Received: 27 June 2022 | Revised: 25 August 2022 | Accepted: 19 September 2022

DOI: 10.1002/pld3.454

ORIGINAL RESEARCH

The plastomes of *Hyalomonas oviformis* and *Hyalogonium fusiforme* evolved dissimilar architectures after the loss of photosynthesis

Alexandra E. DeShaw[1] | Francisco Figueroa-Martinez[2] | Thomas Pröschold[3] | Maike Lorenz[4] | Aurora M. Nedelcu[1] | David R. Smith[5] | Adrián Reyes-Prieto[1]

[1]Department of Biology, University of New Brunswick, Fredericton, New Brunswick, Canada

[2]Department of Biotechnology, CONACyT and Universidad Autónoma Metropolitana, Mexico City, Mexico

[3]Research Department for Limnology Mondsee, University of Innsbruck, Innsbruck, Austria

[4]Experimental Phycology and Culture Collection of Algae, University of Göttingen, Göttingen, Germany

[5]Department of Biology, Western University, London, Ontario, Canada

Correspondence
Adrian Reyes-Prieto, Department of Biology, University of New Brunswick, 10 Bailey Drive Fredericton, New Brunswick E3B 5A3, Canada.
Email: areyes@unb.ca

Funding information
New Brunswick Innovation Foundation, Grant/Award Number: ProjectRIF2012-006; Canada Foundation for Innovation, Grant/Award Number: Project28276; Natural Sciences and Engineering Research Council of Canada, Grant/Award Number: RGPIN-2018-04025

Abstract

The loss of photosynthesis in land plants and algae is typically associated with parasitism but can also occur in free-living species, including chlamydomonadalean green algae. The plastid genomes (ptDNAs) of colorless chlamydomonadaleans are surprisingly diverse in architecture, including highly expanded forms (*Polytoma uvella* and *Leontynka pallida*) as well as outright genome loss (*Polytomella* species). Here, we explore the ptDNAs of *Hyalomonas (Hm.) oviformis* (SAG 62-27; formerly known as *Polytoma oviforme*) and *Hyalogonium (Hg.) fusiforme* (SAG 62-1c), each representing independent losses of photosynthesis within the Chlamydophyceae. The *Hm. oviformis* ptDNA is moderately sized (132 kb) with a reduced gene complement (but still encoding the ATPase subunits) and is in fact smaller than that of its photosynthetic relative *Hyalomonas chlamydogama* SAG 11-48b (198.3 kb). The *Hg. fusiforme* plastome, however, is the largest yet observed in nonphotosynthetic plants or algae (~463 kb) and has a coding repertoire that is almost identical to that of its photosynthetic relatives in the genus *Chlorogonium*. Furthermore, the ptDNA of *Hg. fusiforme* shows no clear evidence of pseudogenization, which is consistent with our analyses showing that *Hg. fusiforme* is the nonphotosynthetic lineage of most recent origin among known colorless Chlamydophyceae. Together, these new ptDNAs clearly show that, in contrast to parasitic algae, plastid genome compaction is not an obligatory route following the loss of photosynthesis in free-living algae, and that certain chlamydomonadalean algae have a remarkable propensity for genomic expansion, which can persist regardless of the trophic strategy.

KEYWORDS
gene losses, genome reduction, *Hyalogonium*, *Hyalomonas*, loss of photosynthesis, nonphotosynthetic algae, plastid genome, *Polytoma*, *Polytomella*

1 | INTRODUCTION

The loss of photosynthesis has occurred many times throughout eukaryotic evolution (Figueroa-Martinez et al., 2015) and is typically associated with a switch to a parasitic or pathogenic lifestyle in land plants and algae (Wicke et al., 2013). Nevertheless, the transition to an obligate nonphotosynthetic lifestyle has also occurred in free-living species, including members of the green algal class Chlamydophyceae. There are currently six known free-living chlamydomonadalean lineages that have independently lost photosynthetic capabilities. These are represented by the following species/genera: *Polytomella* (Pringsheim, 1955), *Leontynka* (Pánek et al., 2022), *Polytoma* (Figueroa-Martinez et al., 2017; Nedelcu, 2001; Vernon et al., 2001), *Hyalogonium fusiforme* (Pringsheim, 1963), *Volvocales* sp. NrCl902 (Kayama et al., 2020), and *Hyalomonas oviformis*, which was formerly known as *Polytoma oviforme* (Ettl, 1983; Rumpf et al., 1996). The genus *Polytoma* represents the colorless parallel to *Chlamydomonas* and currently contains more than 30 described species (Ettl, 1983). Unfortunately, most of these species are not available in public culture collections. Rumpf et al. (1996) and Figueroa-Martinez et al. (2017) demonstrated that the genus *Polytoma* is polyphyletic. For example, the strain SAG 62-27 (former *Polytoma oviforme*) studied here is not close to the type species *Polytoma uvella* and therefore needs to be transferred to another genus, which will be proposed below (*Hyalomonas oviformis* sp. nov.). Originally, the strain SAG 62-27 was listed as *Polytoma* sp. or as *P. oviforme*, a name which has never been validly published.

Previous studies of colorless Chlamydomonadales have identified disparate and intriguing patterns of plastid genome (ptDNA) evolution. *Polytomella* species, for instance, retain a colorless plastid but no longer have ptDNA (Smith & Lee, 2014). *Polytoma uvella*, on the other hand, has a highly expanded plastome (~240 kb), resulting from an abundance of repetitive noncoding DNA (Figueroa-Martinez et al., 2017). In fact, despite having lost all genes related to photosynthesis, this ptDNA is larger than that of its close photosynthetic relative *Chlamydomonas leiostraca* (Figueroa-Martinez et al., 2017). Similarly, the ptDNA of *Leontynka pallida* (>360 kb) no longer harbors genes for photosynthesis but has accumulated huge amounts of noncoding repetitive sequences (Pánek et al., 2022). In contrast, the ptDNA of the colorless *Volvocales* sp. NrCl902 is neither extremely reduced nor highly expanded (176 kb) and has a coding capacity mirroring that of most other nonphotosynthetic plants and alga (Kayama et al., 2020).

To improve our understanding of plastid genome evolution (e.g., coding capacity, size, and architecture) in nonphotosynthetic free-living algae, we characterized the ptDNAs of *Polytoma* sp. SAG 62-27 (hereafter *Hm. oviformis* gen. et sp. nov.; Figure 1a) and *Hg. fusiforme* SAG 62-1c (Figure 1b), which represent two distinct and poorly studied colorless lineages within the Chlamydomonadales.

2 | RESULTS AND DISCUSSION

2.1 | *Hyalomonas*, a new nonphotosynthetic genus

The strain SAG 62-27 was originally assigned to the genus *Polytoma* with the species epithet *oviforme* Pringsheim (nom. nud., without valid description; Koch 1964). Phylogenetic analyses revealed that the strain SAG 62-27 is not closely related to the type species *P. uvella* (Figueroa-Martinez et al., 2015, 2017; Nedelcu, 2001; Rumpf et al., 1996). Closely related to SAG 62-27 is the photosynthetic species originally described as *Chlamydomonas chlamydogama* by Bold (1949). Both taxa represent an own lineage within the Chlamydophyceae and were therefore transferred to the newly erected genus *Hyalomonas*.

Hyalomonas gen. nov.

Description: Cells biflagellated with chloroplast of *Agloe*-type containing a single pyrenoid or reduced parietal leucoplast. Cells surrounded with glycoprotein cell wall. Plastids with prominent eyespot. Asexual reproduction by 4–8 zoospores. Sexual reproduction by isogamy.

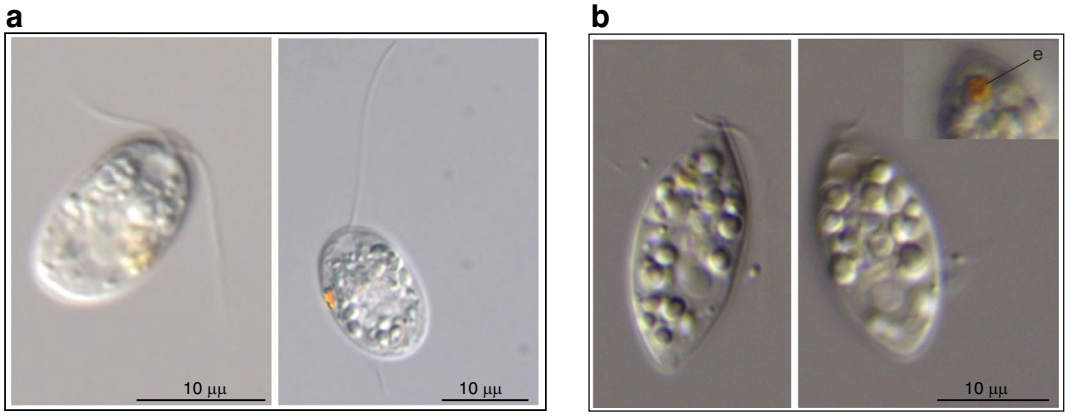

FIGURE 1 Differential interference contrast micrographs of (a) *Hyalomonas oviformis* (SAG 62-27) in low-melting point agarose and (b) *Hyalogonium fusiforme* (SAG 62-1c), with the carotene-rich eyespot (e) indicated in the inset image

Diagnosis: Differs from other genera of biflagellated chlorophytes by SSU-ITS and plastome sequence.

Type species (designated here): *Hm. oviformis* sp. nov.

Hm. oviformis sp. nov.

Description: Cells oviform, 11–15 μm long, 8–10 μm wide, two flagella with 1–1½ body length, with parietal leucoplast containing spot-like anterior eyespot. Nucleus in posterior position. Two contractile vacuoles (Figure 1a).

Diagnosis: Differs from other species of nonphotosynthetic biflagellated chlorophytes by SSU-ITS (GenBank accession OM985702) and plastome (GenBank accession OP392028.1) sequences.

Holotype (designated here): The strain SAG 62-27 cryopreserved in metabolically inactive state at the Culture Collection of Algae (SAG), University Göttingen, Germany.

Hyalomonas chlamydogama (Bold) comb. nov.

Basionym: *Chlamydomonas chlamydogama* Bold, 1949, *Bull. Torrey Bot. Cl.* **76**, 101, Figure 1 (lectotype designated here).

Epitype (designated here): The strain SAG 11-48b cryopreserved in metabolically inactive state at the Culture Collection of Algae (SAG), University Göttingen, Germany.

2.2 | *Hm. oviformis* and *Hg. fusiforme* represent independent losses of photosynthesis within the Chlamydophyceae

Our maximum likelihood (ML) phylogenetic reconstructions using nuclear and plastid SSU rRNA sequences show unambiguously that *Hm. oviforme* and *Hg. fusiforme* lost photosynthesis independently of each other and also separately from the *P. uvella* group, *Polytomella* species, and *Volvocales* sp. NrCl902 (Figure 2a). The closest photosynthetic relatives of *Hm. oviformis* in the SSU rRNA analysis

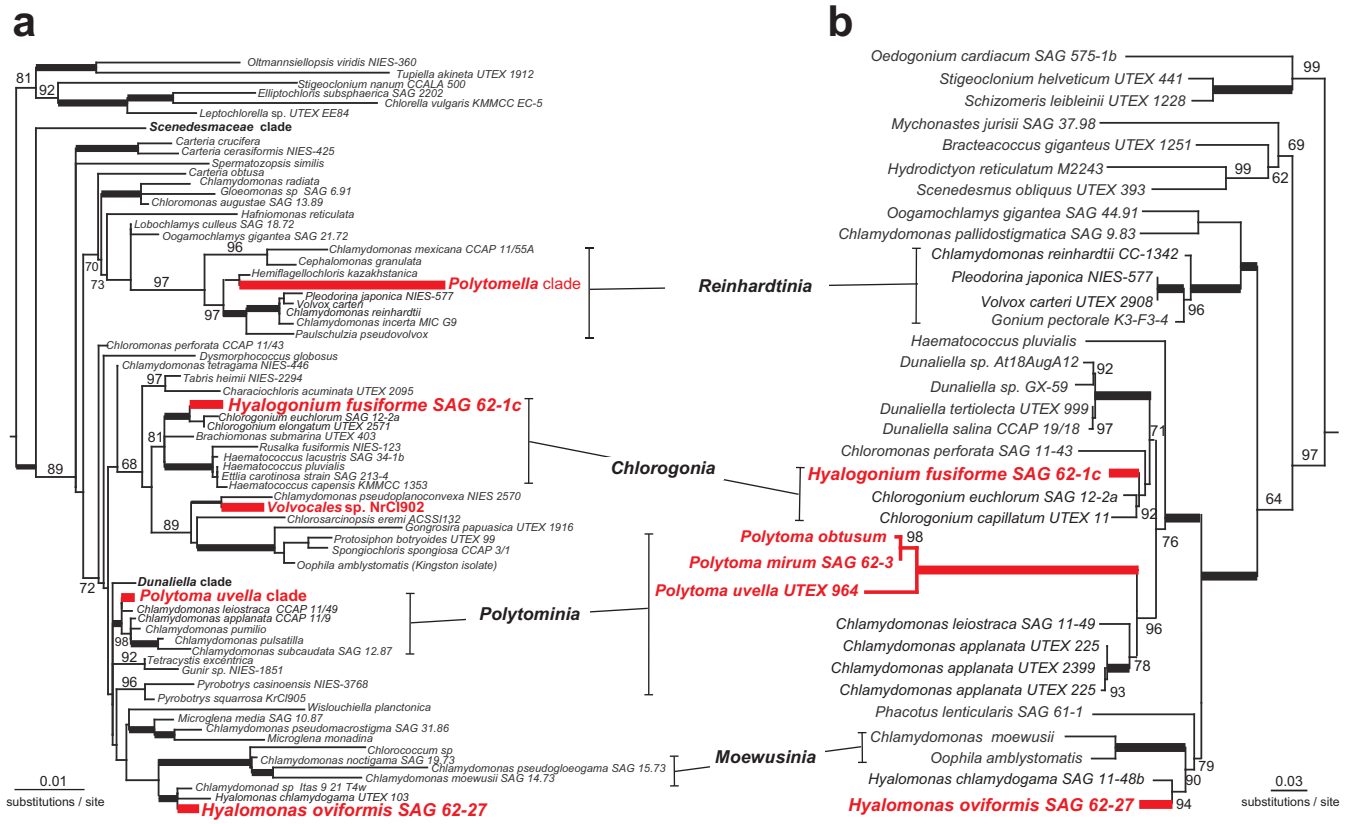

FIGURE 2 Phylogenetic trees estimated from (a) nuclear 18S rRNA and (b) plastid 16S rRNA gene sequences. Both maximum likelihood (ML) trees were estimated with IQ-TREE considering the best fitting nucleotide substitution model in each case (see text for details). The 18S rRNA ML tree was rooted using the Trebouxiophyceae clade as outgroup; in the case of the 16S rRNA tree, the outgroup reference was the three members of the OCC (Oedogoniales, Chaetophorales, and Chaetopeltidales) clade. Nonphotosynthetic chlamydomonadalean species names are indicated in red. Clades recognized within the Chlamydomonadales are also indicated (*Reinhardtinia*, *Moewusinia*, *Polytominia*, and *Chlorogonia*). Branch lengths are proportional to the number of substitutions per site as indicated by the corresponding scale bars. Numbers near nodes indicate values of bootstrap proportion support (10,000 ultra-fast replicates), and thick branches denote 100% bootstrap

are *Hm. chlamydogama* and *Chlamydomonad* sp. Itas 9 21 T4w (99% bootstrap support; BS). *Chlorogonium euchlorum* SAG 12-2a and *Chlorogonium elongatum* UTEX 2571 are the photosynthetic siblings of *Hg. fusiforme* (99% BS; Figure 2a). The recently reported colorless *Volvocales* sp. NrCl902 represents a lineage that is separate from *Hm. oviformis* and *Hg. fusiforme* and is closely related to the photosynthetic *Chlamydomonas pseudoplanoconvexa* NIES-2570 (99% BS; Figure 2a). The ML phylogenetic tree estimated with plastid 16 rRNA sequences is consistent with the nuclear data analysis and also resolves *Hm. oviformis* as sister to *Hm. chlamydogama* (94% BS) and *Hg. fusiforme* as sibling (96% BS) of *Chlorogonium* species (Figure 2b).

Our *RelTime* approximation using as reference the ML phylogenies estimated with five plastid proteins and rRNA sequences (16S rRNA +18S rRNA) consistently recovers *Hg. fusiforme* as the lineage of the most recent origin among the colorless chlamydomonadalean taxa investigated here (Figure S1 and Table S1). Our results also indicate that the divergence of *Hm. oviformis* and *P. uvella* from their respective photosynthetic siblings occurred during close, likely overlapping periods (Figure S1 and Table S1).

2.3 | Two divergent nonphotosynthetic plastomes

To explore the genomic consequences of forfeiting photosynthesis in colorless algae of distinct evolutionary age, we sequenced the ptDNAs of *Hm. oviformis and Hg. fusiforme* and those of their respective close photosynthetic relatives *Hm. chlamydogama* and *Cg. euchlorum* (Figure 2). The plastomes of *Hm. oviformis* (132 kb) and *Hm. chlamydogama* (198.3 kb) were both assembled as circular-mapping molecules (Figure 3 and Table S2). Unlike *Hm. chlamydogama*, the *Hm. oviformis* ptDNA lacks inverted repeats, which is a recurring theme among plastomes of nonphotosynthetic plants and algae (Figueroa-Martinez et al., 2015, 2017; Wicke et al., 2013). The presence of large numbers of repeats prevented us from generating complete assemblies of the *Hg. fusiforme* and *Cg. euchlorum* ptDNAs, despite employing long-read PacBio sequencing. The *Hg. fusiforme* and *Cg. euchlorum* plastome sequences are currently distributed across 43 and 4 scaffolds with cumulative lengths of 463 kb and 314 kb, respectively (Figure S2). The recovery of incomplete plastid chromosomes is not unusual in these cases given Chlamydomonadalean ptDNAs are renowned for being extremely hard to assemble (Gaouda et al., 2018), with many reported assemblies existing in highly fragmented forms (Figueroa-Martinez et al., 2017; Gaouda et al., 2018). The majority of the *Hg. fusiforme* scaffolds contain 1–2 protein-coding regions, and our read coverage analyses indicate that all the *Hg. fusiforme* and *Cg. euchlorum* scaffolds are genuine components of the plastid genome and not ptDNA-like sequences from the mitochondrial or nuclear genomes. The read coverage survey also supports the existence of an inverted repeat in *Hg. fusiforme* (Table S3).

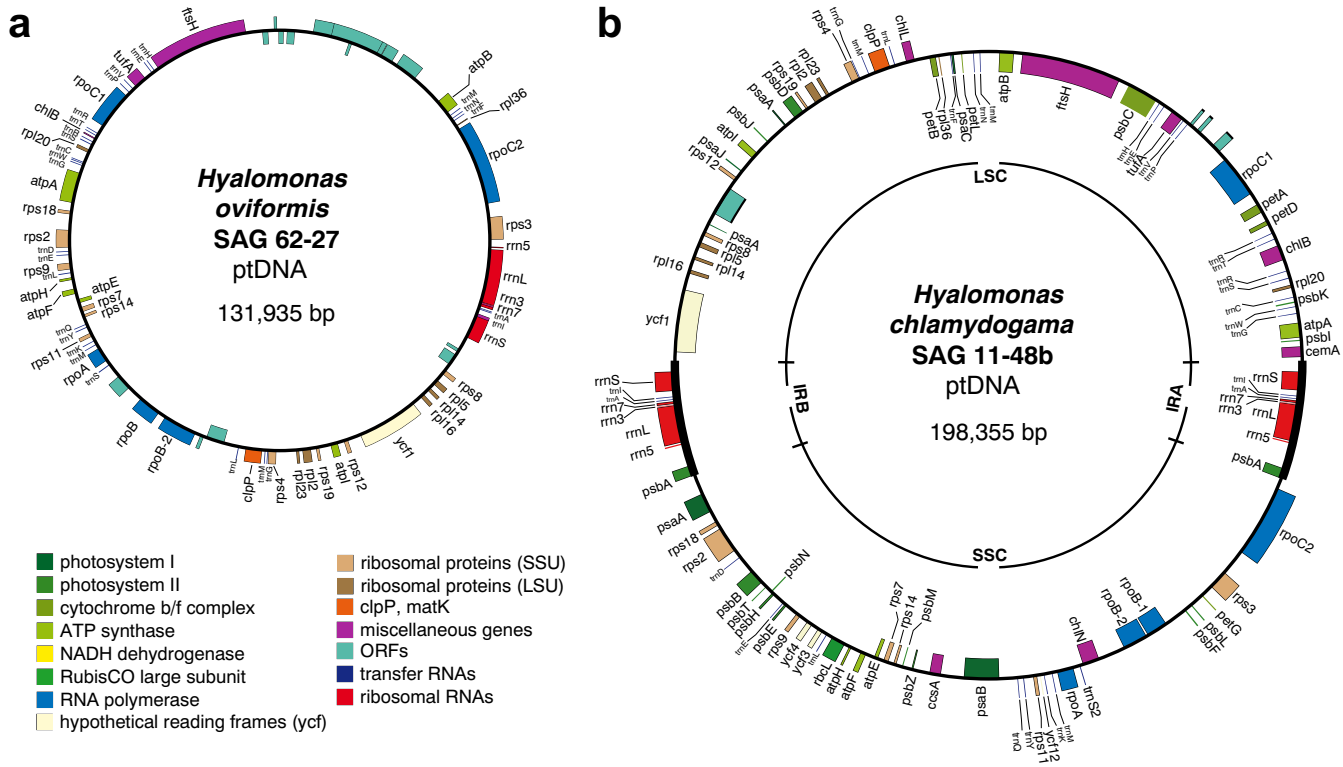

FIGURE 3 Circular maps of the plastid genomes of (a) *Hyalomonas oviformis* (SAG 62-27) and (b) *Hyalomonas chlamydogama* (SAG 11-48b). Large single copy (LSC), small single copy (SSC), and inverted repeat (IR) regions are indicated. The graphic representations were prepared with the OrganellarGenomeDRAW (OGDRAW) v1.3.1 tool (Greiner et al., 2019) implemented in the CHLOROBOX server (https://chlorobox.mpimp-golm.mpg.de/index.html)

2.4 | Comparative plastid genomics among nonphotosynthetic species and their photosynthetic relatives

To better understand the evolutionary mechanisms that shaped the genomic architecture of nonphotosynthetic plastomes, we compared the ptDNAs of *Hm. oviformis* and *Hg. fusiforme* with those of their closest known photosynthetic relatives as well as other colorless species. The *Hm. oviformis* plastome, despite being about 65 kb smaller than that of its photosynthetic relative *Hm. chlamydogama*, is quite still large when compared to the ptDNAs from other colorless plants and algae, which are typically under 75 kb. One reason for this is that *Hm. oviformis* encodes all subunits of the cF_0-CF_1 ATPase (*atpA*, *atpB*, *atpE*, *atpF*, *atpH*, and *atpI*; Figure 4). These six genes have similar lengths to their *Hm. chlamydogama* orthologues, but four (*atpA-E-F-H*) have elevated rates of nonsynonymous substitutions (dN) relative to those from *Hg. fusiforme* and various photosynthetic chlamydomonadaleans (Figure 5). This implies that even though some of the retained *atp* coding regions in the *Hm. oviformis* ptDNA appear functional they are, in fact, under relaxed selective forces and their functions are likely less constrained than in the plastids of related chlamydomonadaleans, including *Hg. fusiforme*.

The genes of the *atp* series complex are not generally retained in colorless plastids, but notable exceptions to varying degrees come from the chlamydomonadalean *L. pallida* (Pánek et al., 2022), some species of the trebouxiophyte genus *Prototheca* (Knauf & Hachtel, 2002), other diverse nonphotosynthetic algae (Tanifuji et al., 2020), and some angiosperms (Wicke et al., 2013). Nevertheless, like the plastomes of other colorless plants and algae, the *Hm. oviformis* ptDNA has lost all genes encoding subunits of the two photosystems, the b_6f complex, and the enzymes for chlorophyll biosynthesis (Figure 4 and Table S4). However, it is still possible to distinguish conserved collinear plastid gene arrangements between *Hm. oviformis* and the photosynthetic *Hm. chlamydogama* (Figure S3).

The presence of the cF_0-CF_1 ATPase in a nonphotosynthetic context is intriguing (Dorrell et al., 2019; Tanifuji et al., 2020). Some have suggested that its hydrolytic activity sustains a proton electrochemical gradient essential for the internalization of proteins into the thylakoid via the twin arginine translocator (TAT) system (Dorrell et al., 2019). Support for this view in *Hm. oviformis* comes from the fact that, in addition to plastid-encoded *atp* genes, we uncovered nuclear genes for *atp*D and *atp*G, which respectively encode the delta and gamma subunits of the plastid cF_0-CF_1 ATPase. The deduced amino acid sequences of *atp*D and *atp*G both contain standard N-terminus plastid transit peptides, and the latter contains the two cysteine residues involved in the thioredoxin-mediated modulation of the enzyme hydrolytic activity (Hisabori et al., 2012) (Figure S4). We also searched the *Hm. oviformis* nuclear contigs for *tat*A, *tat*B, and *tat*C, which encode the TAT protein import system, but only recovered evidence of *tat*C. Similar findings were recently reported in the gene repertoire of *L. pallida*, but in that case *tat*A and *tat*B were identified in transcriptomic data (Pánek et al., 2022).

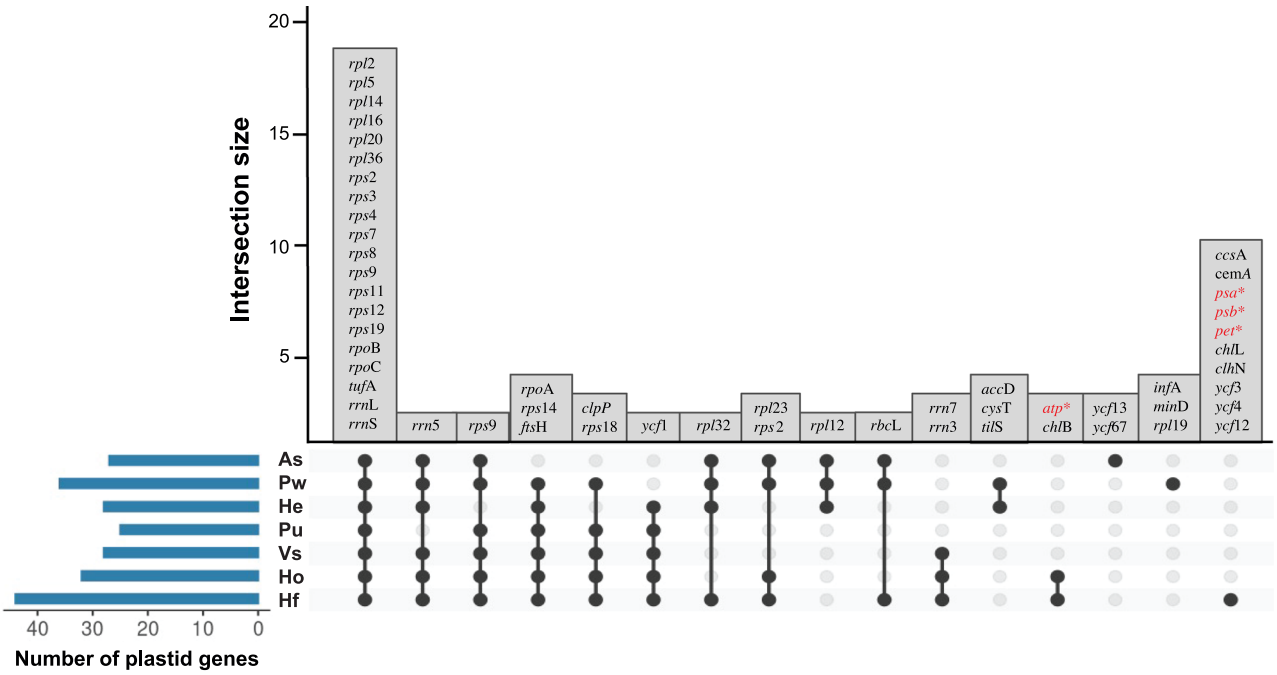

FIGURE 4 Comparative content of protein-coding genes in plastid genomes of diverse nonphotosynthetic green algae and *Euglena (Astasia) longa*, which contains a secondary plastid of green algal origin. Gene names with asterisks refer to genes encoding subunits of the CF-ATPase (*atp* series), the b_6f complex (*pet* series), photosystem I (*psa* series), and photosystem II (*psb* series). Taxa codes: As, *Euglena longa*; Pw, *Prototheca wickerhamii*; He, *Helicosporidium* sp.; Pu, *Polytoma uvella*; Vs, *Volvocales* NrCl902; Ho, *Hyalomonas oviformis*; Hf, *Hyalogonium fusiforme*. Black and gray dots indicate the presence or absence, respectively, of the protein-coding regions in the indicated plastid genomes

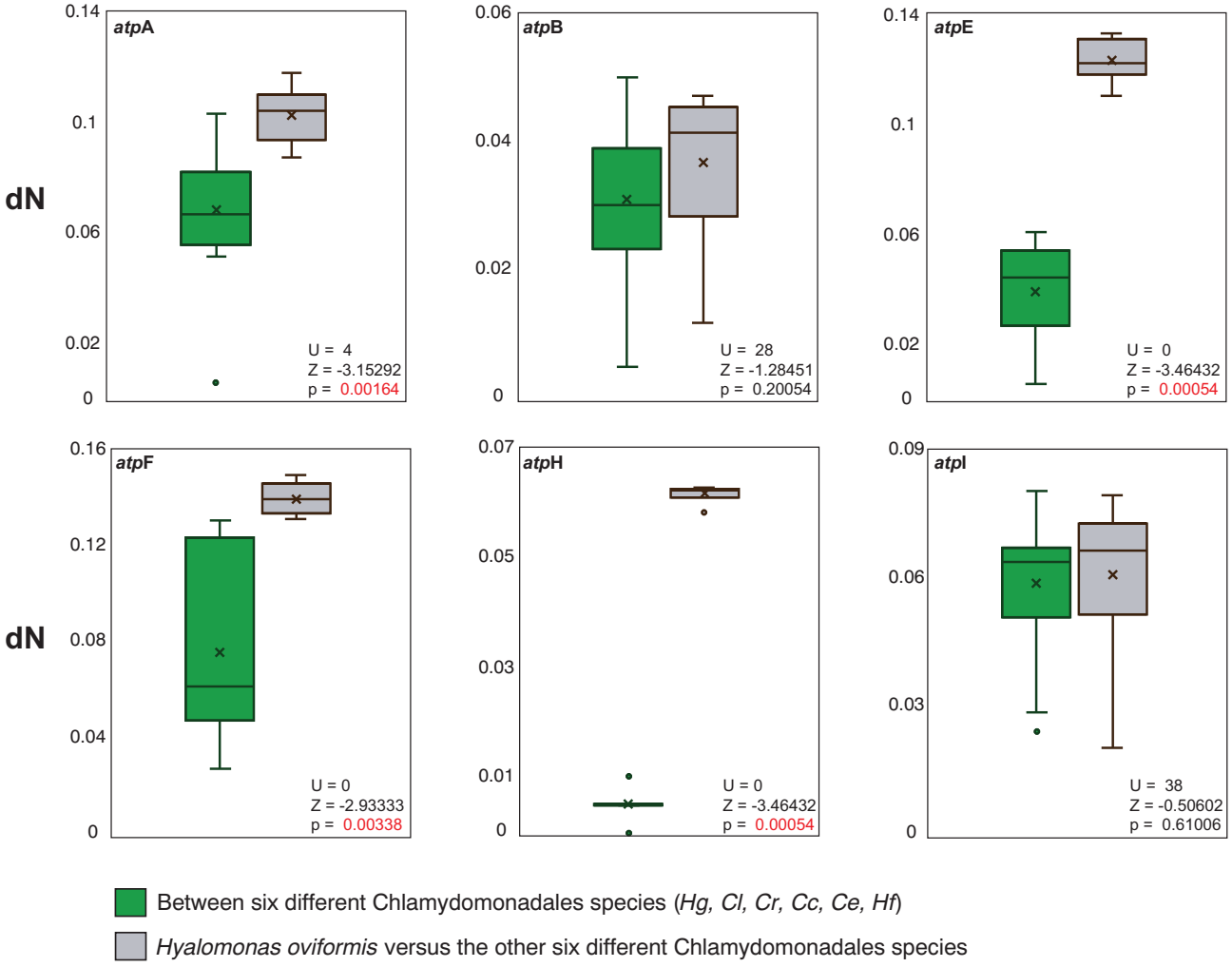

FIGURE 5 Comparison of nonsynonymous (dN) nucleotide substitution rates estimated in the five plastid genes encoding subunits of the F-ATPase. Box plots in green color correspond to nucleotide substitution rates estimated between the six species *Hyalomonas chlamydogama (Hg)*, *Chlamydomonas leiostraca (Cl)*, *Chlamydomonas reinhardtii (Cr)*, *Chlorogonium capillatum (Cc)*, *Chlorogonium euchlorum (Ce)*, and *Hyalogonium fusiforme (Hf)*. Box plots in gray color summarize the nucleotide substitution rates between the mentioned six chlamydomonadalean taxa and *Hyalomonas oviformis*. The *Z* and *U* scores and the *p* values estimated with the Wilcoxon–Mann–Whitney *U* test are indicated in each case

At circa 463 kb, the ptDNA of *Hg. fusiforme* is not only larger than that of its close photosynthetic relative *Cg. euchlorum* (314 kb; intron-rich genome), but it is also the largest plastome ever observed in a colorless plant or alga and larger than the vast majority of sequenced ptDNAs from photosynthetic species. This expanded architecture is primarily the consequence of inflated intergenic regions as well as apparent accumulation of intron-resembling sequences (Tables 1 and S6). Close inspection of the intergenic ptDNA of *Hg. fusiforme* revealed dozens of TrRnEn-like open reading frames (ORFs) belonging to different families of homing endonucleases, including members of the LAGLIDADG (2), GIY-YIG (4), and H-N-H (14) families as well as ORFs matching to reverse transcriptases and maturases (see Table S6). These ORFs were not evidently associated with introns and were one of the key reasons why we were unable to bridge the scaffolds during ptDNA assembly. Other chlamydomonadalean algae, including *Dunaliella salina*, also have a preponderance of free-standing intronic ORFs within their intergenic regions (Smith et al., 2010). The fragmented state of the *Hg. fusiforme* ptDNA assembly prevented synteny analyses with the *Chlorogonium* ptDNAs to identify gene/operon rearrangements. However, the abundance of homing endonucleases as well as the presence of some fragmented genes in the *Hg. fusiforme* ptDNA suggest that considerable plastid gene rearrangements occurred in *Hg. fusiforme* after it diverged from its photosynthetic relatives.

The large size of *Hg. fusiforme* ptDNA is not only a reflection of increased intergenic and intronic regions. Unlike the ptDNAs of *Hm. oviformis* and other nonphotosynthetic chlamydomonadaleans, the *Hg. fusiforme* plastome shows virtually no gene loss as compared to its photosynthetic close relatives (Table S4). Indeed, it encodes most subunits of both photosystems (*psa* and *psb*), the b_6f complex, and the proteins for chlorophyll biosynthesis present in *Chlorogonium* species (Table S4). What is more, the majority of the *Hg. fusiforme* plastid genes typically involved in photosynthesis display no obvious reductions in size or classic signs of pseudogenization (although some

TABLE 1 Plastid genome features of diverse chlorophytes

Taxa	Genbank accession	Size (kb)	Inverted repeat length (kb)	Coding DNA (%)	GC content (%)	Mean intergenic spacer (bp)	Protein-coding genes	rRNA genes	tRNA genes	Introns
Chlorophyceae										
Hyalomonas oviformis[a]	OP392028.1	132	NA	40.6	27.3	617	33	5	27	3
Hyalomonas chlamydogama[a]	OP298011.1	198	12.1	40.9	27.6	886	66	10	29	6
Hyalogonium fusiforme[a]	OP597712.1-OP597753.1	~442	NA	18.4	44	2,454	66	6	21	4
Chlorogonium euchlorum[a]	OP597708.1-OP597711.1	~314	27.9	27.9	36.9	1,692	67	6	31	14
Chlorogonium capillatum	KT625085.1-KT620591.1	~271	NA	34.7	38.8	1,616	65	3	27	10
Polytoma uvella	KX828177.1	~230	NA	26.8	27.3	2,967	24	2	27	1
Chlamydomonas leiostraca	NC_032109	167	13.1	52.4	42.5	600	68	6	28	8
Volvocales sp. *NrCl902*	LC516060.1	176	NA	39.2	40.1	1974	24	5	30	5
Chlamydomonas reinhardtii	NC_005353.1	204	22.2	37.8	34.5	700	68	10	29	6
Volvox carteri	GU084820.1	>420	16.6	<20	42.8	5,103	60	4	27	8
Dunaliella salina	GQ250045.1	269	14	45.5	32.1	935	66	3	28	43
Haematococcus pluvialis	MG677935.1	1,352	NA	~90	~50	NA	~71–125	6	31	NA
Leontynka pallida	OM479425.1	362	NA	18	36.1	4,726	32	3	26	4
Trebouxiophyceae										
Helicosporidium sp.	NC_008100.1	37	NA	95.2	26.9	34	21	3	25	1
Prototheca wickerhamii	KJ000176.1	56	NA	83.17	31.2	135	34	3	26	0

Note: Gray boxes indicate nonphotosynthetic taxa.
NA, no detected or no present.
[a]This work.

of these genes are fragmented in several coding regions; Table S5), with the exception of *chl*B (1068 bp), which is around 130 amino acids shorter than its *Cg. elongatum* counterpart. Still, the *Hg. fusiforme chl*B encodes all the amino acid residues important for the activity of the protochlorophyllide oxidoreductase (Bröcker et al., 2010) (Figure S5). It is notable as well that *psb*E and *psb*Z from *Hg. fusiforme* are split into two pieces by an intergenic spacer (no intron was detected), but their conceptual translations reveal protein products of similar lengths to those of photosynthetic relatives. Altogether, there is no clear evidence of rampant plastid pseudogenization in *Hg. fusiforme*, despite it having forfeited photosynthetic capabilities (i.e., no chlorophyll presence; Figure 2b; Pringsheim, 1963).

Evolutionary rate analyses revealed that certain *Hg. fusiforme* genes typically involved in photosynthesis have significantly ($p < .05$; Mann–Whitney test) higher nonsynonymous substitution rates than their counterparts in photosynthetic Chlamydomonadales (Figure S6 and Table S8). For instance, *psb*A, which encodes protein D1 (one of two components that make up the core of Photosystem II), has accumulated nonsynonymous substitution at higher rate than its counterparts in photosynthetic Chlamydomonadales (Figure S6). Similar trends were recorded for *rbc*L (large subunit of RuBisCO) and the photosystem genes *psb*D and *psb*I (Figure S6). Substitution rates in a representative set of house-keeping genes (*tuf*A, *clp*P, and several ribosomal proteins) indicate that there are no significant dN rate difference between *Hg. fusiforme* and photosynthetic Chlamydomonadales (Table S8). Because the nucleotide composition of the plastid protein-coding regions (Table S9) and the estimated codon usage (Table S10) of *Hg. fusiforme* and the two *Chlorogonium* species studied here are highly similar, it is unlikely that the significant differences in dN rates in those few photosynthesis-related genes are due to compositional or codon preference biases. Overall, these results indicate that certain *Hyalogonium* plastid-encoded proteins involved directly in light reactions (i.e., three *psb* genes) and carbon fixation (i.e., *rbc*L) are accumulating amino acid substitutions at a faster pace than their homologous sequences in photosynthetic relatives.

2.5 | Plastome evolution following the loss of photosynthesis

The loss of photosynthesis has traditionally been associated with plastid gene loss and genomic streamlining. In parallel, plastome gene repertoires of nonphotosynthetic plants and algae, regardless of their trophic strategy, tend to converge on relatively homogeneous small gene sets involved in plastid transcription and translation (Figure 4). Only in rare cases has the plastome remodeling process after the loss of photosynthesis resulted in the complete loss of the organelle genome (Smith & Lee, 2014). But as more and more ptDNA sequences from colorless species become available, it is increasingly obvious that there is not a single narrative for plastome evolution following the loss of photosynthetic capabilities. This is certainly true for colorless members of the Chlamydomonadales, where plastid genome information exists for at least six nonphotosynthetic lineages, each representing an independent loss of photosynthesis within the order. A key theme emerging from these data is that the ptDNAs of colorless plastids, in particular those of free-living Chlamydomonadales, are not always characterized by small sizes, reduced gene contents, and compactness. In fact, diverse combinations of genome size, gene content, and gene density exist in this green algal group. The notable cases include the large gene-rich plastome of *Hg. fusiforme* as well as the large but gene-reduced ptDNAs of *L. pallida* (Pánek et al., 2022) and *P. uvella* (Figueroa-Martinez et al., 2017); nevertheless, all three of these plastid genomes are significantly larger than those of their close photosynthetic relatives. But this group of colorless algae also includes the relatively small and compact plastomes of *Hm. oviformis* and *Volvocales* sp. NrCl902 and the complete loss of ptDNA in *Polytomella* species.

The observed diversity among ptDNAs of colorless chlamydomonadalean algae highlights the fact that various aspects of genome organization—size, compaction, and gene content—evolve independently, even under seemingly similar selective pressures associated with the switch to a strictly heterotrophic lifestyle. This is because they are likely driven by different evolutionary mechanisms and forces. For instance, gene loss is primarily a consequence of (1) reduced functional constraints on photosynthesis-related genes and (2) evolutionary time, whereby the longer the time since the loss of photosynthesis the greater the potential for gene loss. Conversely, changes in genome compaction are largely driven by mutational forces and DNA maintenance machineries acting upon intergenic and intronic regions (Figueroa-Martinez et al., 2015; Wicke et al., 2013). Thus, it is possible for a plastid genome to be large and gene poor, and vice versa.

Our data suggest that *Hg. fusiforme* may have lost photosynthetic capabilities in recent evolutionary time, thus, explaining why it has a near-complete cohort of photosynthetic genes. In the future, it is possible/expected that some of these genes will be lost, but its propensity for genomic expansion may (or may not) persist. Overall, our data suggest *Hg. fusiforme* is in an early stage of ptDNA evolution following the loss of photosynthesis. It should be noted that chlamydomonadalean ptDNAs, including those of photosynthetic species, are particularly prone to genomic inflation, a topic that has been discussed in detail elsewhere (Figueroa-Martinez et al., 2017; Gaouda et al., 2018; Smith et al., 2010). In fact, the largest plastid genome on record belongs to *Haematococcus pluvialis* (>1 Mb) (Smith, 2018), a relatively close relative of *Hg. fusiforme*. As it stands, the data presented here add to the astonishing amount of plastome architectural diversity observed across the Chlamydomonadales (in both photosynthetic and nonphotosynthetic species). How much more diversity remains to be uncovered? Our guess is a lot.

3 | MATERIALS AND METHODS

3.1 | Growth conditions

Hm. oviformis (SAG 62-27) and *Hg. fusiforme* (SAG 62-1c) were grown in ErbsMS media (Pringsheim, 1946) under dark conditions and no

shaking at 16°C. Cultures of *Hm. chlamydogama* (SAG 11-48b) were maintained in standard *Volvox* medium (Kirk et al., 1999) under constant shaking (200 rpm) at 18°C and a 16-h-light/8-h-dark cycle. *Cg. euchlorum* (SAG 12-2a) cultures were grown in MiEg 1:1 and Eg media (Schlösser, 1994).

3.2 | DNA extraction and sequencing

Total DNA of *Hm. chlamydogama* was extracted from a 250-mL culture in exponential growth. The cell pellet was washed three times with saline-ethylenediaminetetra-acetate (EDTA) (50 mM Tris–HCl, pH 8, 50 mM NaCl, and 5 mM EDTA) and then incubated in the presence of proteinase K (33 mg/ml) and .5% SDS for 1 h at 50°C. DNA was extracted by standard phenol-chloroform procedures. Total genomic DNA from *Hm. oviformis*, *Hg. fusiforme*, and *Cg. euchlorum* was extracted using the DNeasy PlantMini Kit (Qiagen, Hilden, Germany) following manufacturer instructions. Paired-end libraries were sequenced using Illumina technology (150 bp length reads; 450-bp insert length; Genome Quebec, McGill University). We also produced low-coverage PacBio reads (PacBio Sequel, v3 chemistry; IMR Dalhousie University) of *Hg. fusiforme* (413,219 reads; circa 604 Mb) and *Cg. euchlorum* (555,390 reads; circa 680 Mb) from selected DNA fragments of ~ 350 Kbp. The 18S rRNA and ITS rRNA sequences of the used strains are available in GenBank (SAG 62-27: OM985702; SAG 11-48b: OM985703; SAG 62-1c: OM985704; SAG 12-2a: OM985705).

3.3 | Genome assembly and sequence annotation

After quality control filtering and trimming, we assembled the paired-end read sets of each alga with Ray v 2.2.0 [k-mer lengths of 21, 31, and 77] (Boisvert et al., 2010). PacBio SMRT reads of *Hg. fusiforme* and *Cg. euchlorum* were corrected and assembled with Canu v 2.1 [predicted genome size of 0.4 megabases, using raw PacBio reads for correction and then corrected reads for assembly] (Koren et al., 2017).

We identified contigs containing typical plastid genes using the Automatic Annotation tool of Geneious Prime v2019.1.3 (Kearse et al., 2012) with published complete ptDNAs from diverse Chlamydomonadales as reference. Manual sequence refinement and read coverage evaluations were carried out with Geneious Prime. Genes encoding tRNAs were predicted with the tRNAscan-SE Search Server (Lowe & Chan, 2016).

3.4 | Phylogenetics, relative divergence time, and nucleotide substitution rates analyses

Sequence multiple alignments for all analyses were prepared with the MAFFT (Katoh & Standley, 2013) tool implemented in Geneious Prime. Resulting raw sequence matrices were visually inspected and manually edited to discard poorly aligned regions. To test the independent origin of the colorless lineages *Hm. oviformis* and *Hg. fusiforme*, we estimated ML trees from multiple alignments of nuclear 18S rRNA (1521 bp long; 121 taxa) and plastid 16S rRNA (1257 bp long and 35 taxa) sequences of diverse chlorophyte green algae. Both ML reconstructions were estimated independently with IQ-TREE v1.6.11 (Nguyen et al., 2015) considering the best fitting nucleotide substitution model (TIM3 + R3 and TVMe+I + G4, selected respectively by the ModelFinder version implemented in IQ-TREE v1.6.11; -m TEST option) and 10,000 ultra-fast bootstrap replicates in each case.

Then, to estimate the relative divergence times (rDT) of the chlamydomonadalean colorless lineages, we prepared additional alignments of 16S rRNA+18S rRNA (2763 nucleotides long; 28 taxa) sequences and of conceptual translations of the protein-coding genes *tuf*A, *rps*4, *rps*7, *rpl*2, and *rpl*5 (1163 amino acid residues long; 28 taxa) considering a reduced sample of chlorophyte algae (see supporting information Table S7). The five protein-coding sequences were selected based on their presence in ptDNAs of the colorless algae analyzed here, the length of the translated protein (>150 aa), and reliability of the multiple alignments (i.e., limited ambiguously aligned or gap-rich regions). ML trees of the 16S rRNA, 18S rRNA, the concatenated 16S rRNA and 18S rRNA sets, and concatenated protein alignments were independently estimated with IQ-TREE considering the best fitting model (Model Finder -m TEST option) for each sequence matrix. Then, we used those ML phylogenetic trees as references to estimate rDT with the RelTime approximation implemented in MEGA X (Mello, 2018). We considered the *Proterocladus* fossils (780–1000 MY) as the minimum divergence time between Ulvophyceae-Bryopsidales and Chlorophyceae as calibration reference point (Tang et al., 2020). Taking into account the limitations of our time calibration approach (i.e., a single fossil data point), taxon sampling, and sequence data (e.g., number of used genes), our rDT estimations were solely designed to provide us with a relative temporal context to contrast and interpret the results of our comparative ptDNA analysis in colorless Chlamydomonadales, but not to provide precise answers about divergence times within this algal group.

Finally, we also prepared multiple alignments at nucleotide sequence level of selected protein-coding regions from diverse photosynthetic and free-living nonphotosynthetic chlamydomonadalean algae to estimate rates of synonymous (dS) and nonsynonymous (dN) substitutions with the yn00 tool of PAMLX (Xu & Yang, 2013).

3.5 | Accession numbers

Data deposition: the new plastome sequences reported here have been deposited in GenBank under accession numbers OP392028.1, OP298011.1, OP597708.1-OP597711.1 and OP597712.1-OP597753.1.

ACKNOWLEDGMENTS

This work was supported by Discovery Grant to ARP from the Natural Sciences and Engineering Research Council (NSERC; project RGPIN-2018-04025) of Canada. ARP was also supported by the Canada

Foundation for Innovation (project 28276) and the New Brunswick Innovation Foundation (project RIF2012-006).

AUTHOR CONTRIBUTIONS

AED performed all laboratory work and computational analyses, in collaboration with FFM. ARP, AED, DRS, AMN, and FFM conceived the original research plan; ML and TP provided cell cultures and purified DNA, performed microscopy documentation, and prepared the taxonomic description; ARP supervised experimental and data analysis; AED, DRS, and ARP led the writing of the article with contributions from all the authors; ARP agrees to serve as the author responsible for contact and communication.

ONE SENTENCE SUMMARY

The plastid genomes of two free-living chlamydomonadalean algae, *Hm. oviformis* and *Hg. fusiforme*, reveal different evolutionary stages following the loss of photosynthesis.

CONFLICT OF INTEREST

The authors did not report any conflict of interest.

ORCID

Alexandra E. DeShaw https://orcid.org/0000-0002-9227-293X
Francisco Figueroa-Martinez https://orcid.org/0000-0003-0037-5091
Maike Lorenz https://orcid.org/0000-0002-2277-3077
Aurora M. Nedelcu https://orcid.org/0000-0002-7517-2419
David R. Smith https://orcid.org/0000-0001-9560-5210
Adrián Reyes-Prieto https://orcid.org/0000-0002-0413-6162

REFERENCES

Armenteros, J. J. A., Salvatore, M., Emanuelsson, O., Winther, O., Heijne, G., von Elofsson, A., & Nielsen, H. (2019). Detecting sequence signals in targeting peptides using deep learning. *Life Science Alliance*, *2*, e201900429. https://doi.org/10.26508/LSA.201900429

Boisvert, S., Laviolette, F., & Corbeil, J. (2010). Ray: Simultaneous assembly of reads from a mix of high-throughput sequencing technologies. *Journal of Computational Biology*, *17*, 1519–1533. https://doi.org/10.1089/cmb.2009.0238

Bold, H. C. (1949). The morphology of *Chlamydomonas chlamydogama*, sp. nov. *Bull Torrey Bot cl*, *76*, 101–108. https://doi.org/10.2307/2482218

Bröcker, M. J., Schomburg, S., Heinz, D. W., Jahn, D., Schubert, W. D., & Moser, J. (2010). Crystal structure of the nitrogenase-like dark operative protochlorophyllide oxidoreductase catalytic complex (ChlN/ChlB)2. *The Journal of Biological Chemistry*, *285*, 27336–27345. https://doi.org/10.1074/jbc.M110.126698

Dorrell, R. G., Azuma, T., Nomura, M., de Kerdrel, G. A., Paoli, L., Yang, S., Bowler, C., Ishii, K., Miyashita, H., Gile, G. H., & Kamikawa, R. (2019). Principles of plastid reductive evolution illuminated by nonphotosynthetic chrysophytes. *Proceedings of the National Academy of Sciences USA*, *116*, 6914–6923. https://doi.org/10.1073/pnas.1819976116

Ettl, H. (1983). Chlorophyta I- Phytomonadina. In H. Ettl, J. Gerloff, H. Heynig, & D. Mollenhauer (Eds.), *Süßwasserflora von Mitteleuropa* (Vol. Bd. 9). Gustav Fischer, Stuttgart.

Figueroa-Martinez, F., Nedelcu, A. M., Smith, D. R., & Reyes-Prieto, A. (2015). When the lights go out: The evolutionary fate of free-living colorless green algae. *New Phytologist*, *206*, 972–982. https://doi.org/10.1111/nph.13279

Figueroa-Martinez, F., Nedelcu, A. M., Smith, D. R., & 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 Physiology*, *173*, 932–943. https://doi.org/10.1104/pp.16.01628

Gaouda, H., Hamaji, T., Yamamoto, K., Kawai-Toyooka, H., Suzuki, M., Noguchi, H., Minakuchi, Y., Toyoda, A., Fujiyama, A., Nozaki, H., & Smith, D. R. (2018). Exploring the limits and causes of plastid genome expansion in Volvocine green algae. *Genome Biology and Evolution*, *10*, 2248–2254. https://doi.org/10.1093/GBE/EVY175

Greiner, S., Lehwark, P., & Bock, R. (2019). OrganellarGenomeDRAW (OGDRAW) version 1.3.1: Expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Research*, *47*, W59–W64. https://doi.org/10.1093/nar/gkz238

Hisabori, T., Sunamura, E. I., Kim, Y., & Konno, H. (2012). The chloroplast ATP synthase features the characteristic redox regulation machinery. *Antioxidants & Redox Signaling*, *19*, 1846–1854. https://doi.org/10.1089/ars.2012.5044

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*, 772–780. https://doi.org/10.1093/MOLBEV/MST010

Kayama, M., Chen, J. F., Nakada, T., Nishimura, Y., Shikanai, T., Azuma, T., Miyashita, H., Takaichi, S., Kashiyama, Y., & Kamikawa, R. (2020). A non-photosynthetic green alga illuminates the reductive evolution of plastid electron transport systems. *BMC Biology*, *18*, 126. https://doi.org/10.1186/s12915-020-00853-w

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., & Drummond, A. (2012). Geneious basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*, 1647–1649. https://doi.org/10.1093/bioinformatics/bts199

Kirk, M. M., Stark, K., Miller, S. M., Muller, W., Taillon, B. E., Gruber, H., Schmitt, R., & Kirk, D. L. (1999). *regA*, a *Volvox* gene that plays a central role in germ-soma differentiation, encodes a novel regulatory protein. *Development*, *126*, 639–647. https://dev.biologists.org/content/126/4/639, https://doi.org/10.1242/dev.126.4.639

Knauf, U., & Hachtel, W. (2002). The genes encoding subunits of ATP synthase are conserved in the reduced plastid genome of the heterotrophic alga *Prototheca wickerhamii*. *Molecular Genetics and Genomics*, *267*, 492–497 https://doi.org/10.1007/s00438-002-0681-6

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., & Phillippy, A. M. (2017). Canu: Scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Research*, *27*, 722–736 https://doi.org/10.1101/gr.215087.116

Lowe, T. M., & Chan, P. P. (2016). tRNAscan-SE on-line: Integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Research*, *44*, W54–W57. https://doi.org/10.1093/nar/gkw413

Mello, B. (2018). Estimating TimeTrees with MEGA and the TimeTree resource. *Molecular Biology and Evolution*, *35*, 2334–2342. https://doi.org/10.1093/molbev/msy133

Nedelcu, A. M. (2001). Complex patterns of plastid 16S rRNA gene evolution in nonphotosynthetic green algae. *Journal of Molecular Evolution*, *53*, 670–679. https://doi.org/10.1007/s002390010254

Nguyen, L. T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, *32*, 268–274. https://doi.org/10.1093/molbev/msu300

Pánek, T., Barcytė, D., Treitli, S. C., Záhonová, K., Sokol, M., Ševčíková, T., Zadrobílková, E., Jaške, K., Yubuki, N., Čepička, I., & Eliáš, M. (2022). A new lineage of non-photosynthetic green algae with extreme organellar genomes. *BMC Biology*, *20*, 66. https://doi.org/10.1186/s12915-022-01263-w

Pringsheim, E. G. (1946). *Pure cultures of algae: Their preparation and maintenance*. Cambridge University Press.

Pringsheim, E. G. (1955). The genus *Polytomella*. *The Journal of Protozoology*, *2*, 137–145. https://doi.org/10.1111/j.1550-7408.1955.tb02413.x

Pringsheim, E. G. (1963). *Farblose Algen*. Gustav Fischer Verlag, Stuttgart, Germany.

Rumpf, R., Vernon, D., Schreiber, D., & Birky, jr C. W. (1996). Evolutionary consequences of the loss of photosynthesis in Chlamydomonadaceae: Phylogenetic analysis of *rrn*18 (18S rDNA) in 13 *Polytoma* strains (Chlorophyta). *Journal of Phycology*, *32*, 119–126. https://doi.org/10.1111/j.0022-3646.1996.00119.x

Schlösser, U. G. (1994). SAG-Sammlung von Algenkulturen at the University of Göttingen. *Catalogue of Strains. Botanica Acta*, *107*, 113–186. https://doi.org/10.1111/j.1438-8677.1994.tb00784.x

Smith, D. R. (2018). *Haematococcus lacustris*: The makings of a giant-sized chloroplast genome. *AoB Plants*, *10*, ply058. https://doi.org/10.1093/aobpla/ply058

Smith, D. R., & Lee, R. W. (2014). A plastid without a genome: Evidence from the nonphotosynthetic green algal genus *Polytomella*. *Plant Physiology*, *164*, 1812–1819. https://doi.org/10.1104/pp.113.233718

Smith, D. R., Lee, R. W., Cushman, J. C., Magnuson, J. K., Tran, D., & Polle, J. E. W. (2010). The *Dunaliella salina* organelle genomes: Large sequences, inflated with intronic and intergenic DNA. *BMC Plant Biology*, *10*, 83. https://doi.org/10.1186/1471-2229-10-83

Tamura, K., Battistuzzi, F. U., Billing-Ross, P., Murillo, O., Filipski, A., & Kumar, S. (2012). Estimating divergence times in large molecular phylogenies. *Proceedings of the National Academy of Sciences USA*, *109*, 19333–19338. https://doi.org/10.1073/pnas.1213199109

Tang, Q., Pang, K., Yuan, X., & Xiao, S. (2020). A one-billion-year-old multicellular chlorophyte. *Nature Ecology & Evolution*, *4*, 543–549. https://doi.org/10.1038/s41559-020-1122-9

Tanifuji, G., Kamikawa, R., Moore, C. E., Mills, T., Onodera, N. T., Kashiyama, Archibald, J.M., Inagaki, Y., & Hashimoto, T. (2020). Comparative plastid genomics of *Cryptomonas* species reveals fine-scale genomic responses to loss of photosynthesis. *Genome Biology and Evolution*, *12*, 3926–3937. https://doi.org/10.1093/gbe/evaa001

Vernon, D., Gutell, R. R., Cannone, J. J., Rumpf, R. W., & Birky, J. (2001). Accelerated evolution of functional plastid rRNA and elongation factor genes due to reduced protein synthetic load after the loss of photosynthesis in the chlorophyte alga *Polytoma*. *Molecular Biology and Evolution*, *18*, 1810–1822. https://doi.org/10.1093/oxfordjournals.molbev.a003968

Wicke, S., Müller, K. F., de Pamphilis, C. W., Quandt, D., Wickett, N. J., Zhang, Y., Renner, S. S., & Schneeweiss, G. M. (2013). Mechanisms of functional and physical genome reduction in photosynthetic and nonphotosynthetic parasitic plants of the broomrape family. *Plant Cell*, *25*, 3711–3725. https://doi.org/10.1105/tpc.113.113373

Xu, B., & Yang, Z. (2013). PamlX: A graphical user interface for PAML. *Molecular Biology and Evolution*, *30*, 2723–2724. https://doi.org/10.1093/molbev/mst179

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: DeShaw, A. E., Figueroa-Martinez, F., Pröschold, T., Lorenz, M., Nedelcu, A. M., Smith, D. R., & Reyes-Prieto, A. (2022). The plastomes of *Hyalomonas oviformis* and *Hyalogonium fusiforme* evolved dissimilar architectures after the loss of photosynthesis. *Plant Direct*, *6*(10), e454. https://doi.org/10.1002/pld3.454