What Happened before Losses of Photosynthesis in Cryptophyte Algae?

Shigekatsu Suzuki [iD],*,[1] Ryo Matsuzaki,[1,2] Haruyo Yamaguchi,[1] and Masanobu Kawachi[1]

[1]Biodiversity Division, National Institute for Environmental Studies, Tsukuba, Ibaraki, Japan
[2]Faculty of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki, Japan

*Corresponding author: E-mail: suzuki.shigekatsu@nies.go.jp.
Associate editor: Michael Purugganan

Abstract

In many lineages of algae and land plants, photosynthesis was lost multiple times independently. Comparative analyses of photosynthetic and secondary nonphotosynthetic relatives have revealed the essential functions of plastids, beyond photosynthesis. However, evolutionary triggers and processes that drive the loss of photosynthesis remain unknown. Cryptophytes are microalgae with complex plastids derived from a red alga. They include several secondary nonphotosynthetic species with closely related photosynthetic taxa. In this study, we found that a cryptophyte, *Cryptomonas borealis*, is in a stage just prior to the loss of photosynthesis. *Cryptomonas borealis* was mixotrophic, possessed photosynthetic activity, and grew independent of light. The plastid genome of *C. borealis* had distinct features, including increases of group II introns with mobility, frequent genome rearrangements, incomplete loss of inverted repeats, and abundant small/medium/large-sized structural variants. These features provide insight into the evolutionary process leading to the loss of photosynthesis.

Key words: cryptophyte, *Cryptomonas borealis*, genome evolution, mixotrophy, photosynthetic loss.

Carbon fixation by photosynthesis is beneficial to organisms. However, secondary losses of photosynthesis have been found in various algae and plants (Krause 2012). Sequencing analyses have revealed the processes of functional and genome reductions as well as the essential functions of the remaining plastid after photosynthetic losses, including roles in carbon fixation and fatty acid, terpenoid, tetrapyrrole, and isoprenoid biosynthesis (Gardner et al. 2002; Borza et al. 2005; Pombert et al. 2014; Janouškovec et al. 2017; Suzuki et al. 2018; Dorrell et al. 2019). However, evolutionary events prior to the loss of photosynthesis and the factors triggering these losses are unclear. Cryptophytes are unicellular microalgae with complex plastids derived from a red alga (Douglas and Penny 1999; Kim et al. 2017). Cryptophytes are generally phototrophic but include some mixotrophic and heterotrophic species (Antia et al. 1973; Gervais 1997) (fig. 1a–g and supplementary table S1, Supplementary Material online). In particular, *Cryptomonas* contains some leucoplast-bearing nonphotosynthetic species, for example, *Cryptomonas paramaecium*. The losses of photosynthesis were occurred independently at least three times in *Cryptomonas* (Hoef-Emden et al. 2005). Therefore, cryptophytes are a useful group to elucidate process of secondary losses of photosynthesis.

The loss of photosynthesis is closely related to a mixotrophic lifestyle (Dorrell et al. 2019). Several *Cryptomonas* species can uptake glucose; however, no known photosynthetic species can be cultured independent of light with a substantial carbon source (Hoef-Emden and Archibald 2017). To find photosynthetic cryptophytes culturable independent of light,

we cultivated five axenic *Cryptomonas* strains for 20 days under heterotrophic conditions (i.e., in the dark with a moderate concentration of organic compound mix [BSM]; fig. 1h). Over 12 days, *Cryptomonas borealis* NIES-276 and *Cryptomonas tetrapyrenoidosa* NIES-348 cells increased more substantially in BSM than in the absence of BSM ($P < 0.05$, *t*-test), indicating that these strains can grow by uptake of BSM without photosynthesis. *Cryptomonas* sp. NIES-3952 increased for the first 4 days with or without BSM; however, the number of cells remained constant after 12 days with BSM, suggesting that cells survived under heterotrophic conditions, as reported previously for *Cryptomonas phaseolus* and *Cryptomonas undulata* (Gervais 1997). *Cryptomonas curvata* NIES-281 and *Cryptomonas* sp. NIES-345 cell counts decreased for the initial 12 days with/without BSM, indicating that these strains need light for growth. *Cryptomonas paramaecium*, *C. borealis*, and NIES-345 are monophyletic (*borealis/paramaecium* clade; fig. 1g). Therefore, the common ancestor of the clade could have possessed heterotrophy (i.e., osmotrophy), with a subsequent loss in NIES-345. To confirm photosynthetic activity in these species, we measured PS II quantum yields (*Fv/Fm*) by pulse amplitude modulation (PAM) fluorometry (fig. 1i). The *Fv/Fm* value for a nonphotosynthetic species, *C. paramaecium* NIES-715, was 0, and the mean values were 0.49–0.79 for the other *Cryptomonas* spp., indicating that they possess photosynthetic activity. The mean *Fv/Fm* value for *C. borealis* was not significantly different from those of the other photosynthetic/mixotrophic strains ($P > 0.05$, *t*-test).
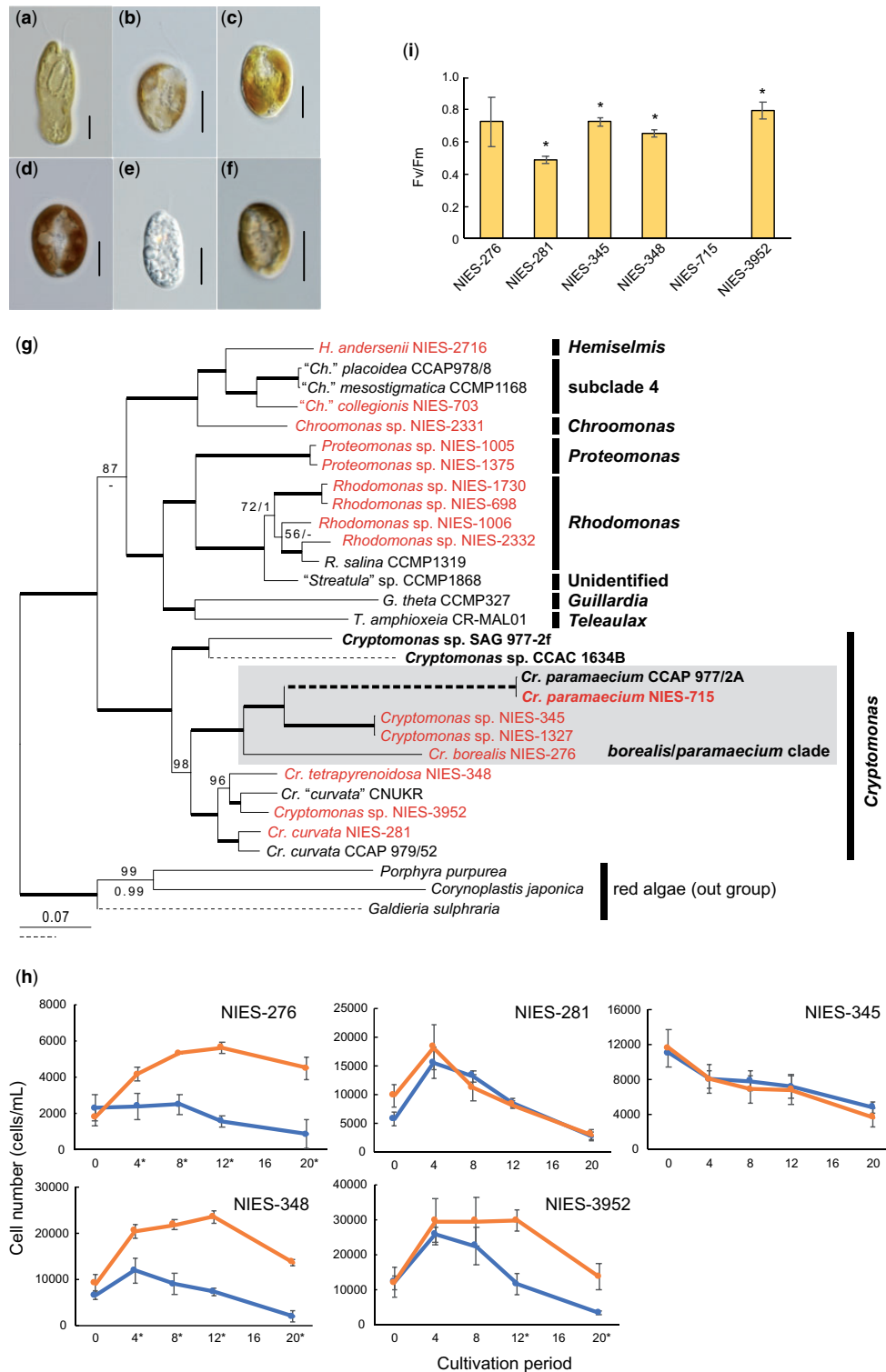
Fɪɢ. 1. Diversity of *Cryptomonas* species. (*a*) *Cryptomonas borealis* NIES-276. (*b*) *Cryptomonas curvata* NIES-281. (*c*) *Cryptomonas* sp. NIES-345. (*d*) *Cryptomonas tetrapyrenoidosa* NIES-348. (*e*) *Cryptomonas paramaecium* NIES-715. (*f*) *Cryptomonas* sp. NIES-3952. (*g*) Maximum likelihood (ML) tree based on 35 plastid genome-encoded proteins. The data set was composed of 6,058 amino acids of 30 operation taxonomic units (OTUs). Bootstrap support values (BP) are shown above each node. Bayesian posterior probabilities (BPP) are shown below each node. Bold lines represent BP=100 and BPP=1. BP<50 or BPP<0.95 are not shown. Red-colored OTUs are species for which the plastid genomes have been sequenced in this study. Bold OTUs are secondary nonphotosynthetic species. (*h*) Growth of five photosynthetic *Cryptomonas* species under dark conditions. Cultivation was performed for 20 days with organic compounds (BSM; red lines) or without BSM (blue lines). Asterisks above cultivation periods indicate significant differences between with and without BSM ($P < 0.05$, *t*-test). Error bars represent standard deviations. Culture experiments were independently performed three times. (*i*) Mean PS II quantum yields (*Fv/Fm*) by PAM fluorometry. NIES-715 is nonphotosynthetic (negative control). Error bars represent standard deviations. Asterisks indicate no significant differences between NIES-276 and other photosynthetic strains ($P > 0.05$, *t*-test). Experiments were performed independently three times.

The structures of plastid genomes are highly conserved in photosynthetic cryptophytes (Khan et al. 2007; Moore et al. 2012; Kim et al. 2015, 2017); however, Tanifuji et al. (2020) suggested that the inverted repeat (IR) loss is accompanied by frequent rearrangements and the loss of photosynthesis. In this study, we sequenced 16 plastid genomes of cryptophytes and detected no rearrangements in most taxa (fig. 1g and supplementary figs. S1–S6, supplementary tables S2–S4, and supplementary text 1–3, Supplementary Material online). However, the plastid genomes of *C. borealis* and NIES-345 showed structural alterations, despite their photosynthetic ability (fig. 2a). We estimated genome rearrangement scenarios of the *borealis/paramaecium* clade using *C. curvata* NIES-281 as a reference. Four, one, and eight putative genome rearrangements were detected in *C. borealis*, NIES-345, and *C. paramaecium*, respectively (supplementary fig. S7 and supplementary text 3, Supplementary Material online). Interestingly, NIES-345 lacks an IR and *C. borealis* had incomplete IRs, which may correspond to an intermediate form of IR loss (fig. 2b and supplementary fig. S8, Supplementary Material online). *Cryptomonas borealis* possessed an intact rRNA operon and pseudogenized 23S rRNA (ψrnl) in distant regions. These results show that IR loss occurs before a loss of photosynthesis in cryptophytes. Similar patterns are observed in apicomplexans, including a few photosynthetic species. The plastid genome of photosynthetic *Vitrella brassicaformis*, without IRs, has frequent genome rearrangements, whereas those of nonphotosynthetic relatives with IRs have conserved genome structures (Kwong et al. 2019). The plastid genome of *Gloeotilopsis sarcinoidea* (Ulvophyceae) shows a similar fragmented *rnl* (Turmel et al. 2016); however, the evolutionary history differs from that of cryptophytes. This alga possesses a 3′ terminus of *rnl* (ψrnl) adjacent to an intact *rnl*, and ψrnl may be produced by an incomplete duplication mediated by a mobile element (Turmel et al. 2016). *Cryptomonas borealis* also showed unique deletions of *ycf19* and *ycf29* (fig. 2c and supplementary fig. S5 and supplementary text 2, Supplementary Material online). Ycf19 is related to plastid division (Kabeya et al. 2010) and Ycf29 is a plastid response regulator of the two-component system, which is related to the transcriptional regulation (Puthiyaveetil and Allen 2009). The deletion of genes related to the maintenance of plastids suggests weak selective pressure on photosynthesis in *C. borealis*. This can be supported by nonsynonymous and synonymous substitution rates (dN/dS), showing that the dN/dS values of conserved orthologs were higher in *C. borealis* and NIES-345 than those in the other photosynthetic *Cryptomonas* strains (supplementary fig. S9, Supplementary Material online). Moreover, codon usage bias was predicted using measure independent of length and composition (MILC) values (Supek and Vlahovicek 2005) (supplementary fig. S10, Supplementary Material online). A higher MILC value indicates stronger codon usage bias and change of the gene expression level. The MILC values of photosynthetic genes to nonphotosynthetic genes in *C. borealis* and NIES-345 were slightly higher than those of the other photosynthetic *Cryptomonas* strains, suggesting that the expression level of photosynthetic genes of

C. borealis and NIES-345 might be changed. Based on its ability to grow independent of light, phylogenetic position, and structural alternations of the plastid genome, we propose that *C. borealis* represents the form just prior to the loss of photosynthesis.

The plastid genome of *C. borealis* was 161.0 kb, which was at least ~15 kb larger than those of other cryptophytes (supplementary table S3, Supplementary Material online). This size difference can be explained by the expansion of intergenic regions and increasing introns. The plastid genome of *C. borealis* had a lower gene density (0.9 genes/kb) than those of the others (1.027–1.201 genes/kb) (supplementary fig. S7, Supplementary Material online). *Cryptomonas borealis* possessed 14 group II introns and one nonfunctional intron in 11 protein-coding genes, which was substantially more than the 0–7 introns observed in other cryptophytes (supplementary table S5 and supplementary text 4, Supplementary Material online). The 14 introns had the six conserved domains in group II introns (supplementary fig. S11, Supplementary Material online), whereas the remaining intron (the *clpC* intron) might be nonfunctional without a similar domain structure or the conserved sequences at the 5′- and 3′-ends (GUGYG and AY, respectively) (Lambowitz and Zimmerly 2011). All of the introns except for the *psbN* and *clpC* introns were spliced in our RNA-seq analysis (supplementary fig. S12, Supplementary Material online). The sequence of the *psbD* intron of *C. borealis* matched that of the *psbN* intron of *C. curvata* NIES-281. Moreover, in *C. borealis*, the introns of *rps6* and *atpB* shared 77.5% sequence similarity. Some introns had five intron-encoded proteins (IEPs) in *C. borealis* (supplementary table S5, Supplementary Material online). To elucidate the origin of the introns, we performed a phylogenetic analysis based on the IEPs (supplementary figs. S13–S15 and supplementary text 4, Supplementary Material online). The IEPs in *petG* of "*Chroomonas*" *placoidea* CCAP 978/8 and "*Chroomonas*" *mesostigmatica* CCMP1168 are sister to an IEP of the prasinophyte *Pyramimonas parkeae* (Kim et al. 2017). Most of the IEPs such as *psbN* IEPs were phylogenetically related to red algae, suggesting that the intron might originate in their plastid. For *C. borealis*, the second *psbB* IEP was closely related to the trebouxiophyceans *Stichococcus bacillaris* and *Jenufa minuta*. The IEP of *atpB* also formed a monophyletic group with the IEPs of some chlorophyceans. However, these two IEPs were not monophyletic, suggesting that some of the group II introns and their IEPs were transferred from green algae several times.

The IEPs of cryptophytes have RT subdomains 0–7 and an X domain, but lack the typical D and En domains (Khan and Archibald 2008) (supplementary fig. S16 and supplementary text 4, Supplementary Material online). Combined with the lack of the YADD motif in RT subdomain 5, these findings suggest that cryptophyte introns are immobile. To confirm the mobility of the group II introns, we evaluated DNA–RNA hybrids uniquely in intronic regions because DNA–RNA hybrids are temporally generated during reverse transcription when the intron RNA is inserted into the genome (Lambowitz and Zimmerly 2011) (fig. 2d and e; supplementary fig. S17, Supplementary Material online). We targeted the
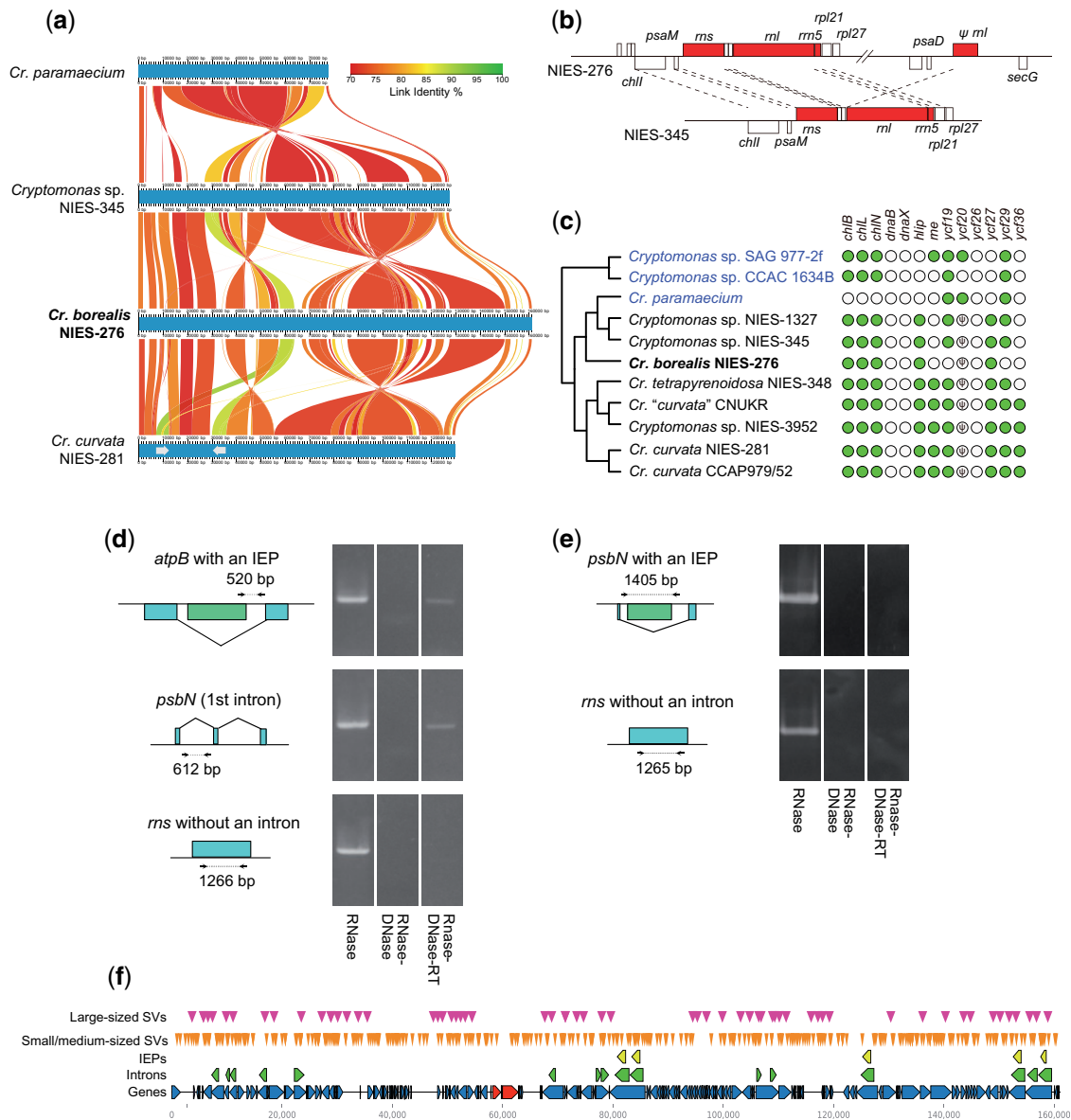
Fig. 2. Structural changes in the plastid genome of *Cryptomonas borealis* NIES-276. (*a*) Genome rearrangements of strains in the *borealis/paramaecium* clade. *Cryptomonas curvata* NIES-281 is used as a reference. Syntenic regions are linked, and colors represent nucleotide identities. Blue bars represent chromosomes of the plastid genomes. (*b*) Genome structures around genes for ribosomal RNAs in *C. borealis* NIES-276 and *Cryptomonas* sp. NIES-345. Red genes are *rns*, *rnl*, and *rrn5*. Orthologous genes are linked with dotted lines. (*c*) Gene deletion in *Cryptomonas* species. Secondarily lost genes in the plastid genome of photosynthetic cryptophytes are shown. Green and white circles represent the presence and absence of genes, respectively. Pseudogenes (ψ) are also shown in white circles. Blue indicates nonphotosynthetic species. (*d*, *e*) Intron mobility and SVs in *C. borealis* NIES-276. Intron–exon structures and tests for intron mobility for *C. borealis* NIES-276 (*d*) and *C. curvata* NIES-281 (*e*). IEPs are represented in green. Arrows represent positions of the specific primers and sizes of the PCR products are shown. The detailed methods are shown in supplementary figure S15, Supplementary Material online. (*f*) Gene models of NIES-276 and positions of predicted small-, medium-, and large-sized SVs. Protein-coding genes, rRNAs, introns, and IEPs are represented in blue, red, green, and yellow, respectively. Orange and pink arrowheads show positions of putative small- and medium-sized indels and large SVs.

introns of *atpB* (with an IEP) and *psbN* (without an IEP) of *C. borealis* and the intron of *psbN* (with an IEP) of *C. curvata* NIES-281. For *C. borealis*, we detected DNA–RNA hybrids in both the *atpB* intron and *psbN* intron; however, a DNA–RNA hybrid was not found in nonintronic regions of 16S rRNA. We did not detect DNA–RNA hybrids in *C. curvata*. These results indicate that the two introns in *C. borealis* are mobile, and the intron in *C. curvata* is immobile. Because the *psbN* intron of *C. borealis* lacked IEPs, it may be transposed via IEPs encoded

by other introns. Therefore, several group II introns might have expanded via transposition in *C. borealis*. The cryptophyte IEPs lacked En domains, which have endonuclease activity to cleave the opposite strand for reverse transcription (San Filippo and Lambowitz 2002). However, the intron RNA of *C. borealis* was reverse-transcribed, implying that reverse transcription can be performed using a nascent strand at a replication fork for priming (Zhong and Lambowitz 2003; Martínez-Abarca et al. 2004). The factors leading to the
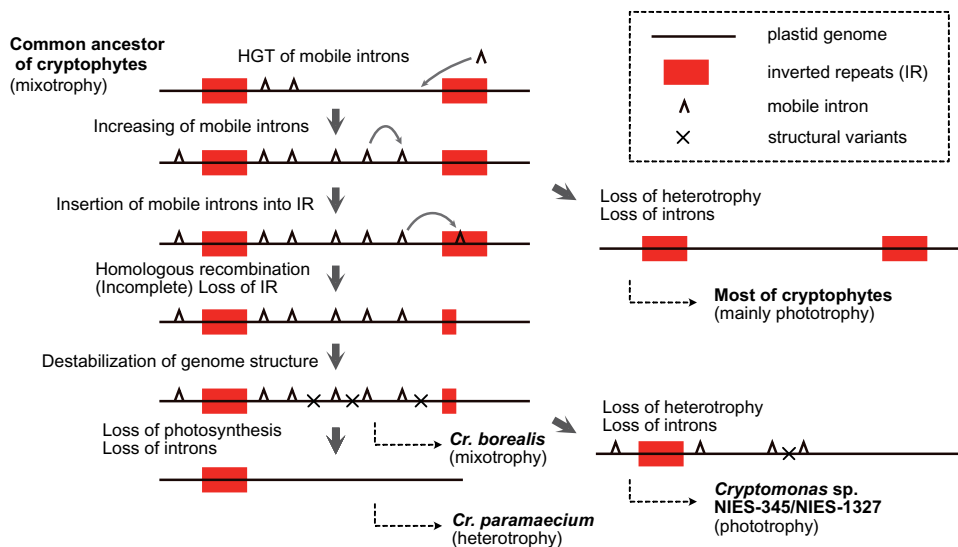
Fig. 3. Evolutionary scenario of plastid genomes prior to the loss of photosynthesis in cryptophytes. The common ancestor of cryptophytes is mixotrophic (photosynthetic and eukaryovorous). Some exogenous group II introns have been inserted in the plastid genome as well as the original introns. These introns retained mobility and expanded in the genome. Most (obligate) photosynthetic cryptophytes lacked heterotrophy and showed losses of introns. These species had highly conserved plastid genomes. The mobile introns could facilitate homologous recombination, thereby inducing genome rearrangements and IR loss. Furthermore, IR loss could prevent error correction of the genome structure, leading to mutation accumulation. *Cryptomonas borealis* is an example. In this condition, some species lacked heterotrophy and showed losses of introns under strong selective pressure, for example, *Cryptomonas* sp. NIES-345, with more conserved plastid genomes than that of *C. borealis*. The loss of photosynthesis can induce genome compaction, reducing the introns (*Cryptomonas paramaecium*). The reduced genome minimizes the opportunity for mutation.

immobility of the intron of *C. curvata* are unclear; however, the IEP of *C. curvata* lacked the FLG motif in RT subdomain 7, which is a primer grip domain (Blocker et al. 2005) (supplementary fig. S16, Supplementary Material online).

We found more structural variants (SVs) in the plastid genome of *C. borealis* than in the plastid genomes of the other relative strains. First, we detected small- to medium-sized indels in the plastid genomes based on Illumina short reads (supplementary tables S6–S9 and supplementary text 5, Supplementary Material online). *Cryptomonas borealis* had 325 small- to medium-sized SVs (1–386 bp), including 289 insertions and 36 deletions. By contrast, the other strains had fewer SVs. *Cryptomonas curvata* NIES-281, NIES-345, and *C. paramaecium* possessed 2–42 SVs. All of the deletions were 1–2 bp, implying that longer deletions are unlikely. The SV frequencies were low, with mean frequencies of 0.55%–3.16% in them. Relatively large numbers of SVs were predicted in *C. borealis*, corresponding to the large number of introns. Indeed, nonphotosynthetic *C. paramaecium* had few SVs and no introns. These data suggest that mobile introns can promote low-frequency insertions. SVs were distributed throughout the genomes (fig. 2f). For example, in the *C. borealis* genome, 56.9%, 27.1%, and 16.0% of SVs were located in genes, intergenic regions, and introns (including IEPs), respectively (supplementary table S6, Supplementary Material online). Moreover, we detected 61 large SVs of 83–1,430 bp (insertions, deletions, inversions, and inverted duplications) in the plastid genome of *C. borealis* using nanopore long reads (fig. 2f and supplementary table S10, Supplementary Material online.). Because essential genes had SVs involving critical

amino acid substitutions, the SVs can exist in some copies of the plastid DNA of individual cells. *Guillardia theta* possesses 130–260 copies of the plastid genome per cell (Hirakawa and Ishida 2014). Therefore, the multicopies of plastid genomes can mask the deleterious effects of SVs.

Based on the genomic features of *C. borealis* and its relatives, we propose an evolutionary scenario explaining the events before the loss of photosynthesis in the cryptophyte plastid genomes (fig. 3). The common ancestor of cryptophytes is mixotrophic with photosynthesis and eukaryovory (Hoef-Emden 2005). Most cryptophytes lacked heterotrophy; however, for some lineages, including secondary heterotrophic species, eukaryovory was converted to osmotrophy (e.g., *C. borealis* and *C. tetrapyrenoidosa*). This conversion has also been inferred in euglenozoans (Yamaguchi et al. 2012). The common ancestor of cryptophytes possessed mobile group II introns originated from its endosymbiont and green algae. The introns expand and are maintained in mixotrophic species but degenerate and decrease in obligate photosynthetic species with strong selective pressure. Under heterotrophic conditions, the increase in group II introns can be retained and trigger frequent SVs in mixotrophic species. IR structures can be disrupted by the insertion of introns or by homologous recombination, and IR losses result in the destabilization of genome structures (Palmer and Thompson 1982; Strauss et al. 1988). Finally, photosynthesis-related genes might be lost in nonphotosynthetic species. The last step will be evaluated by the laboratory evolution of *C. borealis* under heterotrophic conditions over a long period.

Supplementary Material

Acknowledgments

Data Availability

The plastid genome sequences of the 16 cryptophyte strains are available at the DDBJ with the accession numbers LC648950–LC648965. The raw read data are available with the accession numbers DDBJ: DRA012570 and DRA012637. Sequences for 18S rRNA of 15 cryptophyte strains are available with the accession numbers DDBJ: LC647555–LC647560, LC647562–LC647564, and LC647568–LC647573.

References

Antia NJ, Kalley JP, McDonald JT, Bisalputra T. 1973. Ultrastructure of the marine cryptomonad *Chroomonas salina* cultured under conditions of photoautotrophy and glycerol-heterotrophy. *J Protozool*. 20(3):377–385.

Blocker FJ, Mohr G, Conlan LH, Qi L, Belfort M, Lambowitz AM. 2005. Domain structure and three-dimensional model of a group II intron-encoded reverse transcriptase. *RNA* 11(1):14–28.

Borza T, Popescu CE, Lee RW. 2005. Multiple metabolic roles for the nonphotosynthetic plastid of the green alga *Prototheca wickerhamii*. *Eukaryot Cell*. 4(2):253–261.

Dorrell RG, Azuma T, Nomura M, Audren de Kerdrel G, Paoli L, Yang S, Bowler C, Ishii K, Miyashita H, Gile GH, et al. 2019. Principles of plastid reductive evolution illuminated by nonphotosynthetic chrysophytes. *Proc Natl Acad Sci U S A*. 116(14):6914–6923.

Douglas SE, Penny SL. 1999. The plastid genome of the cryptophyte alga, *Guillardia theta*: complete sequence and conserved synteny groups confirm its common ancestry with red algae. *J Mol Evol*. 48(2):236–244.

Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, et al. 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419(6906):498–511.

Gervais F. 1997. Light-dependent growth, dark survival, and glucose uptake by cryptophytes isolated from a freshwater chemocline. *J Phycol*. 33(1):18–25.

Hirakawa Y, Ishida K-I. 2014. Polyploidy of endosymbiotically derived genomes in complex algae. *Genome Biol Evol*. 6(4):974–980.

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.

Hoef-Emden K, Archibald JM. 2017. Cryptophyta (Cryptomonads). In: Archibald JM, Simpson AGB, Slamovits CH, editors. Handbook of the protists. Cham (Switzerland): Springer. p. 851–891.

Hoef-Emden K, Tran HD, Melkonian M. 2005. Lineage-specific variations of congruent evolution among DNA sequences from three genomes, and relaxed selective constraints on rbcL in *Cryptomonas* (Cryptophyceae). *BMC Evol Biol*. 5:1–11.

Janouškovec J, Gavelis GS, Burki F, Dinh D, Bachvaroff TR, Gornik SG, Bright KJ, Imanian B, Strom SL, Delwiche CF, et al. 2017. Major transitions in dinoflagellate evolution unveiled by phylotranscriptomics. *Proc Natl Acad Sci U S A*. 114(2):E171–E180.

Kabeya Y, Nakanishi H, Suzuki K, Ichikawa T, Kondou Y, Matsui M, Miyagishima S. 2010. The YlmG protein has a conserved function related to the distribution of nucleoids in chloroplasts and cyanobacteria. *BMC Plant Biol*. 10:57.

Khan H, Archibald JM. 2008. Lateral transfer of introns in the cryptophyte plastid genome. *Nucleic Acids Res*. 36(9):3043–3053.

Khan H, Parks N, Kozera C, Curtis BA, Parsons BJ, Bowman S, Archibald JM. 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 JI, Moore CE, Archibald JM, Bhattacharya D, Yi G, Yoon HS, Shin W. 2017. Evolutionary dynamics of cryptophyte plastid genomes. *Genome Biol Evol*. 9(7):1859–1872.

Kim JI, Yoon HS, Yi G, Kim HS, Yih W, Shin W. 2015. The plastid genome of the cryptomonad *Teleaulax amphioxeia*. *PLoS One* 10(6):e0129284.

Krause K. 2012. Plastid genomes of parasitic plants: a trail of reductions and losses. In: Bullerwell CE, editor. Organelle genetics: evolution of organelle genomes and gene expression. Berlin Heidelberg (Germany): Springer-Verlag. p. 79–103.

Kwong WK, del Campo J, Mathur V, Vermeij MJA, Keeling PJ. 2019. A widespread coral-infecting apicomplexan with chlorophyll biosynthesis genes. *Nature* 568(7750):103–107.

Lambowitz AM, Zimmerly S. 2011. Group II introns: mobile ribozymes that invade DNA. *Cold Spring Harb Perspect Biol*. 3(8):a003616.

Martínez-Abarca F, Barrientos-Durán A, Fernández-López M, Toro N. 2004. The RmInt1 group II intron has two different retrohoming pathways for mobility using predominantly the nascent lagging strand at DNA replication forks for priming. *Nucleic Acids Res*. 32(9):2880–2888.

Moore CE, Curtis B, Mills T, Tanifuji G, Archibald JM. 2012. Nucleomorph genome sequence of the cryptophyte alga *Chroomonas mesostigmatica* CCMP1168 reveals lineage-specific gene loss and genome complexity. *Genome Biol Evol*. 4(11):1162–1175.

Palmer JD, Thompson WF. 1982. Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. *Cell* 29(2):537–550.

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

Puthiyaveetil S, Allen JF. 2009. Chloroplast two-component systems: evolution of the link between photosynthesis and gene expression. *Proc Biol Sci*. 276(1665):2133–2145.

San Filippo J, Lambowitz AM. 2002. Characterization of the C-terminal DNA-binding/DNA endonuclease region of a group II intron-encoded protein. *J Mol Biol*. 324(5):933–951.

Strauss SH, Palmer JD, Howe GT, Doerksen AH. 1988. Chloroplast genomes of two conifers lack a large inverted repeat and are extensively rearranged. *Proc Natl Acad Sci U S A*. 85(11):3898–3902.

Supek F, Vlahovicek K. 2005. Comparison of codon usage measures and their applicability in prediction of microbial gene expressivity. *BMC Bioinformatics* 6:182.

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 *Prototheca*. *Sci Rep*. 8(1):940.

Tanifuji G, Kamikawa R, Moore CE, Mills T, Onodera NT, Kashiyama Y, Archibald JM, Inagaki Y, Hashimoto T. 2020. Comparative

plastid genomics of *Cryptomonas* species reveals fine-scale genomic responses to loss of photosynthesis. *Genome Biol Evol*. 85:95–127.

Turmel M, Otis C, Lemieux C. 2016. Mitochondrion-to-chloroplast dna transfers and intragenomic proliferation of chloroplast group II introns in *Gloeotilopsis* green algae (Ulotrichales, Ulvophyceae). *Genome Biol Evol*. 8(9):2789–2805.

Yamaguchi A, Yubuki N, Leander BS. 2012. Morphostasis in a novel eukaryote illuminates the evolutionary transition from phagotrophy to phototrophy: description of *Rapaza viridis* n. gen. et sp. (Euglenozoa, Euglenida). *BMC Evol Biol*. 12(1):29.

Zhong J, Lambowitz AM. 2003. Group II intron mobility using nascent strands at DNA replication forks to prime reverse transcription. *EMBO J*. 22(17):4555–4565.