophytes (Read et al. 2013; Dorrell et al. 2017) as well. We must recognize the possibility that nuclear genome mosaicism is due in part to the uptake of a plastid from a eukaryote whose genome was already mosaic due to LGT (not EGT). In addition, we must accept the fact that our phylogenetic reconstructions of ancient evolutionary events are undoubtedly impacted by methodological artifacts and incomplete taxonomic sampling.

Secondary Loss of Photosynthesis: How Widespread?

Given the obvious evolutionary advantages of photosynthesis, it is perhaps surprising that having been gained, the ability to extract energy from sunlight has been lost multiple times independently across a wide range of eukaryotic lineages—from the complex plastids of some stramenopiles (e.g., Beisser et al. 2017; Graupner et al. 2018; Dorrell et al. 2019), cryptophytes (Hoef-Emden 2005; Donaher et al. 2009), apicomplexans (discussed below), and euglenids (e.g., Marin et al. 2003; Záhonová et al. 2018), to the primary plastids of certain species of red algae, green algae, and land plants (fig. 1). Why do plastids usually persist in such organisms? And what happens to their genomes when photosynthesis is lost?

Comparative genomics has shown that nonphotosynthetic plastids and their genomes exhibit a range of characteristics depending on how much time has transpired since photosynthesis was "turned off." For example, the plastid genome of the recently evolved colorless diatom *Nitzschia* sp. has lost all photosystem genes and almost every gene related to photosynthesis; however, a large number of nuclear encoded proteins are targeted to the plastid, indicating that the organelle retains significant metabolic activity (Kamikawa et al. 2015, 2017). The sequence of a second plastid genome from a closely related *Nitzschia* species, as well as phylogenetic analysis of nuclear rDNA and mitochondrial genes from many new nonphotosynthetic *Nitzschia* species, suggests that there was a single loss of photosynthesis within *Nitzschia* spp.; all examined species form a monophyletic group nested within the rest of the photosynthetic *Nitzschia* spp. and diatoms (Onyshchenko et al. 2019). Nuclear genome sequence data for a nonphotosynthetic *Nitzschia* spp. and their closest mixotrophic relatives will hopefully elucidate how the evolutionary transition from phototrophy to heterotrophy occurred in these diatoms.

Apicomplexans such as the malaria parasite *Plasmodium* are an example of a much more ancient loss of photosynthesis. These organisms belong to the Alveolata and (with some exceptions discussed below) harbor a highly reduced remnant plastid called an apicoplast. While nonphotosynthetic, the apicoplast is home to various core metabolic pathways including heme biosynthesis, iron–sulfur cluster synthesis, isoprenoid synthesis, and fatty acid synthesis (Lim and McFadden 2010). Morphological similarities, combined with the discovery of a phylogenetic connection between apicoplasts, the photosynthetic plastids of chromerids, and the peridinin-pigmented plastids of dinoflagellates, strongly suggest that the common ancestor of Apicomplexa and dinoflagellates was photosynthetic and harbored a complex plastid of red algal origin (Moore et al. 2008; Janouškovec et al. 2010).

How many times did apicomplexan-like species evolve within alveolates? Applying single-cell genomics and transcriptomics to uncultivated species, Mathur et al. (2019) and Janouškovec et al. (2019) showed that apicomplexan-like parasites are polyphyletic and that heterotrophic parasites evolved from photosynthetic ancestors on multiple occasions. Kwong et al. (2019) recently described a novel coral-associated lineage called "corallicolids" with apicomplexan-like ultrastructural features and a nonphotosynthetic plastid genome that nevertheless retains ancestral genes for chlorophyll biosynthesis. Combined with inferences gleaned from heterotrophic colpodellids (Janouškovec et al. 2015) and the photosynthetic chromerids (*Chromera* and *Vitrella*; Woo et al. 2015), these studies show that there are in fact multiple paths to genome reduction in nonphotosynthetic plastids, resulting in organelles with overlapping but distinct gene sets and metabolic capacities. Salomaki and Kolisko (2019) describe this process as "endosymbiotic roulette," a game of chance in which the core metabolic pathways and genes that end up being kept in a given organelle are determined by the stochastic nature of EGT, the biochemical capacities of the host cell, and the ease with which key metabolites can be acquired from the environment.

Heterotrophic land plants are either parasites of other plant species or are fully mycoheterotrophic; in both cases, they either have highly reduced plastid genomes (e.g., mycoheterotrophic orchids [Schelkunov et al. 2015] and the plant endoparasite genus *Pilostyles* [Bellot and Renner 2016]) or lack one altogether (e.g., the parasitic flowering plant *Rafflesia* [Molina et al. 2014]). Analysis of plastid genome sequences from these nonphotosynthetic plant species has shown that they are drastically reduced in size and coding content, typically encoding only ribosomal components and a few other housekeeping genes. Those species that are endo-parasites of another photosynthetic host exhibit a greater degree of plastid gene loss. For example, only five or six genes remain in the plastid genome of *Pilostyles* spp. (Bellot and Renner 2016), whereas there is no evidence of a plastid genome at all in *Rafflesia* (Molina et al. 2014) even though a plastid structure remains.

A similar picture is seen within the green algae, where there appears to have been multiple losses of photosynthesis in unrelated lineages, both parasitic and free-living. For example, within the trebouxiophyte green algae, two nonphotosynthetic genera, *Prototheca* and *Helicosporidium*, are closely

related to the photosynthetic genera *Chlorella* and *Auxenochlorella*, respectively. Based on sequencing of the plastid and nuclear genomes of the parasitic *Helicosporidium* (de Koning and Keeling 2006; Pombert et al. 2014) and free-living *Prototheca* (Yan et al. 2015; Suzuki et al. 2018), it was suggested that there have been three independent losses of photosynthesis in this algal group, with convergent gene losses and retained plastid functions (Suzuki et al. 2018). The loss of photosynthesis is typically associated with smaller plastid genomes due to the loss of photosynthesis-related genes. The only known exception is the free-living heterotrophic green alga, *Polytoma uvella* (Figueroa-Martinez et al. 2017). Although the plastid genome of this organism is highly reduced (it has only 25 genes), it is nevertheless still the largest of the nonphotosynthetic plastids currently known and is actually larger than the plastid genome of its closest photosynthetic relative, *Chlamydomonas* (Figueroa-Martinez et al. 2017). This size discrepancy is due to expansions of short repeats in the *Polytoma uvella* genome. Photosynthesis has also been lost in another free-living green alga, but in this case the plastid genome is completely gone (Smith and Lee 2014). Morphological studies support the presence of a colorless plastid in *Polytomella* (Moore et al. 1970), and transcriptome and genome sequencing shows the existence of nuclear encoded, plastid-targeted proteins. Why *Polytomella* has lost its plastid genome when most secondarily heterotrophic lineages retain it is a mystery.

One of the most surprising recent revelations in the field of plastid evolution is the discovery of a group of nonphotosynthetic, predatory flagellates whose closest relatives are the red algae (fig. 1). Gawryluk et al. (2019) sequenced the genomes and transcriptomes of a freshwater and a marine species (*Rhodelphis limneticus* and *Rhodelphis marinus*, respectively) and erected a new phylum: Rhodelphidia. Although no plastid genome was identified in *Rhodelphis* spp., and no plastid could be observed under the microscope, many genes for plastid-targeted proteins were found in the nucleus, some of which are clearly EGTs from the plastid (they are still present in the plastid genomes of typical red algae). Bioinformatic analyses suggest that the relic plastid of *Rhodelphis* spp. is the site of some of the same metabolic processes retained in other nonphotosynthetic organelles, mainly heme biosynthesis and (probably) iron–sulfur cluster biogenesis (Gawryluk et al. 2019). Importantly, combined with its phylogenetic position on the eukaryotic tree, the phagotrophic nature of *Rhodelphis* has important implications for how we envision the biology of the earliest photosynthetic eukaryotes. The evidence suggests that early in archaeplastid evolution, photosynthesis was supplemented by phagotrophy, with mixotrophy persisting until the plastid became fully capable of supporting a photoautotrophic lifestyle (Colp and Archibald 2019). This idea is supported by the fact that certain green algae are capable of ingesting bacterial prey by phagocytosis (Maruyama and Kim 2013).

Plastid Loss

Although the loss of photosynthetic capacity is not uncommon among plants and algae, outright plastid loss appears to be extremely rare. The underlying reasons relate to the essential biochemical functions so often localized to the organelle. In order for a plastid to be lost, alternative ways to obtain essential plastid-derived metabolites are needed; biochemical pathways encoded by, and/or taking place in, the plastid need to be re-engineered or bypassed entirely. Only three unambiguous examples of complete plastid loss have been documented so far, all in parasitic species (two apicomplexans—*Cryptosporidium* and certain gregarines—and the dinoflagellate genus *Hematodinium*). Consideration of these specific cases has proven insightful.

In the case of the human parasite *Cryptosporidium parvum*, genomic analysis suggests that fatty acid synthesis occurs not in the plastid but in the cytosol using an atypical biochemical pathway, and that various metabolites are scavenged from its host; the "core" biochemical pathways typically seen in nonphotosynthetic plastids became dispensable (Zhu et al. 2000). Transcriptome and genome sequence data from the dinoflagellate *Hematodinium* failed to find evidence of a plastid and showed that it has retained the ancestral host pathway for cytosolic fatty acid and tetrapyrrole (related to heme) synthesis (Gornik et al. 2015). *Hematodinium* has also relocated lysine biosynthesis to the cytosol. The final hurdle to plastid loss—isoprenoid synthesis—appears to have been overcome by the scavenging of isoprenoid synthesis intermediates from its host, as many biosynthesis genes appeared to be missing (Gornik et al. 2015).

Recent genomic data obtained for gregarines, intestinal parasites of many invertebrates, confirmed speculation that no plastid or plastid genome exists in terrestrial species (Toso and Omoto 2007; Mathur et al. 2019). Single-cell transcriptomics has shown that some marine gregarines retain a relic plastid that only appears to retain the fatty acid biosynthesis pathway, unlike the apicoplast of related apicomplexans, which retains three additional biochemical pathways, including isoprenoid biosynthesis. These pathways are thought to underly plastid retention (see Janouškovec et al. 2015, 2019; Mathur et al. 2019 for discussion). Other marine gregarine species lack evidence for a relic plastid and are thus suggested to have lost the plastid completely, whereas related blastogregarines and archigregarines have nonphotosynthetic plastids with genomes that are reduced to a level more similar to that of apicomplexans (Janouškovec et al. 2019). Although all known gregarines are parasites, only those that appear to have sufficient alternatives to "core" plastid biosynthesis pathways are able to tolerate complete loss of the organelle.

Kleptomania

Thus far, we have focused on the origins and fates of permanently integrated organelles. But there is also much to learn

from the study of more transient associations between organisms in nature, the phenomenon of kleptoplasty—"plastid stealing"—being one such an example. Kleptoplasty is surprisingly common, having been observed in a variety of protist lineages (particularly foraminiferans, dinoflagellates, and ciliates) and even some animals (sea slugs and some flat worms). These stolen plastids come from varying sources and exhibit a wide range of retention times. In sea slugs, algal plastids are harvested, sequestered, and retained for periods of time ranging from days to weeks and months before being lost (Händeler et al. 2009). Even though only the plastid remains from its algal prey, the kleptoplasts remain photosynthetically active. Transcriptome and genome sequencing have failed to provide evidence of gene transfer from algae to the sacoglossan sea slug nuclear genome, despite suggestions that this had occurred (see Bhattacharya et al. 2013; Rauch et al. 2015 and references therein). In a recently described case of kleptoplasty in two flatworm species, sequestered plastids are stolen from different diatom species and retained for a couple of weeks (Steenkiste et al. 2019). In this instance, transcriptome sequencing suggests that kleptoplast genes are actively expressed, whereas algal nuclear genes are not, supporting the observation that only the plastid is retained and is functional (Steenkiste et al. 2019).

Kleptoplastic tendencies are frequently observed in single-celled eukaryotes and are especially common among dinoflagellates. Approximately half of the described dinoflagellate species maintain the peridinin-pigmented plastid thought to have been present in their common ancestor (Yoon et al. 2002). Many dinoflagellates harbor kleptoplastids of various origins, whereas others have permanently replacing their ancestral plastid (see below). In addition as discussed above, still others completely lack photosynthetic abilities and appear to no longer have a plastid (e.g., *Hematodinium*; Gornik et al. 2015).

Haptophyte-derived kleptoplastids are found in *Phalacroma mitra* (Koike et al. 2005) and the Antarctic Ross Sea (ARS) dinoflagellate (Sellers et al. 2014; Hehenberger et al. 2019). The ARS dinoflagellate is known not only to retain a relic version of its ancestral peridinin plastid but also to be obligately kleptoplastidic and reliant on its haptophyte prey for photosynthesis, which it can maintain for at least 30 months (Gast et al. 2007). Transcriptomics of the ARS dinoflagellate by Hehenberger et al. (2019) showed that plastid functions are split between the relic peridinin plastid and the kleptoplast, with photosynthesis related functions occurring there. Bioinformatic analyses suggest that many nucleus-encoded proteins are targeted to the kleptoplast, most of which do not appear to be derived from the plastid donor itself. Notably, the ARS dinoflagellate is closely related to other dinoflagellate lineages, specifically *Karenia* and *Karlodinium*, who harbor fully integrated, permanent haptophyte-derived plastids. Comparative genomics between the species with permanent haptophyte plastids and the haptophyte-derived

kleptoplasts of the ARS dinoflagellate showed a number of shared gene transfers whose protein products are targeted to the plastid/kleptoplast, suggesting that, in some cases at least, a certain level of genetic integration can precede permanent organelle integration (Hehenberger et al. 2019).

Other dinoflagellates have cryptophyte- or diatom-derived kleptoplasts. In one of the most intricate cases, *Dinophysis* obtains its kleptoplast by preying on a ciliate (*Mesodinium rubrum*) which itself harbors a kleptoplast of cryptophyte origin (either *Geminigera cryophila* [Johnson et al. 2006] or *Teleaulax amphioxeia* [Nishitani et al. 2010]). In *Mesodinium rubrum*, the kleptoplast is transcriptionally active and has been observed to be retained for up to a month with continual replacement (Johnson et al. 2007). Transcriptomics and phylogenetic analysis of plastid-related genes in *Dinophysis fortii* revealed that most such genes are of apparent peridinin plastid origin and thus represent EGTs from the original dinoflagellate plastid (Hongo et al. 2019). Intriguingly, the rest of the plastid-associated genes were not only of cryptophyte kleptoplast origin but also of haptophyte origin—including genes related to those of dinoflagellates with permanent tertiary haptophyte plastids (i.e., fucoxanthin dinoflagellates). This suggests that the ancestors of extant *Dinophysis* engaged in haptophyte kleptoplasty at some point during their evolutionary history (Hongo et al. 2019).

Finally, dinoflagellates are known to harbor plastids derived from diatoms. These so-called "dinotoms" are retained permanently in the cytosol and the evidence suggests that the host cell exhibits some level of control over their division (e.g., Hehenberger et al. 2016; Yamada et al. 2017). To date, there is only one known example of a kleptoplasty-derived dinotom. In this case, the dinoflagellate *Durinskia capensis* appears to retain its diatom only temporarily, for ~2 months, and does not appear to control its cell division (Yamada et al. 2019). Further genomic and transcriptomic investigations of the myriad ways in which dinoflagellates acquire and recycle their plastids will no doubt continue to provide insight into the molecular, biochemical, and cell biological factors underlying the endosymbiont-to-organelle transition.

Out with the Old, In with the New

As mentioned above, only about half of the known species of dinoflagellates obviously retain their original peridinin-pigmented plastid, although it seems increasingly likely that all free-living, nonparasitic dinoflagellates harbor a plastid of some kind, regardless of whether or not it is photosynthetic (see Janouškovec et al. 2017 for recent discussion). Some dinoflagellates exhibit kleptoplasty, whereas others have permanently replaced their ancestral plastid with something new. For example, dinoflagellates with diatom-derived plastids typically retain the diatom in their cytoplasm permanently at an intermediate level of integration (Yamada et al. 2017). Transcriptomics has shown that although the host

dinoflagellate is able to retain the diatom by controlling its division, genome reduction appears not to have occurred (at least in *Durinskia baltica* and *Glenodinium foliaceum*; Hehenberger et al. 2016). The apparent lack of host–endosymbiont genetic integration (along with the general retention of cellular features such as the diatom's ER, cytosol, mitochondrion, and nucleus) suggests that the association was established relatively recently, and while obligate and permanent, it is still early on in the transition to fully fledged organelle (Hehenberger et al. 2016). Phylogenetic analysis of the rDNA of a variety of "dinotoms" has shown that there are at least 11 different diatom species found throughout these closely related dinoflagellates, providing evidence that the dinotoms have been acquired and replaced on multiple occasions (Yamada et al. 2017).

There are two instances of permanent plastid replacement in dinoflagellates outside the dinotoms discussed above: 1) a haptophyte-derived plastid in the Kareniaceae (e.g., *Karenia*, *Karlodinium*, and *Takayama*) and 2) the green alga-derived plastid seen in *Lepidodinium*. On the basis of plastid and nuclear rDNA phylogenies, the *Karenia* and *Karlodinium* tertiary plastids appear to be derived from two different haptophytes (Tengs et al. 2000). Sequencing the genome of the haptophyte-derived plastid in *Karlodinium veneficum* identified substantial genome rearrangement and gene loss compared with the haptophyte plastid itself (convergent with but not quite to the level of reduction seen in a peridinin plastid) (Gabrielsen et al. 2011), whereas transcriptome-based studies of both *Karenia brevis* and *Karlodinium veneficum* have identified 90 haptophyte-to-dinoflagellate EGTs (Burki et al. 2014). Genes for plastid-targeted proteins from a variety of sources were also found (Nosenko et al. 2006; Patron et al. 2006), providing evidence for host–endosymbiont integration.

In the case of *Lepidodinium*, the presence of a green alga-type plastid has been recognized for quite some time (Watanabe et al. 1990). It was not, however, until recently (with increased plastid genome sampling and phylogenetic analysis) that the exact source of the serial secondary plastid was identified as a pedinophyte green alga (Kamikawa et al. 2015; Jackson et al. 2018). Sarai et al. (2020) recently provided molecular, biochemical, and microscopic evidence for the existence of a pedinophyte-derived plastid and nucleomorph-like organelle in two evolutionarily distinct dinoflagellate strains, presently dubbed MGD and TGD. Taken as a whole, host and endosymbiont phylogenies suggest that their plastids evolved independent of the *Lepidodinium* plastid; the MGD and TGD plastids (and associated "nucleomorphs") may themselves in fact represent separate acquisitions (Sarai et al. 2020). Endosymbiont-to-host gene transfer was also documented, although the extent of genetic integration between the MGD and TGD hosts and endosymbionts is still unclear.

Conclusion

Three photosynthetic lineages—green algae plus land plants, red algae, and glaucophyte algae—harbor plastids that appear to stem directly from a primary endosymbiotic event with a cyanobacterium a billion-plus years ago (Parfrey et al. 2011; Shih and Matzke 2013). A wealth of data reveals that the plastids of green and red algae subsequently spread far and wide across the tree of eukaryotes by higher-order endosymbiotic mergers between eukaryotic hosts and endosymbionts. Here, we have focused on insights gleaned from the perspective of genomics, but it is important to note that advances in our understanding of the cell biology, biochemistry, and metabolism of diverse algae and plants have contributed greatly to the broad picture of plastid evolution (see, e.g., Gould et al. 2015; Kim and Archibald 2009 and references therein). The challenge is combining all of these data into a single, coherent picture of the birth and spread of plastids. Despite 20-plus years of molecular phylogenetic and genomic investigation, we are still in the dark about how many eukaryote–eukaryote endosymbioses have occurred and, in many cases, who the partner cells even were (although not discussed here, interested readers should refer to Kim and Maruyama [2014] for a provocative discussion of the possibility that the plastid found in green algae and land plants is of secondary endosymbiotic origin).

With phylogenomics, the deep structure of the eukaryotic tree of life has become ever more resolved (e.g., Strassert et al. 2019; Burki et al. 2020; fig. 1), thereby providing a framework for mapping plastid gains and losses across the tree. However, the same genomic data that have enabled construction of taxonomically rich phylogenomic trees also reveal that the nuclear genomes of complex algae are mosaics of genes of both red and green algal ancestry. Putative LGTs from bacteria are also increasingly described in the nuclear and plastid genomes of phototrophs representing the full breadth of algal diversity (e.g., Khan et al. 2007; Dorrell et al. 2017; Ševčíková et al. 2019; Novák Vanclová et al. 2020). Why (and how) this is so is not yet clear. As we have seen, lineages such as the dinoflagellates and apicomplexans provide a window into the dynamics of plastid gain, loss, and replacement over recent evolutionary timescales. Moving forward, these data will help us to generate and test hypotheses with which to elucidate much older events in plastid evolution. Which complex algal lineage was the initial recipient of the primordial red algal secondary plastid? How many secondary, tertiary, and (possibly) quaternary endosymbioses gave rise to the full breadth of algal biodiversity, and who were the plastid donors and recipients? Answers to these questions will require even greater efforts to improve taxon sampling among both heterotrophic and photosynthetic lineages, as well as creative solutions to hard bioinformatic problems.

Acknowledgments

Literature Cited

Beisser D, et al. 2017. Comprehensive transcriptome analysis provides new insights into nutritional strategies and phylogenetic relationships of chrysophytes. PeerJ 5:e2832.

Bellot S, Renner SS. 2016. The plastomes of two species in the endoparasite genus *Pilostyles* (Apodanthaceae) each retain just five or six possibly functional genes. Genome Biol Evol. 8(1):189–201.

Bhattacharya D, Pelletreau KN, Price DC, Sarver KE, Rumpho ME. 2013. Genome analysis of *Elysia chlorotica* egg DNA provides no evidence for horizontal gene transfer into the germ line of this kleptoplastic mollusc. Mol Biol Evol. 30(8):1843–1852.

Bodył A. 2018. Did some red alga-derived plastids evolve via kleptoplastidy? A hypothesis. Biol Rev. 93(1):201–222.

Bodył A, Stiller JW, Mackiewicz P. 2009. Chromalveolate plastids: direct descent or multiple endosymbioses? Trends Ecol Evol. 24(3):119–121.

Burki F. 2017. The convoluted evolution of eukaryotes with complex plastids. Adv Bot Res. 84:1–30.

Burki F, Okamoto N, Pombert JF, Keeling PJ. 2012. The evolutionary history of haptophytes and cryptophytes: phylogenomic evidence for separate origins. Proc R Soc B 279(1736):2246–2254.

Burki F, Roger AJ, Brown MW, Simpson A. 2020. The new tree of eukaryotes. Trends Ecol Evol. 35(1):43–55.

Burki F, et al. 2014. Endosymbiotic gene transfer in tertiary plastid-containing dinoflagellates. Eukaryot Cell 13(2):246–255.

Burki F, et al. 2016. Untangling the early diversification of eukaryotes: a phylogenomic study of the evolutionary origins of Centrohelida, Haptophyta and Cryptista. Proc R Soc B 283(1823):20152802.

Cavalier-Smith T. 1999. Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. J Eukaryotic Microbiol. 46(4):347–366.

Cavalier-Smith T, Lee JJ. 1985. Protozoa as hosts for endosymbioses and the conversion of symbionts into organelles. J Protozool. 32(3):376–379.

Cenci U, et al. 2017. Biotic host–pathogen interactions as major drivers of plastid endosymbiosis. Trends Plant Sci. 22(4):316–328.

Cenci U, et al. 2018. Nuclear genome sequence of the plastid-lacking cryptomonad *Goniomonas avonlea* provides insights into the evolution of secondary plastids. BMC Biol. 16(1):137.

Colp MJ, Archibald JM. 2019. Evolution: new protist predators under the sun. Curr Biol. 29(19):R936–R938.

Curtis BA, et al. 2012. Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs. Nature 492(7427):59–65.

Dagan T, et al. 2013. Genomes of stigonematalean cyanobacteria (subsection V) and the evolution of oxygenic photosynthesis from prokaryotes to plastids. Genome Biol Evol. 5(1):31–44.

de Koning AP, Keeling PJ. 2006. The complete plastid genome sequence of the parasitic green alga *Helicosporidium* sp. is highly reduced and structured. BMC Biol. 4(1):12.

Delaye L, Valadez-Cano C, Pérez-Zamorano B. 2016. How really ancient is *Paulinella chromatophora*? PLoS Curr Tree Life 8:ecurrents.tol.e68a099364bb1a1e129a17b4e06b0c6b.

Deschamps P, Moreira D. 2012. Reevaluating the green contribution to diatom genomes. Genome Biol Evol. 4(7):683–688.

Deusch O, et al. 2008. Genes of cyanobacterial origin in plant nuclear genomes point to a heterocyst-forming plastid ancestor. Mol Biol Evol. 25(4):748–761.

Donaher N, et al. 2009. The complete plastid genome sequence of the secondarily nonphotosynthetic alga *Cryptomonas paramecium*: reduction, compaction, and accelerated evolutionary rate. Genome Biol Evol. 1:439–448.

Dorrell RG, et al. 2017. Chimeric origins of ochrophytes and haptophytes revealed through an ancient plastid proteome. eLife 6:e23717.

Dorrell RG, et al. 2019. Principles of plastid reductive evolution illuminated by nonphotosynthetic chrysophytes. Proc Natl Acad Sci U S A. 116(14):6914–6923.

Figueroa-Martinez F, Nedelcu AM, Smith DR, Reyes-Prieto A. 2017. The plastid genome of *Polytoma uvella* is the largest known among colorless algae and plants and reflects contrasting evolutionary paths to nonphotosynthetic lifestyles. Plant Physiol. 173(2):932–943.

Gabrielsen TM, et al. 2011. Genome evolution of a tertiary dinoflagellate plastid. PLoS One 6(4):e19132.

Gast RJ, Moran DM, Dennett MR, Caron DA. 2007. Kleptoplasty in an Antarctic dinoflagellate: caught in evolutionary transition? Environ Microbiol. 9(1):39–45.

Gawryluk RM, et al. 2019. Non-photosynthetic predators are sister to red algae. Nature 572(7768):240–243.

Gornik SG, et al. 2015. Endosymbiosis undone by stepwise elimination of the plastid in a parasitic dinoflagellate. Proc Natl Acad Sci U S A. 112:5767–5772.

Gould SB, Maier UG, Martin WF. 2015. Protein import and the origin of red complex plastids. Curr Biol. 25(12):R515–R521.

Gould SB, Waller RF, McFadden GI. 2008. Plastid evolution. Annu Rev Plant Biol. 59(1):491–517.

Graupner N, et al. 2018. Evolution of heterotrophy in chrysophytes as reflected by comparative transcriptomics. FEMS Microbiol Ecol. 94(4).

Händeler K, Grzymbowski YP, Krug PJ, Wägele H. 2009. Functional chloroplasts in metazoan cells—a unique evolutionary strategy in animal life. Front Zool. 6(1):28.

Hehenberger E, Burki F, Kolisko M, Keeling PJ. 2016. Functional relationship between a dinoflagellate host and its diatom endosymbiont. Mol Biol Evol. 33(9):2376–2390.

Hehenberger E, Gast RJ, Keeling PJ. 2019. A kleptoplastidic dinoflagellate and the tipping point between transient and fully integrated plastid endosymbiosis. Proc Natl Acad Sci U S A. 116(36):17934–17942.

Hoef-Emden K. 2005. Multiple independent losses of photosynthesis and differing evolutionary rates in the genus Cryptomonas (Cryptophyceae): combined phylogenetic analyses of DNA sequences of the nuclear and the nucleomorph ribosomal operons. J Mol Evol. 60(2):183–195.

Hongo Y, Yabuki A, Fujikura K, Nagai S. 2019. Genes functioned in kleptoplastids of *Dinophysis* are derived from haptophytes rather than from cryptophytes. Sci Rep. 9(1):9009.

Howe C, Barbrook A, Nisbet R, Lockhart P, Larkum A. 2008. The origin of plastids. Philos Trans R Soc B 363(1504):2675–2685.

Jackson C, Knoll AH, Chan C, Verbruggen H. 2018. Plastid phylogenomics with broad taxon sampling further elucidates the distinct evolutionary origins and timing of secondary green plastids. Sci Rep. 8(1):1523.

Janouskovec J, Horak A, Obornik M, Lukes J, Keeling PJ. 2010. A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. Proc Natl Acad Sci U S A. 107(24):10949–10954.

Janouškovec J, et al. 2015. Factors mediating plastid dependency and the origins of parasitism in apicomplexans and their close relatives. Proc Natl Acad Sci U S A. 112(33):10200–10207.

Janouškovec J, et al. 2017. Major transitions in dinoflagellate evolution unveiled by phylotranscriptomics. Proc Natl Acad Sci U S A. 114(2):E171–E180.

Janouškovec J, et al. 2019. Apicomplexan-like parasites are polyphyletic and widely but selectively dependent on cryptic plastid organelles. Elife 8:e49662.

Johnson MD, Oldach D, Delwiche CF, Stoecker DK. 2007. Retention of transcriptionally active cryptophyte nuclei by the ciliate Myrionecta rubra. Nature 445(7126):426–428.

Johnson MD, Tengs T, Oldach D, Stoecker DK. 2006. Sequestration, performance, and functional control of cryptophyte plastids in the Ciliate myrionecta Rubra (Ciliophora). J Phycol. 42(6):1235–1246.

Kamikawa R, et al. 2015. Plastid genome-based phylogeny pinpointed the origin of the green-colored plastid in the dinoflagellate Lepidodinium chlorophorum. Genome Biol Evol. 7(4):1133–1140.

Kamikawa R, et al. 2017. A non-photosynthetic diatom reveals early steps of reductive evolution in plastids. Mol Biol Evol. 34(9):2355–2366.

Khan H, et al. 2007. Plastid genome sequence of the cryptophyte alga Rhodomonas salina CCMP1319: lateral transfer of putative DNA replication machinery and a test of chromist plastid phylogeny. Mol Biol Evol. 24(8):1832–1842.

Kim E, Archibald JM. 2009. Diversity and evolution of plastids and their genomes. In: Aronsson H, Sandelius AS, editors. The chloroplast—interactions with the environment. Berlin: Springer-Verlag. p. 1–39.

Kim E, Archibald JM. 2013. Ultrastructure and molecular phylogeny of the cryptomonad Goniomonas avonlea sp. nov. Protist 164(2):160–182.

Kim E, Maruyama S. 2014. A contemplation on the secondary origin of green algal and plant plastids. Acta Soc Bot Pol. 83(4):331–336.

Kim J, et al. 2017. Evolutionary dynamics of cryptophyte plastid genomes. Genome Biol Evol. 9(7):1859–1872.

Koike K, et al. 2005. A novel type of kleptoplastidy in Dinophysis (Dinophyceae): presence of haptophyte-type plastid in Dinophysis mitra. Protist 156(2):225–237.

Kwong WK, et al. 2019. A widespread coral-infecting apicomplexan with chlorophyll biosynthesis genes. Nature 568(7750):103–107.

Larkum A, Lockhart PJ, Howe CJ. 2007. Shopping for plastids. Trends Plant Sci. 12(5):189–195.

Lauterborn R. 1895. Protozoenstudien II. Paulinella chromatophora nov gen., nov. spec., ein beschalter Rhizopode des Süßwassers mit blaugrünen chromatophorenartigen Einschlüssen. Z Wiss Zool. 59:537–544.

Lhee D, et al. 2019. Evolutionary dynamics of the chromatophore genome in three photosynthetic Paulinella species. Sci Rep. 9(1):2560.

Lim L, McFadden G. 2010. The evolution, metabolism and functions of the apicoplast. Philos Trans R Soc B 365(1541):749–763.

Marin B, Nowack EC, Melkonian M. 2005. A plastid in the making: evidence for a second primary endosymbiosis. Protist 156(4):425–432.

Marin B, Palm A, Klingberg M, Melkonian M. 2003. Phylogeny and taxonomic revision of plastid-containing euglenophytes based on SSU rDNA sequence comparisons and synapomorphic signatures in the SSU rRNA secondary structure. Protist 154(1):99–145.

Maruyama S, Kim E. 2013. A modern descendant of early green algal phagotrophs. Curr Biol. 23(12):1081–1084.

Mathur V, et al. 2019. Multiple independent origins of apicomplexan-like parasites. Curr Biol. 29(17):2936–2941.

Molina J, et al. 2014. Possible loss of the chloroplast genome in the parasitic flowering plant Rafflesia lagascae (Rafflesiaceae). Mol Biol Evol. 31(4):793–803.

Moore J, Cantor MH, Sheeler I, Kahn W. 1970. The ultrastructure of Polytomella agilis. J Protozool. 17(4):671–676.

Moore KR, et al. 2019. An expanded ribosomal phylogeny of cyanobacteria supports a deep placement of plastids. Front Microbiol. 10:1612.

Moore RB, et al. 2008. A photosynthetic alveolate closely related to apicomplexan parasites. Nature 451(7181):959–963.

Morozov AA, Galachyants YP. 2019. Diatom genes originating from red and green algae: implications for the secondary endosymbiosis models. Mar Genomics. 45:72–78.

Moustafa A, et al. 2009. Genomic footprints of a cryptic plastid endosymbiosis in diatoms. Science 324(5935):1724–1726.

Muñoz-Gómez SA, et al. 2017. The new red algal subphylum Proteorhodophytina comprises the largest and most divergent plastid genomes known. Curr Biol. 27(11):1677–1684.

Nishitani G, et al. 2010. High-level congruence of Myrionecta rubra prey and Dinophysis species plastid identities as revealed by genetic analyses of isolates from Japanese coastal waters. Appl Environ Microbiol. 76(9):2791–2798.

Nosenko T, et al. 2006. Chimeric plastid proteome in the Florida "red tide" dinoflagellate Karenia brevis. Mol Biol Evol. 23(11):2026–2038.

Novák Vanclová AMG, et al. 2020. Metabolic quirks and the colourful history of the Euglena gracilis secondary plastid. New Phytol. 225(4):1578–1592.

Nowack E, Melkonian M, Glöckner G. 2008. Chromatophore genome sequence of Paulinella sheds light on acquisition of photosynthesis by eukaryotes. Curr Biol. 18(6):410–418.

Nowack EC, et al. 2016. Gene transfers from diverse bacteria compensate for reductive genome evolution in the chromatophore of Paulinella chromatophora. Proc Natl Acad Sci U S A. 113(43):12214–12219.

Ohyama K, et al. 1986. Chloroplast gene organization deduced from complete sequence of liverwort Marchantia polymorpha chloroplast DNA. Nature 322(6079):572–574.

Onyshchenko A, Ruck EC, Nakov T, Alverson AJ. 2019. A single loss of photosynthesis in the diatom order Bacillariales (Bacillariophyta). Am J Bot. 106(4):560–572.

Parfrey LW, Lahr DJG, Knoll AH, Katz LA. 2011. Estimating the timing of early eukaryotic diversification with multigene molecular clocks. Proc Natl Acad Sci U S A. 108(33):13624–13629.

Patron NJ, Waller RF, Keeling PJ. 2006. A tertiary plastid uses genes from two endosymbionts. J Mol Biol. 357(5):1373–1382.

Petersen J, et al. 2014. Chromera velia, endosymbioses and the rhodoplex hypothesis—plastid evolution in cryptophytes, alveolates, stramenopiles, and haptophytes (CASH Lineages). Genome Biol Evol. 6(3):666–684.

Pombert JF, Blouin N, Lane C, Boucias D, Keeling PJ. 2014. A lack of parasitic reduction in the obligate parasitic green alga Helicosporidium. PLoS Genet. 10(5):e1004355.

Ponce-Toledo RI, López-García P, Moreira D. 2019. Horizontal and endosymbiotic gene transfer in early plastid evolution. New Phytol. 224(2):618–624.

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. Mol Biol Evol. 35(9):2198–2204.

Ponce-Toledo RI, et al. 2017. An early-branching freshwater cyanobacterium at the origin of plastids. Curr Biol. 27(3):386–391.

Qiu H, Price DC, et al. 2013. Assessing the bacterial contribution to the plastid proteome. Trends Plant Sci. 18(12):680–687.

Qiu H, Yoon H, Bhattacharya D. 2013. Algal endosymbionts as vectors of horizontal gene transfer in photosynthetic eukaryotes. Front Plant Sci. 4:366.

Rauch C, et al. 2015. Why it is time to look beyond algal genes in photosynthetic slugs. Genome Biol Evol. 7(9):2602–2607.

Read BA, et al. 2013. Pan genome of the phytoplankton Emiliania underpins its global distribution. Nature 499(7457):209–213.

Salomaki ED, Kolisko M. 2019. There is treasure everywhere: reductive plastid evolution in apicomplexa in light of their close relatives. Biomolecules 9(8):378.

Sanchez-Puerta VM, Delwiche CF. 2008. A hypothesis for plastid evolution in chromalveolates. J Phycol. 44(5):1097–1107.

Sarai C, et al. 2020. Dinoflagellates with relic endosymbiont nuclei as models for elucidating organellogenesis. Proc Natl Acad Sci U S A. 117(10):5364–5375.

Schelkunov MI, et al. 2015. Exploring the limits for reduction of plastid genomes: a case study of the mycoheterotrophic orchids *Epipogium aphyllum* and *Epipogium roseum*. Genome Biol Evol. 7(4):1179–1191.

Sellers GC, Gast RJ, Sanders RW. 2014. Selective feeding and foreign plastid retention in an Antarctic dinoflagellate. J Phycol. 50(6):1081–1088.

Ševčíková T, et al. 2015. Updating algal evolutionary relationships through plastid genome sequencing: did alveolate plastids emerge through endosymbiosis of an ochrophyte? Sci Rep. 5(1):10134.

Ševčíková T, et al. 2019. Plastid genomes and proteins illuminate the evolution of eustigmatophyte algae and their bacterial endosymbionts. Genome Biol Evol. 11(2):362–379.

Shih PM, Matzke NJ. 2013. Primary endosymbiosis events date to the later Proterozoic with cross-calibrated phylogenetic dating of duplicated ATPase proteins. Proc Natl Acad Sci U S A. 110(30): 12355–12360.

Shih PM, et al. 2013. Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. Proc Natl Acad Sci U S A. 110(3):1053–1058.

Shinozaki K, et al. 1986. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J. 5(9):2043–2049.

Singer A, et al. 2017. Massive protein import into the early-evolutionary-stage photosynthetic organelle of the amoeba *Paulinella chromatophora*. Curr Biol. 27(18):2763–2773.e5.

Smith D, Lee RW. 2014. A plastid without a genome: evidence from the nonphotosynthetic green algal genus *Polytomella*. Plant Physiol. 164(4):1812–1819.

Steenkiste NW, et al. 2019. A new case of kleptoplasty in animals: marine flatworms steal functional plastids from diatoms. Sci Adv. 5:eaaw4337.

Stiller JW, Reel DC, Johnson JC. 2003. A single origin of plastids revisited: convergent evolution in organellar genome content. J Phycol. 39(1):95–105.

Stiller JW, et al. 2014. The evolution of photosynthesis in chromist algae through serial endosymbioses. Nat Commun. 5(1):5764.

Strassert JF, Jamy M, Mylnikov AP, Tikhonenkov DV, Burki F. 2019. New phylogenomic analysis of the enigmatic phylum telonemia further resolves the eukaryote tree of life. Mol Biol Evol. 36(4):757–765.

Suzuki S, Endoh R, Manabe R, Ohkuma M, Hirakawa Y. 2018. Multiple losses of photosynthesis and convergent reductive genome evolution in the colourless green algae. Sci Rep. 8(1):940.

Tengs T, et al. 2000. Phylogenetic analyses indicate that the 19′hexanoyloxy-fucoxanthin-containing dinoflagellates have tertiary plastids of haptophyte origin. Mol Biol Evol. 17(5):718–729.

Timmis JN, Ayliffe MA, Huang CY, Martin W. 2004. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Rev Genet. 5(2):123–135.

Toso MA, Omoto CK. 2007. *Gregarina niphandrodes* may lack both a plastid genome and organelle. J Eukaryotic Microbiol. 54(1):66–72.

Watanabe MM, Suda S, Inouya I, Sawaguchi T, Chihara M. 1990. *Lepidodinium viride* gen. sp. nov. (Gymnodinaiales, Dinophyta), a green dinoflagellate with a chlorophyll A- and B-containing endosymbiont. J Phycol. 26(4):741–751.

Woo YH, et al. 2015. Chromerid genomes reveal the evolutionary path from photosynthetic algae to obligate intracellular parasites. eLife 4:e06974.

Yabuki A, et al. 2014. *Palpitomonas bilix* represents a basal cryptist lineage: insight into the character evolution in Cryptista. Sci Rep. 4(1):4641.

Yamada N, Sym SD, Horiguchi T. 2017. Identification of highly divergent diatom-derived chloroplasts in dinoflagellates, including a description of *Durinskia kwazulunatalensis* sp. nov. (Peridiniales, Dinophyceae). Mol Biol Evol. 34(6):1335–1351.

Yamada N, et al. 2019. Discovery of a kleptoplastic 'dinotom' dinoflagellate and the unique nuclear dynamics of converting kleptoplastids to permanent plastids. Sci Rep. 9(1):10474.

Yan D, et al. 2015. *Auxenochlorella protothecoides* and *Prototheca wickerhamii* plastid genome sequences give insight into the origins of nonphotosynthetic algae. Sci Rep. 5(1):14465.

Yoon H, Hackett JD, Bhattacharya D. 2002. A single origin of the peridinin- and fucoxanthin-containing plastids in dinoflagellates through tertiary endosymbiosis. Proc Natl Acad Sci U S A. 99(18):11724–11729.

Záhonová K, et al. 2018. Peculiar features of the plastids of the colourless alga *Euglena longa* and photosynthetic euglenophytes unveiled by transcriptome analyses. Sci Rep. 8(1):17012.

Zhu G, Marchewka MJ, Keithly JS. 2000. *Cryptosporidium parvum* appears to lack a plastid genome. Microbiology 146(2):315–321.

Associate editor: Geoff McFadden