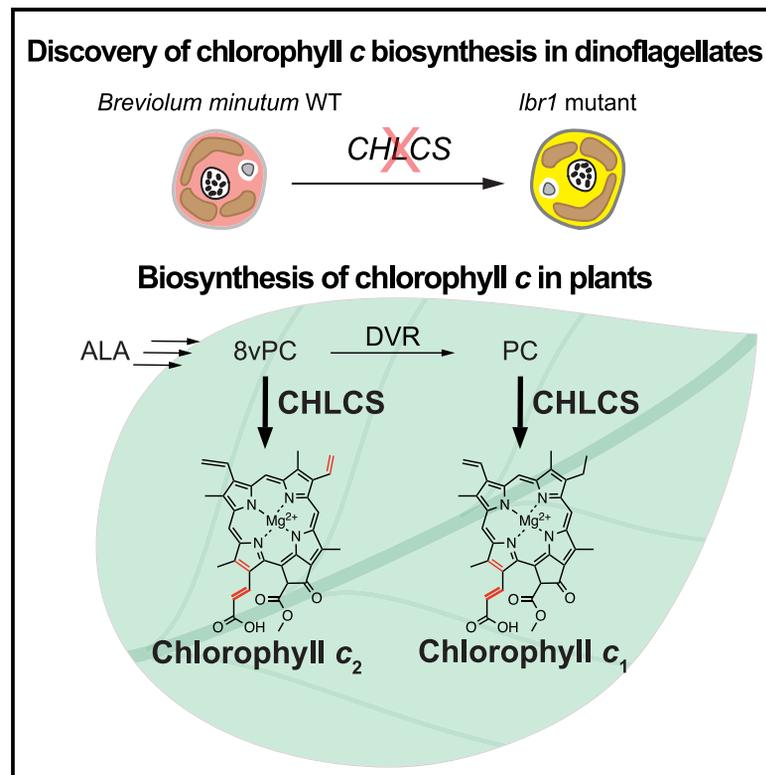


# Biosynthesis of chlorophyll *c* in a dinoflagellate and heterologous production *in planta*

## Graphical abstract



## Authors

Robert E. Jinkerson,  
Daniel Poveda-Huertes,  
Elizabeth C. Cooney, ...,  
Patrick J. Keeling, Tingting Xiang,  
Johan Andersen-Ranberg

## Correspondence

robert.jinkerson@ucr.edu (R.E.J.),  
txiang@ucr.edu (T.X.)

## In brief

Jinkerson et al. discover CHLOROPHYLL C SYNTHASE (CHLCS), a multidomain enzyme in dinoflagellates catalyzing chlorophyll *c* biosynthesis. The 2OG dioxygenase domain is essential for chlorophyll *c* catalysis. Expressing CHLCS in plants leads to chlorophyll *c*<sub>1</sub> and *c*<sub>2</sub> production, offering avenues for enhancing plant pigment profiles with algal traits.

## Highlights

- CHLCS is a multidomain enzyme that produces chlorophyll *c* in dinoflagellates
- 2OG dioxygenase domain is essential for chlorophyll *c* catalysis
- Phylogeny reveals CHLCS evolution patterns in chlorophyll *c*-containing algae
- CHLCS expression in plants results in chlorophyll *c*<sub>1</sub> and *c*<sub>2</sub> accumulation



Article

# Biosynthesis of chlorophyll *c* in a dinoflagellate and heterologous production *in planta*

Robert E. Jinkerson,<sup>1,2,6,7,9,\*</sup> Daniel Poveda-Huertes,<sup>3,6</sup> Elizabeth C. Cooney,<sup>4</sup> Anna Cho,<sup>4</sup> Rocio Ochoa-Fernandez,<sup>3</sup> Patrick J. Keeling,<sup>4</sup> Tingting Xiang,<sup>5,8,\*</sup> and Johan Andersen-Ranberg<sup>3</sup>

<sup>1</sup>Department of Chemical and Environmental Engineering, University of California, Riverside, Riverside, CA 92521, USA

<sup>2</sup>Center for Plant Cell Biology, Department of Botany and Plant Sciences, University of California, Riverside, Riverside, CA 92521, USA

<sup>3</sup>Department of Plant and Environmental Science, University of Copenhagen, 1871 Frederiksberg C, Denmark

<sup>4</sup>Department of Botany, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

<sup>5</sup>Department of Bioengineering, University of California, Riverside, Riverside, CA 92521, USA

<sup>6</sup>These authors contributed equally

<sup>7</sup>X (formerly Twitter): @JinkersonLab

<sup>8</sup>X (formerly Twitter): @Xiang\_lab

<sup>9</sup>Lead contact

\*Correspondence: robert.jinkerson@ucr.edu (R.E.J.), txiang@ucr.edu (T.X.)

<https://doi.org/10.1016/j.cub.2023.12.068>

## SUMMARY

Chlorophyll *c* is a key photosynthetic pigment that has been used historically to classify eukaryotic algae. Despite its importance in global photosynthetic productivity, the pathway for its biosynthesis has remained elusive. Here we define the *CHLOROPHYLL C SYNTHASE (CHLCS)* discovered through investigation of a dinoflagellate mutant deficient in chlorophyll *c*. *CHLCSs* are proteins with chlorophyll *a/b* binding and 2-oxoglutarate-Fe(II) dioxygenase (2OGD) domains found in peridinin-containing dinoflagellates; other chlorophyll *c*-containing algae utilize enzymes with only the 2OGD domain or an unknown synthase to produce chlorophyll *c*. 2OGD-containing synthases across dinoflagellate, diatom, cryptophyte, and haptophyte lineages form a monophyletic group, 8 members of which were also shown to produce chlorophyll *c*. Chlorophyll *c*<sub>1</sub> to *c*<sub>2</sub> ratios in marine algae are dictated in part by chlorophyll *c* synthases. *CHLCS* heterologously expressed *in planta* results in the accumulation of chlorophyll *c*<sub>1</sub> and *c*<sub>2</sub>, demonstrating a path to augment plant pigment composition with algal counterparts.

## INTRODUCTION

Photosynthetic organisms have evolved a diverse array of pigments to effectively capture light across a wide range of wavelengths. Chlorophyll *c* is a blue-green pigment found in diatoms, brown algae, dinoflagellates, and many other lineages of eukaryotic algae. These chlorophyll *c*-containing algae perform a major fraction of primary production in marine ecosystems.<sup>1,2</sup> The presence of chlorophyll *c* has historically been used to classify and infer the evolutionary relationships of heterokont, haptophyte, cryptophyte, and dinoflagellate algae, making chlorophyll *c* a defining characteristic of the proposed kingdom Chromista.<sup>3</sup> Despite the ecological and evolutionary significance of chlorophyll *c*, it is the last major chlorophyll whose biosynthetic pathway has not been elucidated.<sup>4</sup> Here we report the discovery of a protein in dinoflagellates that catalyzes the biosynthesis of chlorophyll *c*<sub>1</sub> and *c*<sub>2</sub>.

previously isolated a series of *B. minutum* mutants with altered color phenotypes. One pale yellow mutant, *LESS BROWN 1 (lbr1)*, could not grow photoautotrophically, had disrupted photosynthetic function,<sup>6</sup> and exhibited less absorbance in the blue-green region (Figures 1A and S1A). Evaluation of the pigment composition of *lbr1* using ultra-high-performance liquid chromatography-high-resolution mass spectrometry (UHPLC-HRMS) revealed that it does not produce chlorophyll *c*<sub>2</sub> (Figures 1B and S1B). To identify the causative mutation responsible for the loss of chlorophyll *c* phenotype, we sequenced the transcriptome of *lbr1*. Analysis of sequencing data revealed at least 79 mutations identifiable in the transcriptome (Figure S2A; Table S1). One of the mutations, a single nucleotide deletion, caused a frameshift in a gene (s6\_3623) encoding a peptide with a predicted chloroplast targeting sequence and putative chlorophyll *a/b* binding and 2 oxoglutarate-Fe(II) dioxygenase (2OGD) superfamily domains (Figures 1C, 1D, and S2).

## RESULTS

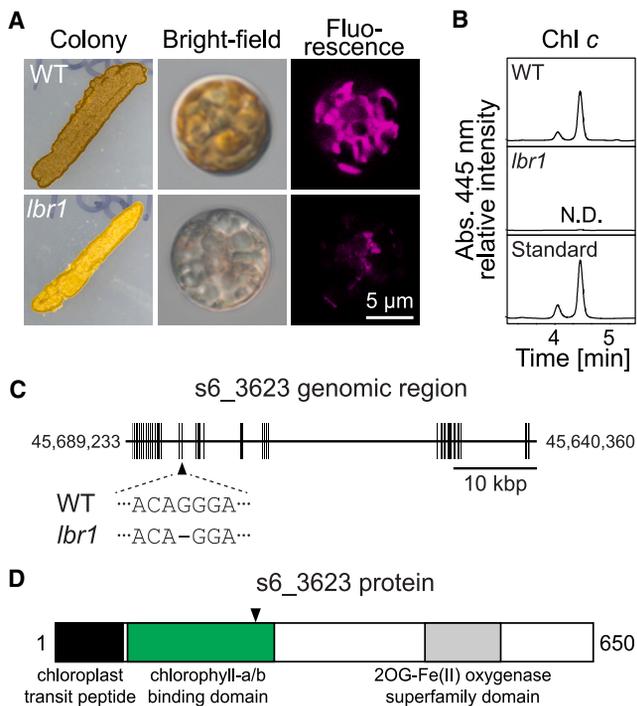
### *Breviolum* mutant deficient in chlorophyll *c* production

*Breviolum minutum* is a chlorophyll *c*<sub>2</sub>-containing dinoflagellate alga<sup>5</sup> that is brown in color. In a mutagenesis screen, we

### Heterologous *in planta* biosynthesis of chlorophyll *c*

To elucidate the potential role of s6\_3623 in chlorophyll *c*<sub>2</sub> biogenesis, we initially considered genetic complementation approaches. However, due to the current inability to achieve





**Figure 1. *B. minutum lbr1* mutant is deficient in chlorophyll c production**

(A) *Breviolum minutum* wild type (WT) and *lbr1* mutant grown heterotrophically on solid agar media (left). Representative microscopy images of WT and *lbr1* mutant cells showing bright-field (middle) and chlorophyll autofluorescence (right). Scale bar, 5  $\mu$ m.

(B) UHPLC-DAD analysis of pigment extracts from WT and *lbr1* mutant and authentic standard of chlorophyll  $c_2$ . Identity of chlorophyll  $c_2$  was confirmed by high-resolution mass spectrometry and authentic standards (Figure S1). N.D., not detected.

(C) Deletion site in *lbr1* mutant indicated as black solid triangle in the genomic region of s6\_3623 with exons shown in black.

(D) The primary structure of the s6\_3623 protein divided into the predicted chloroplast transit peptide (cTP, black), chlorophyll-a/b binding domain (Chl\_a-b, green), and 2 oxoglutarate-Fe(II) dioxygenase superfamily domain (2OGD, gray). Arrow indicates where the protein is altered in *lbr1* due to the frameshift.

See also Figures S1 and S2 and Table S1.

stable genetic transformation in *B. minutum*,<sup>6,7</sup> we employed a heterologous plant expression system as an alternative. We transiently expressed a codon-optimized version of the gene in *Nicotiana benthamiana* (Figure S3A; Table S2).<sup>8,9</sup> Subsequent UHPLC-HRMS analysis of pigment extracts from *N. benthamiana* leaves expressing the s6\_3623 coding sequence revealed a very small peak displaying an absorption spectrum characteristic of porphyrin-type chlorophyll, which was absent in the negative control (Figure 2A). The new peak had the same retention time and absorption spectra similar to chlorophyll  $c_1$  and chlorophyll  $c_2$  authentic standards, which coelute during our chromatographic separation (Figure S1). This correspondence in retention time and spectral properties indicates that the peak represents chlorophyll c. The detection of chlorophyll c supported that s6\_3623 encodes an enzyme capable of catalyzing the formation of chlorophyll c. We therefore designated the gene linked to the s6\_3623 transcript as *CHLOROPHYLL C*

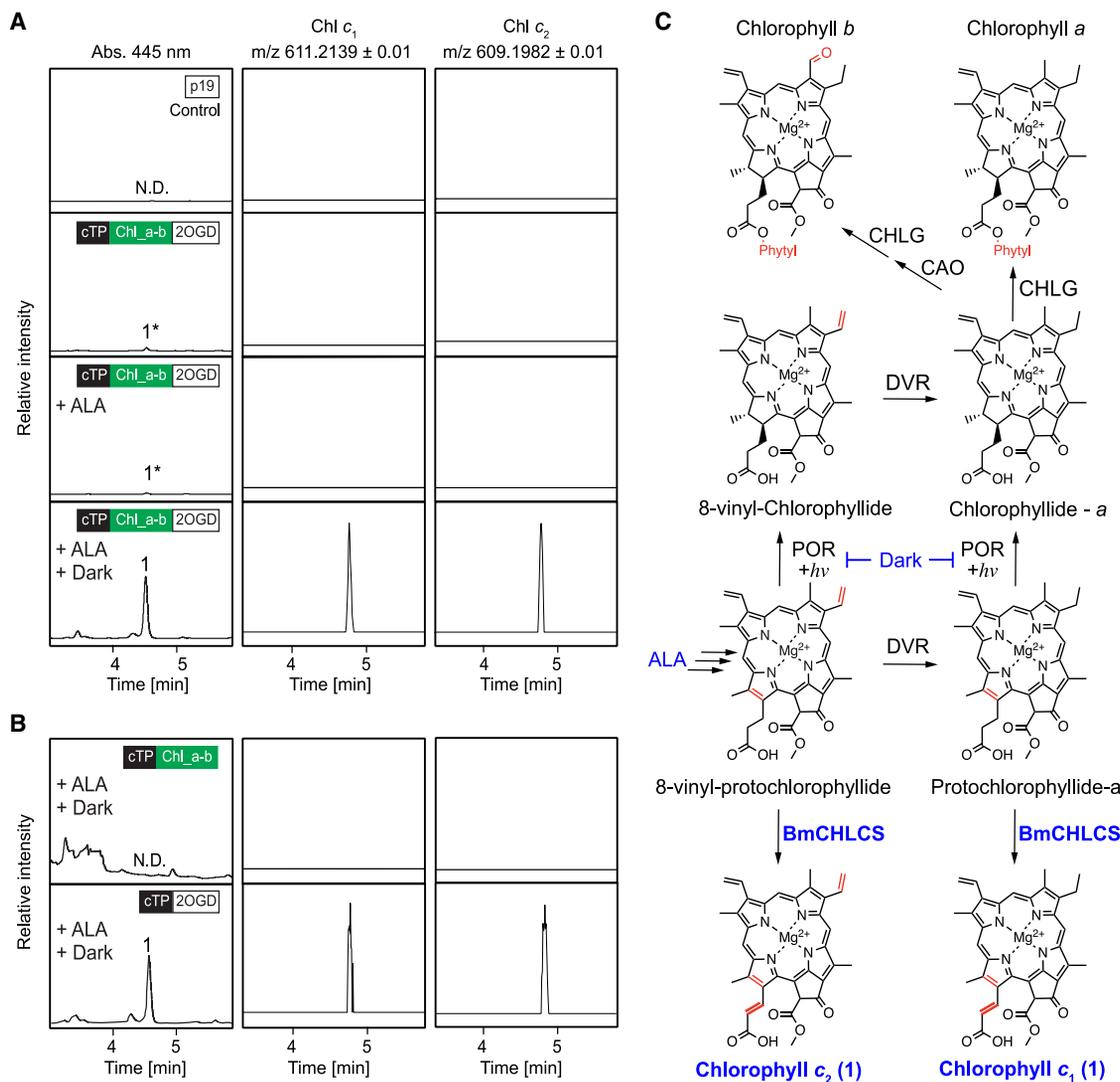
*SYNTHASE (BmCHLCS)*. This represents the first documented case of heterologous biosynthesis of chlorophyll c in a phototrophic organism and the inaugural instance of a non-plant accessory pigment being synthesized *in planta*.

### 8-vinyl-protochlorophyllide a and protochlorophyllide a are substrates for BmCHLCS

Because only trace amounts of chlorophyll c were found in the *N. benthamiana* extracts, we postulated that this limited accumulation might be attributable to the lack of substrate availability for the BmCHLCS enzyme. It has been hypothesized that 8-vinyl-protochlorophyllide a (8vPC) and protochlorophyllide a (PC) are precursors for chlorophyll  $c_2$  and chlorophyll  $c_1$ , respectively.<sup>10,11</sup> To increase the concentration of 8vPC and PC, we immersed *N. benthamiana* leaf discs in a buffer supplemented with the protochlorophyllide precursor  $\delta$ -aminolevulinic acid (ALA)<sup>12</sup> (Figure S3B). This, however, did not enhance chlorophyll c production (Figure 2A). It is notable that both 8vPC and PC are substrates for the protochlorophyllide oxidoreductase (POR) *in planta*, catalyzing their conversion to 8-vinylchlorophyllide and chlorophyllide a, respectively<sup>13</sup> (Figure 2C). POR is a light-dependent enzyme, so in addition to supplementing with ALA, *N. benthamiana* leaf discs expressing BmCHLCS were also placed in the dark to facilitate accumulation of 8vPC and PC. This dual intervention led to a pronounced increase in the chlorophyll c peak (Figure 2A), suggesting that substrate availability *in planta* directly impacts the ability of BmCHLCS to catalyze the formation of chlorophyll c. This increase also enabled MS detection of adducts with mass-to-charge ratios of  $m/z$  611.2139 and  $m/z$  609.1982. This aligns with the pseudo molecular ions  $[M + H]^+$  of chlorophyll  $c_1$  and chlorophyll  $c_2$ , respectively, with their identities verified against authentic standards (Figures S1D and S1E). Notably, whereas only chlorophyll  $c_2$  is observed in *B. minutum* wild type (WT) (Figure S1C), *in planta* expression of BmCHLCS results in both chlorophyll  $c_1$  and  $c_2$  production. Overall, these results indicate that BmCHLCS operates as a light-independent enzyme and, within the context of *N. benthamiana*, utilizes 8vPC and possibly PC as substrates for synthesizing chlorophyll  $c_1$  and chlorophyll  $c_2$ , respectively.

### 2OG dioxygenase domain is the catalytic domain that produces chlorophyll c

BmCHLCS encodes two distinct domains: a putative chlorophyll a/b binding domain<sup>14</sup> and a domain belonging to the 2-oxoglutarate-Fe(II) dioxygenase (2OGD) superfamily domain.<sup>15</sup> However, it was unclear whether the chlorophyll a/b binding or the 2OGD domain, if not both, is essential for BmCHLCS catalytic activity. To resolve this ambiguity, we selectively deleted each domain to assess their individual functional contributions to chlorophyll c production in *N. benthamiana*. When the 2OGD domain was deleted, no chlorophyll  $c_1$  or  $c_2$  was detected in extracts from leaves (Figure 2B). In contrast, deletion of the chlorophyll a/b binding domain had no effect on chlorophyll  $c_1$  and  $c_2$  production, with levels produced similar to those in the full length BmCHLCS (Figure 2B). These results indicate that, *in planta*, the 2OGD domain functions as the catalytic site for chlorophyll  $c_1$  and  $c_2$  biosynthesis, while the chlorophyll a/b binding domain is non-essential.



**Figure 2. Heterologous biosynthesis of chlorophyll *c* in planta requires the 2OG dioxygenase domain**

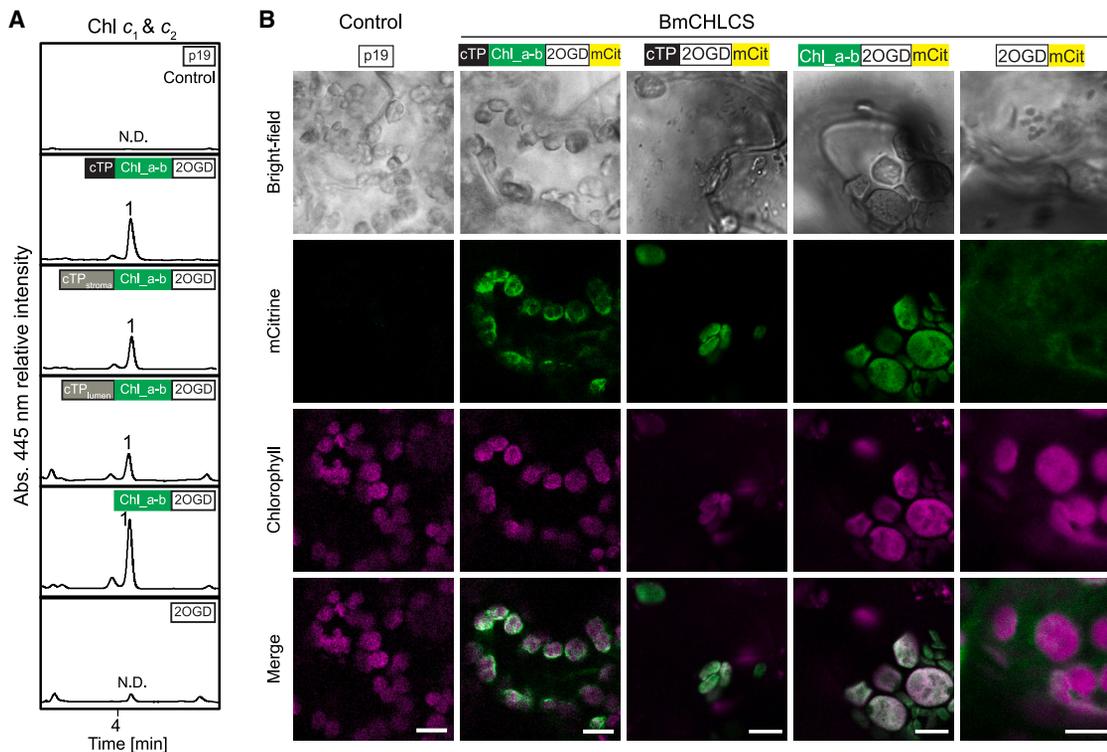
(A and B) UHPLC-HRMS analysis of pigment extracts from *Nicotiana benthamiana* leaf disks expressing constructs indicated in the insets. For each sample, absorbance at 445 nm and the corresponding extracted ion chromatogram for chlorophyll  $c_1$  and  $c_2$  are shown. Chlorophyll  $c$  peak indicated with (1); trace amount detected was indicated as asterisk (\*); not detected as N.D. Leaf disk incubation conditions are indicated in plots, such as with  $\delta$ -aminolevulinic acid (ALA) or in the dark. The empty vector p19 served as a control.

(C) Chlorophyll biosynthesis pathway in *N. benthamiana* with BmCHLCS expression. Chemical changes driven by BmCHLCS, protochlorophyllide oxidoreductase (POR), divinyl reductase (DVR), chlorophyll synthase (CHLG), and chlorophyllide *a* oxygenase (CAO) are highlighted in red. POR is a light-dependent enzyme. Interventions made to the pathway are indicated in blue. See also Figure S3 and Table S2.

### Chloroplast localization of BmCHLCS is required for chlorophyll *c* biosynthesis

Subcellular localization of BmCHLCS is likely foundational to its functional role, especially as the late stages of chlorophyll biosynthesis occur within the chloroplast in plants. In contrast to plants, marine eukaryotic algae derived from a secondary or more endosymbiotic event(s), including dinoflagellates, utilize a bipartite N-terminal pre-sequence consisting of a signal peptide followed by a chloroplast transit peptide to facilitate protein import into the chloroplast.<sup>16</sup> Our predictive analyses, utilizing HECTAR<sup>17</sup> and DeepLoc 2.0,<sup>18</sup> identified a signal peptide followed by chloroplast transit peptide at BmCHLCS's N terminus. At the C terminus of the predicted chloroplast transit

peptide, a cleavage motif (AXA)<sup>19</sup> was observed (Figure S2B). Remarkably, BmCHLCS, even with its native dinoflagellate bipartite pre-sequence, can synthesize chlorophyll *c* in *N. benthamiana* (Figures 2A and 3A), suggesting the dinoflagellate bipartite pre-sequence can facilitate chloroplast localization in *planta*. To determine if BmCHLCS is actually localized to the chloroplast in *N. benthamiana*, we expressed BmCHLCS fused with the yellow fluorescent protein mCitrine.<sup>20</sup> Epidermal cells imaged with confocal microscopy revealed localization of mCitrine fluorescence within chloroplasts (Figure 3B). This was further corroborated by identifying BmCHLCS in isolated chloroplasts (Figure S3C). This information prompted us to experiment with transit peptide modifications, revealing that even



**Figure 3. *In planta* accumulation of chlorophyll  $c_1$  and  $c_2$  is dependent on chloroplast localization of BmCHLCS**

(A) UHPLC-HRMS analysis of pigment extracts from *N. benthamiana* leaf disks expressing constructs indicated in the insets. For each sample, absorbance at 445 nm is shown. Chlorophyll  $c$  peak indicated with (1); not detected as N.D. Leaf disks were incubated with  $\delta$ -aminolevulinic acid (ALA) and in the dark. The empty vector p19 served as a control. Black box represents the native *B. minutum* chloroplast transit peptide (cTP) sequence. Gray boxes indicate chloroplast transit peptide (cTP) of plant origin.

(B) Representative confocal microscopy images of *N. benthamiana* epidermal cells expressing full-length or truncated constructs of BmCHLCS:mCitrine, as indicated at top. Black box represents the native *B. minutum* chloroplast transit peptide (cTP) sequence. Green, mCitrine fluorescence; magenta, chlorophyll fluorescence. The empty vector p19 was used as a control. Scale bars, 10  $\mu$ m.

See also Table S4.

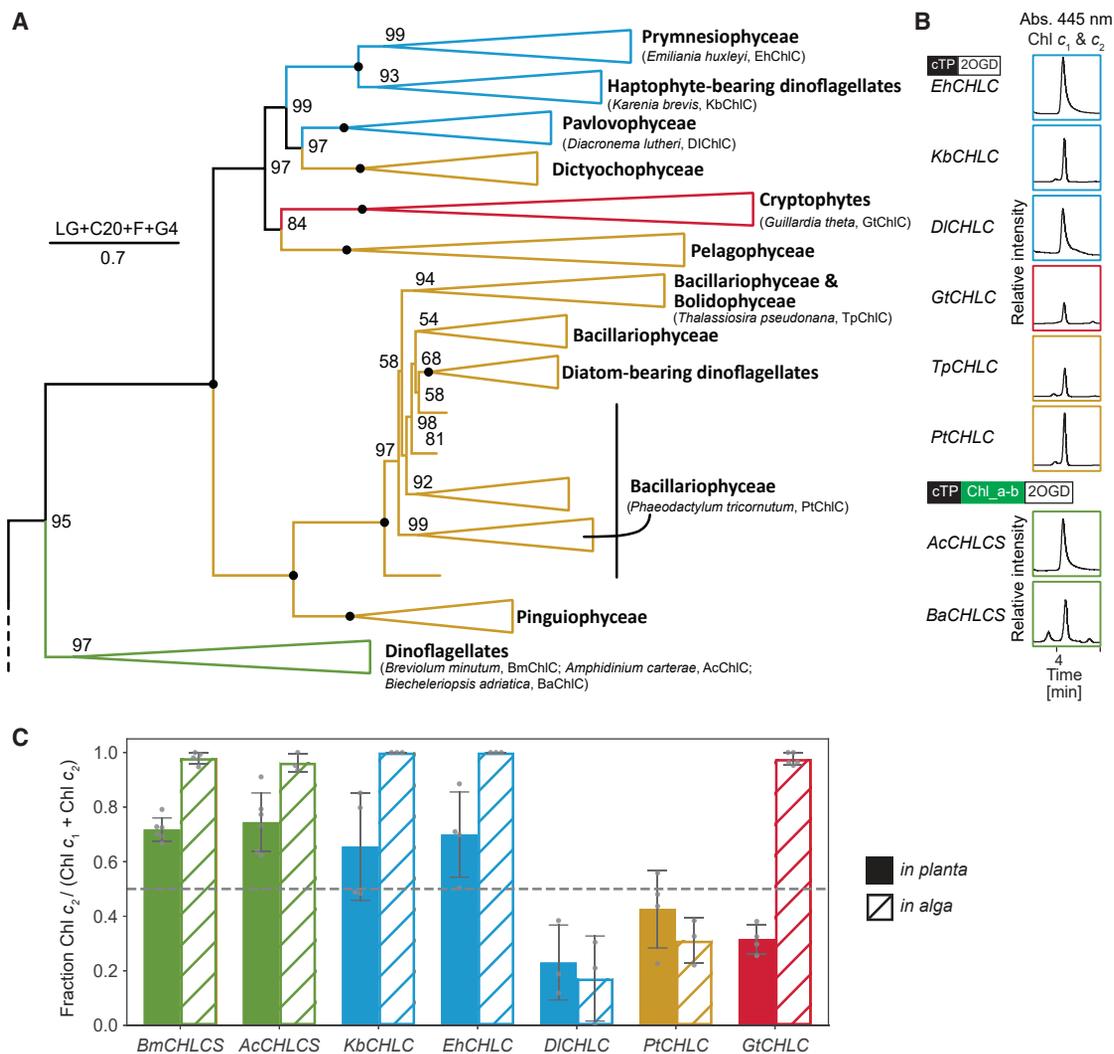
when swapping the native dinoflagellate bipartite pre-sequence with plant-specific stromal and luminal targeting sequences,<sup>21</sup> chlorophyll  $c$  was still produced (Figure 3A). Intriguingly, the omission of the dinoflagellate bipartite pre-sequence did not impede either chlorophyll  $c$  production or BmCHLCS's chloroplast localization. However, stripping both the bipartite pre-sequence and the chlorophyll  $a/b$  binding domain abolished chloroplast localization and chlorophyll  $c$  production (Figures 3A and 3B), solidifying that appropriate subcellular positioning of BmCHLCS is imperative for chlorophyll  $c$  synthesis. Fusion of the bipartite pre-sequence and the 2OGD domain also leads to chlorophyll  $c$  accumulation, showing that the bipartite pre-sequence facilitates chloroplast translocation independently from the chlorophyll  $a/b$  binding domain (Figure 3A). In summary, BmCHLCS is directed to the chloroplast *in planta* via the dinoflagellate bipartite pre-sequence, and BmCHLCS localization to the chloroplast is essential for chlorophyll  $c$  biosynthesis.

### Chlorophyll $c$ biosynthesis across algal lineages

To gain insights into chlorophyll  $c$  biosynthesis across eukaryotic algae, we sought to examine CHLCS distribution and function in species harboring this pigment. Using BLAST to search data

from the Marine Microbial Eukaryote Transcriptome Sequencing Project,<sup>22</sup> we identified transcripts that share sequence similarity to the full-length CHLCS in all peridinin-containing dinoflagellate species available (Figure 4A; Data S1; Table S3). To determine their chlorophyll  $c$  production capability, we chose transcripts from the extensively researched dinoflagellate, *Amphidinium carterae*, and *Biecheleriopsis adriatica*, which is deemed a basal lineage in this dinoflagellate clade.<sup>23,24</sup> These were then codon-optimized and expressed in *N. benthamiana*. Both genes proved successful in catalyzing the production of chlorophyll  $c_1$  and  $c_2$  (Figures 4B and S4A). A broader phylogenetic scan<sup>25</sup> revealed a conspicuous absence of genes outside of the peridinin-containing dinoflagellates that simultaneously harbor both the chlorophyll  $a/b$  binding and 2OGD domains (Figure 4A; Data S1).

Significantly, however, all chlorophyll  $c$ -dependent algal lineages were found to encode 2OGD-containing proteins, and in phylogenetic analyses these formed a monophyletic group with strong support (Figure 4A; Data S1; Table S3). We evaluated the function of genes harboring the 2OGD domain across major algal lineages, given that the 2OGD domain of BmCHLCS alone can drive the biosynthesis of chlorophyll  $c$ . Genes were selected from representatives of major chlorophyll  $c$ -containing algal groups (*Karenia brevis* [Kareniaceae],



**Figure 4. Chlorophyll c biosynthesis across diverse algal lineages is catalyzed by genes with 2OGD domains**

(A) Collapsed phylogenetic tree of 2-oxoglutarate dioxygenase (2OGD) domains identified in chlorophyll c-containing eukaryotic algae. Complete phylogeny can be found in [Data S1](#). Scale bar indicates the estimated number of amino acid substitutions per site and shows the model used to infer the tree.

(B) UHPLC-DAD chromatograms of *N. benthamiana* extracts expressing *CHLCS* or *CHLC* from select species, representing the groups depicted in the phylogeny. Species include *Emiliana huxleyi* (Eh), *Diacronema lutheri* (Dl), *Guillardia theta* (Gt), *Thalassiosira pseudonana* (Tp), *Phaeodactylum tricornutum* (Pt), *Amphidinium carterae* (Ac), and *Biecheleriopsis adriatica* (Ba). Algal groups indicated by color: haptophyte and haptophyte-bearing dinoflagellate (blue), Cryptophyte (red), Diatomista and diatom-bearing dinoflagellates (brown), and peridinin-containing dinoflagellate (green). For each sample absorbance at 445 nm is shown.

(C) Level of chlorophyll  $c_2$  as a fraction of total chlorophyll c (chlorophyll  $c_1$  + chlorophyll  $c_2$ ) in extracts of *N. benthamiana* expressing the selected *CHLCS*s and *CHLC*s and in extracts from algae grown in culture. Fraction was based on peak areas from extracted ion chromatograms specific for chlorophyll  $c_2$  ( $m/z$  609.1982  $\pm$  0.01) and chlorophyll  $c_1$  ( $m/z$  611.2139  $\pm$  0.01) ([Figure S4A](#)). Dotted line indicates equal levels of chlorophyll  $c_2$  and chlorophyll  $c_1$ . Error bars indicate standard deviation and gray dot individual replicates ( $n = 3, 4, \text{ or } 6$ ). Bar color indicates algal group within condensed phylogeny. See also [Figure S4](#), [Tables S2](#) and [S3](#), and [Data S1](#).

*Emiliana huxleyi* [Haptophyceae], *Diacronema lutheri* [Haptophyceae], *Phaeodactylum tricornutum* [Diatomista], *Thalassiosira pseudonana* [Diatomista], and *Guillardia theta* [Cryptophyceae], codon-optimized, and expressed in *N. benthamiana*. Both chlorophyll  $c_1$  and chlorophyll  $c_2$  were detected in the *N. benthamiana* assays for all genes evaluated; thus, we refer to these genes with only a 2OGD domain as *CHLC* ([Figures 4B](#) and [S4A](#)). In total, we show that genes similar to *BmCHLCS* that possess both the chlorophyll *a/b* binding and 2OGD domain

and those that only possess the 2OG dioxygenase domain, *CHLC*, catalyze the biosynthesis of chlorophyll c across diverse algal lineages.

These algae all acquired their secondary plastids from red algae, so we searched all available genomic data from red algae but recovered no *CHLCS* homologs. This suggests that the 2OGD domain in *CHLCS* most likely arose after the acquisition of the red algal plastid and not from the red algal ancestor of the plastid.

### Chlorophyll *c* synthases influence the ratio of chlorophyll *c*<sub>1</sub> to *c*<sub>2</sub>

During the *in planta* expression of these three *CHLCS* and six *CHLC* genes, we observed variations in the ratio of chlorophyll *c*<sub>1</sub> to chlorophyll *c*<sub>2</sub> produced across species. Given that all the heterologous expression was conducted within a constant host, *N. benthamiana*, the observed variability suggested there may be inherent differences among chlorophyll *c* synthases that give rise to different chlorophyll *c*<sub>1</sub>/*c*<sub>2</sub> ratios. To evaluate if these distinct ratios observed *in planta* are naturally present in the algae that harbor these genes, pigments were extracted and analyzed from seven algal species grown in culture (Figures 4C and S4A). Comparison of both sets of data revealed a significant correlation between chlorophyll *c*<sub>1</sub>/*c*<sub>2</sub> ratios *in alga* and *in planta* (Pearson correlation coefficient > 0.60, *p* = 0.0004). For *D. lutheri* and *P. tricornutum*, the chlorophyll *c*<sub>1</sub>/*c*<sub>2</sub> ratios observed *in alga* closely matched those *in planta* and were statistically indistinguishable (*p* = 0.57, two-tailed *t* test). Pigments extracted from *B. minutum*, *A. ceterae*, *K. brevis*, and *E. huxleyi* cultures only contained chlorophyll *c*<sub>2</sub> (Figure 4C). When chlorophyll *c* synthase genes from these four species were expressed *in planta*, chlorophyll *c*<sub>2</sub> was primarily produced. However, in a departure from what is observed *in alga*, chlorophyll *c*<sub>1</sub> was also detected (Figure 4C). Interestingly, all four pigment extracts contained similar levels of chlorophyll *c*<sub>1</sub>, ~29% of the total chlorophyll *c* (*p* = 0.74, one-way ANOVA). *G. theta*, however, diverged from this trend (Figure 4C). These data collectively imply that observed variations in chlorophyll *c*<sub>1</sub>/*c*<sub>2</sub> ratios among species appear to be influenced, in part, by enzymatic differences in their specific chlorophyll *c* synthases.

### CHLC has been lost in a subgroup of ochrophytes

The one exception to the strict correlation between the presence of chlorophyll *c* and *CHLCS* or *CHLC* homologs is found in a subset of ochrophytes. Ochrophytes encompass many subgroups and include both photosynthetic and non-photosynthetic species. One major subgroup is the Diatomista,<sup>26</sup> which are chlorophyll *c* dependent<sup>27</sup> and encode *CHLC* (Figure 5A). The other major subgroup includes a diversity of both photosynthetic and non-photosynthetic lineages; some photosynthetic lineages are chlorophyll *c* dependent, while others depend on chlorophyll *a* alone. No *CHLC* homolog was identified in any genome or transcriptome from any of these species, regardless of their pigment composition. This most likely suggests that alternative chlorophyll *c* biosynthetic enzyme(s) evolved in this lineage, and based on the scattered presence of chlorophyll *c*<sub>1</sub> and *c*<sub>2</sub> in this group (Figure 5A), that this pathway was subsequently lost in the Euglostomatophyte and Xanthophyte lineages (Figure 5A). Interestingly, in ochrophytes the presence of chlorophyll *c*<sub>1</sub> and *c*<sub>2</sub> also correlates with the presence of the secondary pigment fucoxanthin. Recently, the *ZEP1*, *VDL2*, and *CRTISO5* genes were all shown to be essential for fucoxanthin biosynthesis in Diatomista and Haptophyceae,<sup>28,29</sup> but we now find that these genes are also lacking in the same subgroup of ochrophytes lacking *CHLC* (Figure 5A), suggesting an even more widespread remodeling of their pigment biosynthesis pathways took place after their divergence from Diatomista.

## DISCUSSION

### Linking phenotype to genotype in a dinoflagellate

Dinoflagellates are unique organisms that have had a non-conventional evolutionary trajectory and occupy key ecological niches.<sup>34–37</sup> The lack of genetic tools to probe their distinct biology has limited our ability to make discoveries about these organisms. Recently we have shown that UV exposure can generate mutants with a variety of phenotypes in the dinoflagellate *B. minutum*.<sup>6</sup> However, this advance is limited without the ability to link a mutant phenotype to a genotype, something commonly done with other organisms, but not yet demonstrated in dinoflagellates. Here we show how transcriptome sequencing coupled with heterologous expression can be used to determine gene function and link a mutant genotype to a phenotype in a dinoflagellate. Achieving this milestone will hopefully mark the beginning of a new era of biological discovery in dinoflagellates, spanning from elucidation of novel biosynthetic pathways to deepening our understanding of coral-dinoflagellate symbiosis.

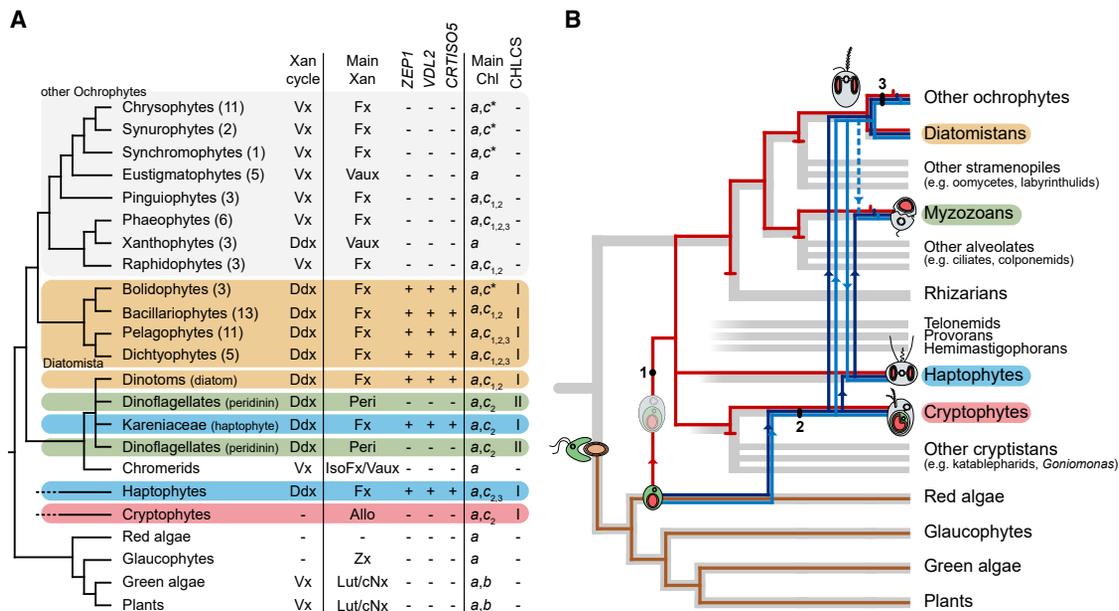
### Chemistry of chlorophyll *c* biosynthesis

The 2OGD domain is the catalytic domain of chlorophyll *c* synthase. This domain shares weak homology (E value: 2.02e–05 for *BmCHLCS*, NCBI Conserved Domain Database) to a 2OGD domain from phytyl-CoA dioxygenase enzymes (EC 1.14.11.18).<sup>15</sup> 2OGDs have been identified in all kingdoms of life and are commonly involved in catalyzing diverse oxidative reactions across nature, such as hydroxylation of aliphatic or aromatic C-H bonds, forming or cleaving C-C bonds, demethylation, ring formation or cleavage, and desaturation.<sup>15,38–40</sup> Formation of chlorophyll *c*<sub>2</sub> and chlorophyll *c*<sub>1</sub> from 8vPC and PC, respectively, requires a desaturation of C17<sup>1</sup>-C17<sup>2</sup> (Figures 2C and S1G). Desaturation of aliphatic carbons with 2OGD typically requires an adjacent heteroatom,<sup>41</sup> although none are present in 8vPC or PC. While examples are limited, 2OGD-catalyzed desaturation independent of the presence of adjacent heteroatoms has been reported for a gibberellin oxidase/desaturase (*CsGA1ox/ds*) involved in plant hormone biosynthesis.<sup>42</sup> This discovery that the 2OGD domain from *CHLCS* catalyzes the requisite desaturation to form chlorophyll *c* broadens our understanding of 2OGD enzymatic versatility. This atypical route for a biochemical desaturation offers prospects to delve deeper into the enzyme's mechanism and holds promise for a single-step enzymatic reaction to desaturate aliphatic molecules, a process with broad potential in green chemistry.

### Biosynthesis of chlorophyll *c*

Expression of *BmCHLCS* in *N. benthamiana* is the first heterologous production of chlorophyll *c* and allows for investigation of the biology and chemistry of chlorophyll *c* biosynthesis.

Chlorophyll *c* has been known since 1864,<sup>43</sup> but not until 1968 was it determined that the observed “chlorophyll *c*” peak was actually made up of both chlorophyll *c*<sub>1</sub> and *c*<sub>2</sub>, which have different spectral properties.<sup>44</sup> At that time, it was also shown that different species of algae possess different ratios of chlorophyll *c*<sub>1</sub> and *c*<sub>2</sub>; for example, both chlorophylls *c*<sub>1</sub> and *c*<sub>2</sub> are often present in brown algae and diatoms while only chlorophyll *c*<sub>2</sub> is typically found in dinoflagellates and cryptomonads.<sup>5,10,45</sup> Algae



**Figure 5. Insights into the evolution of chlorophyll c-containing algae from CHLCS and CHLC**

(A) Photosynthetic eukaryotic phylum/class phylogeny. Each phylum/class displays number of species in each ochrophyte lineage with available genomes or transcriptomes (in brackets), plastid type in dinoflagellates (in brackets), xanthophyll cycle type (Xan cycle), primary light-harvesting xanthophyll(s) (Main Xan), fucoxanthin biosynthesis gene presence, chlorophyll composition (Main Chl), and CHLCS homolog presence. Vx, violaxanthin-dependent cycle; Ddx, diadinoxanthin-dependent cycle; Fx, fucoxanthin; Vaux, vaucheriaxanthin; Peri, peridinin; IsoFx, isofucoxanthin; Allo, alloxanthin; Lut, lutein; cNx, *cis*-neoxanthin; ZEP1, type 1 zeaxanthin epoxidase; VDL2, type 2 violaxanthin deepoxidase-like enzyme; CRTISO5, type 5 carotenoid isomerase; a, chlorophyll a; c<sub>1</sub>, chlorophyll c<sub>1</sub>; c<sub>2</sub>, chlorophyll c<sub>2</sub>; c<sub>3</sub>, chlorophyll c<sub>3</sub>; b, chlorophyll b. Xanthophyll and chlorophyll compositions are sourced from Bai et al.<sup>28,29</sup> and referenced literature. Chlorophyll c<sup>\*</sup> refers to studies where chlorophyll c<sub>1</sub> and c<sub>2</sub> were not differentiated. For CHLCS: I, only CHLC domain present; II, both chlorophyll a/b binding domain and 2OGD domain are present.

(B) Three evolutionary scenarios consistent with the single origin of chlorophyll c synthase. The gray cladogram is a generalized phylogeny of major lineages. Unresolved relationships are shown as disconnected branches with a faded root. Color-highlighted lineage names carry the chlorophyll c synthases described in this study. The brown line shows the proliferation of primary plastids. The red line depicts the scenario proposed by the chromalveolate hypothesis<sup>3</sup> while blue lines depict two possible rhodoplex hypothesis scenarios that are consistent with a timescale analysis performed by Strasser et al.<sup>30</sup>: dark blue corresponds to Bodyl et al.,<sup>31</sup> light blue represents Stiller et al.,<sup>32</sup> and dashed light blue represents Ševčíková et al.<sup>33</sup> Arrows show the direction of new plastid acquisitions between lineages. The distribution of chlorophyll c synthases is consistent with all three scenarios of a single origin of red algal plastids, but not with any scenario of multiple acquisitions. Hatches terminating red lines indicate cases of plastid loss for an entire lineage while half-hatches represent cases in which some but not all members of a lineage have undergone plastid loss. The black markers at 1 and 2 indicate when chlorophyll c synthase would have arisen in the chromalveolate and rhodoplex hypotheses, respectively. Marker 3 indicates where, in all cases, chlorophyll c synthase must have been lost in the lineage of ochrophytes sister to the diatomists.

may regulate the ratio of chlorophyll c<sub>1</sub> to c<sub>2</sub> through various mechanisms, including chlorophyll c synthase substrate specificity or activity, substrate availability or localization, or by modulating the specificity/activity of other enzymes involved in chlorophyll biosynthesis, such as DVR and POR. In the heterologous plant system we used, all these factors remain consistent, except for the chlorophyll c synthases, facilitating a direct comparison of chlorophyll c synthases from diverse algal species. Expression of these interspecific synthases from different algal groups resulted in a variety of chlorophyll c<sub>1</sub>/c<sub>2</sub> ratios. Intriguingly, the ratios documented *in planta* frequently mirrored those observed in the native algal species from which the enzymes originated, suggesting that intrinsic enzymatic differences in chlorophyll c synthases play a role in dictating c<sub>1</sub>/c<sub>2</sub> balance. The heterologous plant system, however, does not perfectly replicate *in alga* conditions. When chlorophyll c synthases are expressed from dinoflagellates and haptophytes that typically only possess chlorophyll c<sub>2</sub>, *in planta* chlorophyll c<sub>1</sub> is observed. This implies a unique aspect of the heterologous environment

that is different from the chlorophyll c<sub>2</sub>-producing algae. Potential explanations might include (1) the presence of protochlorophyllide a *in planta* at levels not typically available *in alga*; (2) the capability of the plant DVR, but not the algal DVR, to potentially catalyze the conversion of chlorophyll c<sub>2</sub> to c<sub>1</sub>; and/or (3) different spatial localization of the enzyme or substrates. Similar levels of chlorophyll c<sub>1</sub> production from these enzymes from different chlorophyll c<sub>2</sub>-only algae hint at a common mechanism of c<sub>1</sub> production in this heterologous context. The recent *in vitro* demonstration of chlorophyll c biosynthesis<sup>46</sup> could be leveraged to ascertain more detailed enzymatic parameters of these synthases. Understanding how chlorophyll c synthases can affect chlorophyll c<sub>1</sub>/c<sub>2</sub> ratio gives more insights into the molecular mechanism of how algae tune and optimize pigment composition for a given environmental niche.

As chlorophyll c is absent from plants, the heterologous biosynthesis of this pigment within a plant model system provides a unique opportunity to explore its biochemical interactions with endogenous plant enzymatic pathways. In contrast

to chlorophyll *a*, *b*, *d*, and *f*, chlorophyll *c*<sub>1</sub> and *c*<sub>2</sub> are not esterified by a phytol moiety at the C17<sup>3</sup> carbonyl (Figure S1G). Chlorophyll synthase (CHLG) esterifies chlorophyllide *a* and chlorophyllide *b* with phytol to produce chlorophyll *a* and chlorophyll *b*,<sup>47</sup> respectively, demonstrating this enzyme can function on multiple substrates,<sup>48</sup> but it was unknown if CHLG could esterify chlorophyll *c* with a phytol. Some studies report the identification of chlorophyll *c* with a phytol group isolated from marine algae samples,<sup>49</sup> so the biosynthesis of this molecule may be possible. Therefore, we employed UHPLC-HRMS to investigate the presence of phytol-esterified forms of chlorophyll *c*<sub>1</sub> or *c*<sub>2</sub>. These compounds are anticipated to have retention times similar to chlorophyll *a* and are characterized by distinct pseudomolecular ion masses with *m/z* values of 889.5029 and 887.4873, respectively. Esterified versions of chlorophyll *c*<sub>1</sub> and *c*<sub>2</sub> were not detected in the pigment extracts from *N. benthamiana* expressing *BmCHLCS* (Figure S4B), suggesting that the native enzymes of *N. benthamiana* are incapable of esterifying chlorophyll *c*<sub>1</sub> or *c*<sub>2</sub>. One possibility is that the unique C17<sup>1</sup>-C17<sup>2</sup> and C17-C18 double bonds in porphyrins of chlorophyll *c*<sub>1</sub> and *c*<sub>2</sub>, unlike chlorins, may introduce steric hindrance that restricts their esterification by the native CHLG enzyme *in planta*. Another chemical transformation that takes place during chlorophyll biosynthesis is when chlorophyllide *a* is converted to chlorophyllide *b* via chlorophyllide *a* oxygenase (CAO)<sup>50</sup> to generate an aldehyde group (–CHO) at the C-3 position. Conceivably chlorophyll *c* could serve as a substrate for this enzyme, which would produce aldehyde forms of chlorophyll *c*. However, no compounds with *m/z* corresponding to the pseudomolecular ions of aldehyde derivatives of chlorophyll *c*<sub>1</sub> (*m/z* 625.1921) and *c*<sub>2</sub> (*m/z* 623.1765) were detected (Figure S4C). Previous work with *A. thaliana* showed that CAO is specific for non-esterified chlorophyllide *a* and that chlorophyll *a* and protochlorophyllide *a* are not substrates for this enzyme.<sup>51</sup> Our results suggest that chlorophyll *c* also may not be able to serve as a substrate for CAO. In summary, chlorophyll *c* produced in *N. benthamiana* is unmodified through esterification or formylation, demonstrating the enzymatic specificity of endogenous chlorophyll biosynthetic pathways in plants, with the dominant forms produced *in planta* being chlorophyll *c*<sub>1</sub> and *c*<sub>2</sub>.

In the plant kingdom, enzymes with 2-oxoglutarate (2OG) dioxygenase domains biosynthesize a diverse array of metabolites, ranging from plant hormones to alkaloids to many other natural products.<sup>52–55</sup> While most 2OGs function in the cytosol or mitochondria,<sup>54,56,57</sup> their ability to function in the chloroplast is less clear. A survey of the *Arabidopsis thaliana* proteome (Araport11) reveals that, among the 132 enzymes predicted to possess a 2OG dioxygenase domain, only four are predicted to be localized to the chloroplast (DeepLoc 2.0). None of these four have been experimentally verified to localize to or be functional within the chloroplast (Table S4). Beyond *Arabidopsis*, 2OGs in other plant species are also overwhelmingly predicted to be cytosol and mitochondria localized, with only small fractions (~3.5%) predicted to have chloroplast localization.<sup>58</sup> Our results demonstrate that 2OG dioxygenase domain-containing proteins are functional in the chloroplast. This implies all of the requisite co-factors (Fe<sup>2+</sup>) and co-substrates (2-oxo-glutarate, ascorbic acid, etc.) are also maintained within the

chloroplast. This highlights that chloroplasts are not only a metabolic hub for photosynthesis but are also a previously unrecognized site for 2OGD-dependent biosynthetic pathways. Our functional heterologous expression of 9 chlorophyll *c* synthases from diverse algae demonstrates that the chloroplast is a promising target for engineering 2OGD-associated chemical reactions.

### Evolutionary origin of *CHLCS* and insights into plastid endosymbiosis

The recognition that plastids moved from one eukaryotic lineage to another by secondary endosymbiosis solved in one stroke a number of evolutionary puzzles surrounding the distribution of photosynthesis in the tree of eukaryotes. But it also raised new questions, and one that has proved particularly difficult is the number of secondary endosymbiosis events involving red algae. The chromalveolate hypothesis crystallized this debate by suggesting all algae with red secondary plastids evolved from a single common endosymbiosis, in part due to their unique shared use of chlorophyll *c*.<sup>3</sup> This hypothesis was supported by plastid phylogeny,<sup>59</sup> but the phylogeny of the host lineages has more recently argued against it,<sup>30</sup> and attempts have now been made to synthesize these apparently conflicting evolutionary histories with more complex models evoking a single origin from red algae in one lineage, followed by a serial spread of this plastid to other lineages by “tertiary” or even “quaternary” endosymbioses.<sup>30</sup>

Our demonstration that every lineage with red algal-derived plastids ancestrally shares orthologous chlorophyll *c* biosynthesis that is absent in red algae reinforces the conclusion that red algal-derived plastids must ultimately trace back to a single endosymbiosis, but is equally compatible with either a single origin followed by repeated plastid loss or with a single origin followed by subsequent serial endosymbioses (Figure 5B).

The absence of *CHLC* homologs in red algae may reflect poor genomic sampling of that lineage, but more likely suggests the system arose after the initial secondary endosymbiosis. This could be because *CHLC* evolved from a gene in the eukaryotic host of the secondary event, or that the 2OGD domain is derived from a horizontal gene transfer event. Indeed, many of the distant relatives of the chlorophyll *c*-producing 2OGD domain are bacterial homologs of other 2OG dioxygenase domain-containing proteins (Data S1; Figure S4D). However, functionally divergent homologs within this protein family were also found across many eukaryotic groups, suggesting that the 2OGD chlorophyll *c* synthase domain might have arisen after a duplication event within a eukaryote in a related gene, followed by a change of function in one of the duplicates.

### Application of chlorophyll *c* synthase for improving plant photosynthesis

Given the urgent need to improve agricultural yields in the face of escalating climate perturbations, the optimization of photosynthesis has emerged as a critical goal. Diversifying the photosynthetic pigment portfolio of plants by drawing from the broad light-harvesting strategies of various photosynthetic organisms offers a promising approach. While the production of non-native chlorophylls in plants, such as chlorophyll *f* to capture energy in

the far-red light spectrum, has been proposed,<sup>60–62</sup> no production of these pigments has been previously demonstrated. Discovery of the BmCHLCS, its functional expression, and ultimate production of chlorophyll *c* in a plant serves as a first step to augment the photosynthetic palette of plants. To build an operational light-harvesting complex that utilizes chlorophyll *c*, the heterologous production of fucoxanthin-chlorophyll *a/c* binding protein (FCP) and fucoxanthin *in planta* are the next milestones to achieve. The confluence of recent breakthroughs in understanding fucoxanthin biosynthesis<sup>28,29</sup> and the production of chlorophyll *c* in this study brings the goal of harnessing a wider range of the light spectrum for plant photosynthesis increasingly within reach.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Dinoflagellate strains and growth conditions
  - Plant growth conditions
- METHOD DETAILS
  - Mutant isolation and characterization
  - Pigment extraction
  - UHPLC-APCI-qTOF-MS analysis
  - Bioinformatic identification of CHLCS
  - Heterologous expression of CHLCS
  - Localization of BmCHLCS in *N. benthamiana*
  - Chloroplast isolation from tobacco leaves
  - Signal peptide and transit peptide analysis
  - Western blot of isolated chloroplasts
  - Phylogenetic analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.cub.2023.12.068>.

## ACKNOWLEDGMENTS

We thank Prof. Poul Erik Jensen (University of Copenhagen) for input on the experimental setup for chlorophyll *c* biosynthesis *in planta*, Prof. Birger Lindberg Møller for valuable discussions on chloroplast targeting of biosynthetic enzymes in plants, Prof. Christiane Funk and Luisa Corredor Arias for providing us a culture of *Guillardia theta*, and Weichao Huang for early discussions about the gene discovered in this study. We thank Joseph Russo and Andrea Kirk for assisting with *lbr1* cultures. We also thank members of the Jinkerson, Xiang, and Andersen-Ranberg laboratories for helpful discussions. This work was supported by funds provided by Novo Nordisk Foundation grant NNF20OC0061048 to J.A.-R. and D.P.-H.; Novo Nordisk Foundation grant NNF21OC0070602 to D.P.-H.; the Hakai Institute and Gordon and Betty Moore Foundation (<https://doi.org/10.37807/GBMF9201>) to P.J.K.; the University of California, Riverside to T.X.; NSF-IOS EDGE Award to T.X.

(2308644); a Simons Fellowship of the Life Sciences Research Foundation to R.E.J.; and a University of California Regents' Faculty Fellowship to R.E.J.

## AUTHOR CONTRIBUTIONS

R.E.J., T.X., and J.A.-R. conceived the study; R.E.J. isolated *lbr1* and performed bioinformatic and phylogenetic analyses; T.X. phenotyped *lbr1* and performed sequencing; R.O.-F. and D.P.-H. performed chlorophyll *c* synthase localization experiments; J.A.-R. performed pigment analysis; J.A.-R. and D.P.-H. performed heterologous expression of chlorophyll *c* synthase; E.C.C., A.C., and P.J.K. performed the phylogenetic analysis and formulated the evolutionary theory; and R.E.J., T.X., P.J.K., and J.A.-R. wrote the manuscript with contributions and edits from all authors.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work the author(s) used ChatGPT 4 in order to improve readability and language of the work. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Received: November 10, 2023

Revised: December 18, 2023

Accepted: December 20, 2023

Published: December 28, 2023

## REFERENCES

1. Field, C.B., Behrenfeld, M.J., Randerson, J.T., and Falkowski, P. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* 281, 237–240.
2. Nelson, D.M., Tréguer, P., Brzezinski, M.A., Leynaert, A., and Quéguiner, B. (1995). Production and dissolution of biogenic silica in the ocean: revised global estimates, comparison with regional data and relationship to biogenic sedimentation. *Global Biogeochem. Cycles* 9, 359–372.
3. 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. Eukaryot. Microbiol.* 46, 347–366.
4. Green, B.R. (2011). After the primary endosymbiosis: an update on the chromalveolate hypothesis and the origins of algae with Chl *c*. *Photosynth. Res.* 107, 103–115.
5. Jeffrey, S.W. (1976). The occurrence of chlorophyll *c*<sub>1</sub> and *c*<sub>2</sub> in algae. *J. Phycol.* 12, 349–354.
6. Jinkerson, R.E., Russo, J.A., Newkirk, C.R., Kirk, A.L., Chi, R.J., Martindale, M.Q., Grossman, A.R., Hatta, M., and Xiang, T. (2022). Cnidarian-Symbiodiniaceae symbiosis establishment is independent of photosynthesis. *Curr. Biol.* 32, 2402–2415.e4.
7. Gornik, S.G., Maegele, I., Hambleton, E.A., Voss, P.A., Waller, R.F., and Guse, A. (2022). Nuclear transformation of a dinoflagellate symbiont of corals. *Front. Mar. Sci.* 9, 1035413.
8. Bach, S.S., Bassard, J.-E., Andersen-Ranberg, J., Møldrup, M.E., Simonsen, H.T., and Hamberger, B. (2014). High-throughput testing of terpenoid biosynthesis candidate genes using transient expression in *Nicotiana benthamiana*. *Methods Mol. Biol.* 1153, 245–255.
9. Sparkes, I.A., Runions, J., Kearns, A., and Hawes, C. (2006). Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nat. Protoc.* 1, 2019–2025.
10. Myśliwa-Kurdiel, B., Latowski, D., and Strzałka, K. (2019). Chapter Three - Chlorophylls *c*—Occurrence, synthesis, properties, photosynthetic and evolutionary significance. In *Advances in Botanical Research Advances in Botanical Research*, B. Grimm, ed. (Academic Press), pp. 91–119.

11. Beale, S.I. (1999). Enzymes of chlorophyll biosynthesis. *Photosynth. Res.* **60**, 43–73.
12. Granick, S. (1959). Magnesium porphyrins formed by barley seedlings treated with  $\delta$ -aminolevulinic acid. *Plant Physiol* **34**. xvi–xxi.
13. Fujita, Y., and Bauer, C.E. (2000). Reconstitution of light-independent protochlorophyllide reductase from purified bchl and BchN-BchB subunits. In vitro confirmation of nitrogenase-like features of a bacteriochlorophyll biosynthesis enzyme. *J. Biol. Chem.* **275**, 23583–23588.
14. Liu, Z., Yan, H., Wang, K., Kuang, T., Zhang, J., Gui, L., An, X., and Chang, W. (2004). Crystal structure of spinach major light-harvesting complex at 2.72 Å resolution. *Nature* **428**, 287–292.
15. Islam, M.S., Leissing, T.M., Chowdhury, R., Hopkinson, R.J., and Schofield, C.J. (2018). 2-Oxoglutarate-dependent oxygenases. *Annu. Rev. Biochem.* **87**, 585–620.
16. Bolte, K., Bullmann, L., Hempel, F., Bozarth, A., Zauner, S., and Maier, U.-G. (2009). Protein targeting into secondary plastids. *J. Eukaryot. Microbiol.* **56**, 9–15.
17. Gschloessl, B., Guermeur, Y., and Cock, J.M. (2008). HECTAR: a method to predict subcellular targeting in heterokonts. *BMC Bioinform.* **9**, 393.
18. Thumhuri, V., Almagro Armenteros, J.J., Johansen, A.R., Nielsen, H., and Winther, O. (2022). DeepLoc 2.0: multi-label subcellular localization prediction using protein language models. *Nucleic Acids Res.* **50**, W228–W234.
19. Bassham, D.C., Bartling, D., Mould, R.M., Dunbar, B., Weisbeek, P., Herrmann, R.G., and Robinson, C. (1991). Transport of proteins into chloroplasts. Delineation of envelope “transit” and thylakoid “transfer” signals within the pre-sequences of three imported thylakoid lumen. *J. Biol. Chem.* **266**, 23606–23610.
20. Zacharias, D.A., Violin, J.D., Newton, A.C., and Tsien, R.Y. (2002). Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science* **296**, 913–916.
21. Dautermann, O., Lyska, D., Andersen-Ranberg, J., Becker, M., Fröhlich-Nowoisky, J., Gartmann, H., Krämer, L.C., Mayr, K., Pieper, D., Rij, L.M., et al. (2020). An algal enzyme required for biosynthesis of the most abundant marine carotenoids. *Sci. Adv.* **6**, eaaw9183.
22. Keeling, P.J., Burki, F., Wilcox, H.M., Allam, B., Allen, E.E., Amaral-Zettler, L.A., Armbrust, E.V., Archibald, J.M., Bharti, A.K., Bell, C.J., et al. (2014). The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. *PLoS Biol.* **12**, e1001889.
23. Kang, H.C., Jeong, H.J., Park, S.A., Ok, J.H., You, J.H., Eom, S.H., Park, E.C., Jang, S.H., and Lee, S.Y. (2021). Comparative transcriptome analysis of the phototrophic dinoflagellate *Biecheleropsis adriatica* grown under optimal temperature and cold and heat stress. *Front. Mar. Sci.* **8**, 761095.
24. LaJeunesse, T.C., Parkinson, J.E., Gabrielson, P.W., Jeong, H.J., Reimer, J.D., Voolstra, C.R., and Santos, S.R. (2018). Systematic revision of Symbiodiniaceae highlights the antiquity and diversity of coral endosymbionts. *Curr. Biol.* **28**, 2570–2580.e6.
25. Burki, F., Roger, A.J., Brown, M.W., and Simpson, A.G.B. (2020). The new tree of eukaryotes. *Trends Ecol. Evol.* **35**, 43–55.
26. Derelle, R., López-García, P., Timpano, H., and Moreira, D. (2016). A phylogenomic framework to study the diversity and evolution of stramenopiles (=Heterokonts). *Mol. Biol. Evol.* **33**, 2890–2898.
27. Jeffrey, S.W., Wright, S.W., and Zapata, M. (2011). Microalgal classes and their signature pigments. In *Phytoplankton Pigments* (Cambridge University Press), pp. 3–77.
28. Bai, Y., Cao, T., Dautermann, O., Buschbeck, P., Cantrell, M.B., Chen, Y., Lein, C.D., Shi, X., Ware, M.A., Yang, F., et al. (2022). Green diatom mutants reveal an intricate biosynthetic pathway of fucoxanthin. *Proc. Natl. Acad. Sci. USA* **119**, e2203708119.
29. Cao, T., Bai, Y., Buschbeck, P., Tan, Q., Cantrell, M.B., Chen, Y., Jiang, Y., Liu, R.-Z., Ries, N.K., Shi, X., et al. (2023). An unexpected hydratase synthesizes the green light-absorbing pigment fucoxanthin. *Plant Cell* **35**, 3053–3072.
30. Strasser, J.F.H., Irisarri, I., Williams, T.A., and Burki, F. (2021). Author Correction: A molecular timescale for eukaryote evolution with implications for the origin of red algal-derived plastids. *Nat. Commun.* **12**, 3574.
31. Bodyl, A., Stiller, J.W., and Mackiewicz, P. (2009). Chromalveolate plastids: direct descent or multiple endosymbioses? *Trends Ecol. Evol.* **24**, 119–121.
32. Stiller, J.W., Schreiber, J., Yue, J., Guo, H., Ding, Q., and Huang, J. (2014). The evolution of photosynthesis in chromist algae through serial endosymbioses. *Nat. Commun.* **5**, 5764.
33. Ševčíková, T., Horák, A., Klimeš, V., Zbránková, V., Demir-Hilton, E., Sudek, S., Jenkins, J., Schmutz, J., Příbyl, P., Fousek, J., et al. (2015). Updating algal evolutionary relationships through plastid genome sequencing: did alveolate plastids emerge through endosymbiosis of an ochrophyte? *Sci. Rep.* **5**, 10134.
34. Keeling, P.J. (2004). Diversity and evolutionary history of plastids and their hosts. *Am. J. Bot.* **91**, 1481–1493.
35. Saldarriaga, J.F., Max Taylor, F.J.R., Cavalier-Smith, T., Menden-Deuer, S., and Keeling, P.J. (2004). Molecular data and the evolutionary history of dinoflagellates. *Eur. J. Protistol.* **40**, 85–111.
36. Keeling, P.J. (2010). The endosymbiotic origin, diversification and fate of plastids. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **365**, 729–748.
37. Dorrell, R.G., and Howe, C.J. (2015). Integration of plastids with their hosts: lessons learned from dinoflagellates. *Proc. Natl. Acad. Sci. USA* **112**, 10247–10254.
38. Loenarz, C., and Schofield, C.J. (2008). Expanding chemical biology of 2-oxoglutarate oxygenases. *Nat. Chem. Biol.* **4**, 152–156.
39. Aik, W., McDonough, M.A., Thalhammer, A., Chowdhury, R., and Schofield, C.J. (2012). Role of the jelly-roll fold in substrate binding by 2-oxoglutarate oxygenases. *Curr. Opin. Struct. Biol.* **22**, 691–700.
40. Fletcher, S.C., and Coleman, M.L. (2020). Human 2-oxoglutarate-dependent oxygenases: nutrient sensors, stress responders, and disease mediators. *Biochem. Soc. Trans.* **48**, 1843–1858.
41. Martinez, S., and Hausinger, R.P. (2015). Catalytic mechanisms of Fe(II)- and 2-oxoglutarate-dependent oxygenases. *J. Biol. Chem.* **290**, 20702–20711.
42. Pimenta Lange, M.J., Szperlinski, M., Kalix, L., and Lange, T. (2020). Cucumber gibberellin 1-oxidase/desaturase initiates novel gibberellin catabolic pathways. *J. Biol. Chem.* **295**, 8442–8448.
43. Stokes, G.G. (1864). On the supposed identity of biliverdin with chlorophyll, with remarks on the constitution of chlorophyll. *Proc. R. Soc. Lond.* **13**, 144–145.
44. Jeffrey, S.W. (1968). Two spectrally distinct components in preparations of chlorophyll c. *Nature* **220**, 1032–1033.
45. Jeffrey, S.W. (1969). Properties of two spectrally different components in chlorophyll c preparations. *Biochim. Biophys. Acta* **177**, 456–467.
46. Jiang, Y., Cao, T., Yang, Y., Zhang, H., Zhang, J., and Li, X. (2023). A chlorophyll c synthase widely co-opted by phytoplankton. *Science* **382**, 92–98.
47. Gaubier, P., Wu, H.J., Laudé, M., Delseny, M., and Grellet, F. (1995). A chlorophyll synthetase gene from *Arabidopsis thaliana*. *Mol. Gen. Genet.* **249**, 58–64.
48. Rüdiger, W. (2007). Biosynthesis of chlorophylls a and b: the last steps. In *Advances in Photosynthesis and Respiration* (Springer Netherlands), pp. 189–200.
49. Nelson, J.R., and Wakeham, S.G. (1989). A phytol-substituted chlorophyll c from *Emiliania huxleyi* (Prymnesiophyceae). *J. Phycol.* **25**, 761–766.
50. Tanaka, A., Ito, H., Tanaka, R., Tanaka, N.K., Yoshida, K., and Okada, K. (1998). Chlorophyll a oxygenase (CAO) is involved in chlorophyll b formation from chlorophyll a. *Proc. Natl. Acad. Sci. USA* **95**, 12719–12723.
51. Oster, U., Tanaka, R., Tanaka, A., and Rüdiger, W. (2000). Cloning and functional expression of the gene encoding the key enzyme for chlorophyll b biosynthesis (CAO) from *Arabidopsis thaliana*. *Plant J.* **21**, 305–310.

52. Lau, W., and Sattely, E.S. (2015). Six enzymes from mayapple that complete the biosynthetic pathway to the etoposide aglycone. *Science* **349**, 1224–1228.
53. Araújo, W.L., Martins, A.O., Fernie, A.R., and Tohge, T. (2014). 2-Oxoglutarate: linking TCA cycle function with amino acid, glucosinolate, flavonoid, alkaloid, and gibberellin biosynthesis. *Front. Plant Sci.* **5**, 552.
54. Wang, Y., Shi, Y., Li, K., Yang, D., Liu, N., Zhang, L., Zhao, L., Zhang, X., Liu, Y., Gao, L., et al. (2021). Roles of the 2-oxoglutarate-dependent dioxygenase superfamily in the flavonoid pathway: a review of the functional diversity of F3H, FNS I, FLS, and LDOX/ANS. *Molecules* **26**, 6745.
55. Zhang, X., Wang, D., Elberse, J., Qi, L., Shi, W., Peng, Y.-L., Schuurink, R.C., Van den Ackerveken, G., and Liu, J. (2021). Structure-guided analysis of Arabidopsis JASMONATE-INDUCED OXYGENASE (JOX) 2 reveals key residues for recognition of jasmonic acid substrate by plant JOXs. *Mol. Plant* **14**, 820–828.
56. Bailey, P.S.J., Ortmann, B.M., Martinelli, A.W., Houghton, J.W., Costa, A.S.H., Burr, S.P., Antrobus, R., Frezza, C., and Nathan, J.A. (2020). ABHD11 maintains 2-oxoglutarate metabolism by preserving functional lipoylation of the 2-oxoglutarate dehydrogenase complex. *Nat. Commun.* **11**, 4046.
57. Huergo, L.F., and Dixon, R. (2015). The emergence of 2-oxoglutarate as a master regulator metabolite. *Microbiol. Mol. Biol. Rev.* **79**, 419–435.
58. Jiang, D., Li, G., Chen, G., Lei, J., Cao, B., and Chen, C. (2021). Genome-wide identification and expression profiling of 2OGD superfamily genes from three Brassica plants. *Genes* **12**, 1399.
59. Keeling, P.J. (2013). The number, speed, and impact of plastid endosymbioses in eukaryotic evolution. *Annu. Rev. Plant Biol.* **64**, 583–607.
60. Ho, M.-Y., Shen, G., Canniffe, D.P., Zhao, C., and Bryant, D.A. (2016). Light-dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II. *Science* **353**, aaf9178.
61. Shen, G., Canniffe, D.P., Ho, M.-Y., Kurashov, V., van der Est, A., Golbeck, J.H., and Bryant, D.A. (2019). Characterization of chlorophyll f synthase heterologously produced in *Synechococcus* sp. PCC 7002. *Photosynth. Res.* **140**, 77–92.
62. Kato, K., Shinoda, T., Nagao, R., Akimoto, S., Suzuki, T., Dohmae, N., Chen, M., Allakhverdiev, S.I., Shen, J.-R., Akita, F., et al. (2020). Structural basis for the adaptation and function of chlorophyll f in photosystem I. *Nat. Commun.* **11**, 238.
63. van der Fits, L., Deakin, E.A., Hoge, J.H., and Memelink, J. (2000). The ternary transformation system: constitutive virG on a compatible plasmid dramatically increases *Agrobacterium*-mediated plant transformation. *Plant Mol. Biol.* **43**, 495–502.
64. Hansen, N.L., Kjaerulf, L., Heck, Q.K., Forman, V., Staerk, D., Møller, B.L., and Andersen-Ranberg, J. (2022). Tripterygium wilfordii cytochrome P450s catalyze the methyl shift and epoxidations in the biosynthesis of triptonide. *Nat. Commun.* **13**, 5011.
65. Peyret, H., and Lomonosoff, G.P. (2013). The pEAQ vector series: the easy and quick way to produce recombinant proteins in plants. *Plant Mol. Biol.* **83**, 51–58.
66. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760.
67. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., and Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008.
68. Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* **6**, 80–92.
69. Cingolani, P., Patel, V.M., Coon, M., Nguyen, T., Land, S.J., Ruden, D.M., and Lu, X. (2012). Using *Drosophila melanogaster* as a model for genotoxic chemical mutational studies with a new program. *Front. Genet.* **3**, 35.
70. Russo, J.A., Xiang, T., and Jinkerson, R.E. (2023). Protocol for the generation of Symbiodiniaceae mutants using UV mutagenesis. *STAR Protoc.* **4**, 102627.
71. Bijttebier, S., D'Hondt, E., Noten, B., Hermans, N., Apers, S., and Voorspoels, S. (2014). Ultra high performance liquid chromatography versus high performance liquid chromatography: Stationary phase selectivity for generic carotenoid screening. *J. Chromatogr. A* **1332**, 46–56.
72. Xiang, T. (2018). Total RNA extraction from dinoflagellate Symbiodinium cells. *Bio-protocol* **8**, e2866.
73. Xiang, T., Nelson, W., Rodriguez, J., Tolleter, D., and Grossman, A.R. (2015). Symbiodinium transcriptome and global responses of cells to immediate changes in light intensity when grown under autotrophic or mixotrophic conditions. *Plant J.* **82**, 67–80.
74. Xiang, T., Lehnert, E., Jinkerson, R.E., Clowez, S., Kim, R.G., DeNofrio, J.C., Pringle, J.R., and Grossman, A.R. (2020). Symbiont population control by host-symbiont metabolic interaction in Symbiodiniaceae-cnidarian associations. *Nat. Commun.* **11**, 108.
75. The Pandas Development Team (2019). “Pandas-dev/pandas: Pandas”. <https://doi.org/10.5281/zenodo.3509135>.
76. Park, H.-Y., Seok, H.-Y., Park, B.-K., Kim, S.-H., Goh, C.-H., Lee, B.-H., Lee, C.-H., and Moon, Y.-H. (2008). Overexpression of Arabidopsis ZEP enhances tolerance to osmotic stress. *Biochem. Biophys. Res. Commun.* **375**, 80–85.
77. Tottey, S., Block, M.A., Allen, M., Westergren, T., Albrieux, C., Scheller, H.V., Merchant, S., and Jensen, P.E. (2003). Arabidopsis CHL27, located in both envelope and thylakoid membranes, is required for the synthesis of protochlorophyllide. *Proc. Natl. Acad. Sci. USA* **100**, 16119–16124.
78. Robinson, C., and Mant, A. (2002). Import of proteins into isolated chloroplast and thylakoid membranes. In *Molecular Plant Biology: A Practical Approach*, P.M. Gilmartin, and C. Bowler, eds. (Oxford University Press), pp. 123–146.
79. Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567–580.
80. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.
81. Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780.
82. Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973.
83. Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2 - approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, e9490.
84. Richter, D.J., Berney, C., Strasser, J.F.H., Poh, Y.-P., Herman, E.K., Muñoz-Gómez, S.A., Wideman, J.G., Burki, F., and de Vargas, C. (2022). EukProt: A database of genome-scale predicted proteins across the diversity of eukaryotes. *Peer Community J.* **2**, e56.
85. Tice, A.K., Žihala, D., Pánek, T., Jones, R.E., Salomaki, E.D., Nenarokov, S., Burki, F., Eliáš, M., Eme, L., Roger, A.J., et al. (2021). PhyloFisher: A phylogenomic package for resolving eukaryotic relationships. *PLoS Biol.* **19**, e3001365.
86. Cho, A., Tikhonenkov, D.V., Hehenberger, E., Karnkowska, A., Mylnikov, A.P., and Keeling, P.J. (2022). Monophyly of diverse Bigyromonadea and their impact on phylogenomic relationships within stramenopiles. *Mol. Phylogenet. Evol.* **171**, 107468.
87. Azuma, T., Pánek, T., Tice, A.K., Kayama, M., Kobayashi, M., Miyashita, H., Suzuki, T., Yabuki, A., Brown, M.W., and Kamikawa, R. (2022). An enigmatic stramenopile sheds light on early evolution in Ochrophyta plastid organellogenesis. *Mol. Biol. Evol.* **39**, msac065.

88. Dorrell, R.G., Azuma, T., Nomura, M., Audren de Kerdrel, G., Paoli, L., Yang, S., Bowler, C., Ishii, K.-I., Miyashita, H., Gile, G.H., and Kamikawa, R. (2019). Principles of plastid reductive evolution illuminated by nonphotosynthetic chrysophytes. *Proc. Natl. Acad. Sci. USA* *116*, 6914–6923.
89. Beisser, D., Graupner, N., Bock, C., Wodniok, S., Grossmann, L., Vos, M., Sures, B., Rahmann, S., and Boenigk, J. (2017). Comprehensive transcriptome analysis provides new insights into nutritional strategies and phylogenetic relationships of chrysophytes. *PeerJ* *5*, e2832.
90. Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* *32*, 268–274.
91. Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., and Vinh, L.S. (2018). UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* *35*, 518–522.
92. Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., and Jermini, L.S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* *14*, 587–589.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
$\alpha$ -FLAG antibody	Sigma-Aldrich	F3165; RRID: AB_259529
<b>Bacterial and virus strains</b>		
<i>Agrobacterium tumefaciens</i> AGL1	van der Fits et al. <sup>63</sup>	GV3850
<b>Biological samples</b>		
<i>Emiliana huxleyi</i>	Bigelow, National Center for Marine Algae and Microbiota	CCMP375
<i>Guillardia theta</i>	Bigelow, National Center for Marine Algae and Microbiota	CCMP2712
<i>Karenia brevis</i>	Bigelow, National Center for Marine Algae and Microbiota	CCMP2281
<i>Diacronema lutheri</i>	Bigelow, National Center for Marine Algae and Microbiota	CCMP1251
<i>Amphidinium carterae</i>	Bigelow, National Center for Marine Algae and Microbiota	CCMP3177
<i>Phaeodactylum tricornutum</i>	Bigelow, National Center for Marine Algae and Microbiota	CCMP632
<b>Chemicals, peptides, and recombinant proteins</b>		
$\delta$ -Aminolevulinic acid (ALA) (5451-09-2)	Thermo Scientific Chemicals	103920050
<b>Deposited data</b>		
RNA-Seq data of <i>lbr1</i>	NCBI SRA	SRA: PRJNA1054151 (Biosample: SAMN38909378)
The mRNA sequence of <i>CHLCS</i>	NCBI GenBank	GenBank: OR978340
<b>Experimental models: Organisms/strains</b>		
Clonal axenic <i>Breviolum minutum</i> (Clade B) strain SSB01	The Aiptasia Symbiosis Resource	SSB01
<i>Breviolum minutum lbr1</i>	Jinkerson et al. <sup>6</sup>	<i>lbr1</i>
<i>Nicotiana benthamiana</i>	University of Copenhagen. Dept. of Plant & Environmental Sciences	N/A
<b>Oligonucleotides</b>		
Primers for cloning, see <a href="#">Table S2</a>	IDT	N/A
CHLCS and CHLC sequences, see <a href="#">Table S2</a>	Twist Biosciences	N/A
<b>Recombinant DNA</b>		
new_pLIFE	Hansen et al. <sup>64</sup>	N/A
pEAQ-HT	Peyret et al. <sup>65</sup>	N/A
<b>Software and algorithms</b>		
ImageJ	NIH – public domain	<a href="https://imagej.nih.gov/ij/">https://imagej.nih.gov/ij/</a>
Python (2.7.11)	Python Software Foundation	<a href="https://www.python.org">https://www.python.org</a>
SciPy (0.17.0)		<a href="https://www.scipy.org">https://www.scipy.org</a>
Pandas (0.18.1)		<a href="https://pandas.pydata.org/">https://pandas.pydata.org/</a>
Excel (16.80)	Microsoft	<a href="https://www.microsoft.com/en-us/microsoft-365/excel">https://www.microsoft.com/en-us/microsoft-365/excel</a>

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
BWA (0.7.17)	Li et al. <sup>66</sup>	<a href="https://github.com/lh3/bwa">https://github.com/lh3/bwa</a>
SAMtools and BCftools	Danecek et al. <sup>67</sup>	<a href="https://www.htslib.org/">https://www.htslib.org/</a>
SnEff	Cingolani et al. <sup>68</sup>	<a href="https://pcingola.github.io/SnpEff/">https://pcingola.github.io/SnpEff/</a>
SnpSift	Cingolani et al. <sup>69</sup>	<a href="https://pcingola.github.io/SnpEff/">https://pcingola.github.io/SnpEff/</a>
<b>Other</b>		
<i>B. minutum</i> WT strain SSB01 Symb6 transcriptome assembly	NCBI GenBank	GICE00000000
Transcript sequence of s6_3623	NCBI GenBank	GICE01003545.1

**RESOURCE AVAILABILITY**

**Lead contact**

Further information and requests for resources and reagent should be directed to and will be fulfilled by the lead contact, Robert Jinkerson ([robert.jinkerson@ucr.edu](mailto:robert.jinkerson@ucr.edu))

**Materials availability**

Materials generated in this study are available from the [lead contact](#) upon request. Distribution of lines is governed by the appropriate material transfer agreements (MTAs) and availability of algae material is dependent on provision of appropriate import permits acquired by the receiver.

**Data and code availability**

- All sequencing data have been deposited at NCBI Sequence Read Archive (SRA) and are publicly available as of the date of publication. Accession numbers are listed in the [key resources table](#).
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

**EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS**

**Dinoflagellate strains and growth conditions**

Clonal and axenic *Breviolum minutum* (Clade B) strain SSB01, mutant *lbr1*, and microalgae strains purchased from the Bigelow collection (National Center for Marine Algae and Microbiota, <https://ncma.bigelow.org/>) including *Amphidinium carterae* CCMP3177, *Emiliana huxleyi* CCMP375, *Guillardia theta* CCMP2712, *Karenia brevis* CCMP2281, *Diacronema lutheri* CCMP1251, *Amphidinium carterae* CCMP3177, and *Phaeodactylum tricornutum* CCMP632, were used in this study. SSB01 and *lbr1* were grown in 37.4 g L<sup>-1</sup> marine broth (MB) (Millipore-Sigma 76448) medium supplemented with 10 g L<sup>-1</sup> glucose (Millipore-Sigma G8270). SSB01 WT and *lbr1* cultures were incubated without agitation at 27°C on a 4 hr-light/20 hr-dark cycle with an irradiance of ~10 μmol photons m<sup>-2</sup> s<sup>-1</sup> of photosynthetically active radiation (PAR) provided by Percival SciWhite LED tiles. The microalgae strains purchased from the Bigelow collection were grown in f/2 or L1 - Si media according to Bigelow protocols.

**Plant growth conditions**

*Nicotiana benthamiana* was used in this study. The plants were grown on soil in the green house at the University of Copenhagen (18°C-25°C). For transient expression analysis, *Nicotiana benthamiana* leaves were harvested and placed in 6-well plates in a buffer supplemented with δ-aminolevulinic acid (ALA) overnight in the dark or light as indicated in [method details](#).

**METHOD DETAILS**

**Mutant isolation and characterization**

*lbr1* was isolated as previously described.<sup>6</sup> Briefly, clonal and axenic liquid cultures of *Breviolum minutum* (Clade B) strain SSB01 were subjected to UV mutagenesis.<sup>70</sup> Cells were plated on agar plates of marine broth (MB) media (37.4 g L<sup>-1</sup>, Millipore-Sigma 76448) supplemented with 10 g L<sup>-1</sup> glucose (Millipore-Sigma G8270) and kept in the dark. Colonies with an altered color were isolated and transferred to autotrophic conditions (MB in the light) to identify mutants deficient in autotrophic growth. One mutant identified, *lbr1*, had a light pale yellow color, was incapable of autotrophic growth, and then subjected to pigment analysis. *Emiliana huxleyi* CCMP375, *Guillardia theta* CCMP2712, *Karenia brevis* CCMP2281, *Diacronema lutheri* CCMP1251, *Amphidinium carterae*

CCMP3177, *Thalassiosira pseudonana* CCMP1335 and *Phaeodactylum tricorutum* CCMP632 microalgae were purchased from the Bigelow collection (National Center for Marine Algae and Microbiota, <https://ncma.bigelow.org/>) and grown according to Bigelow protocols.

### Pigment extraction

For pigment extraction from algae, cells were grown in MB supplemented with 10 g L<sup>-1</sup> glucose in the dark. Log phase cells were collected by centrifugation, and lyophilized in the dark. Extraction solvent (90% MeOH:H<sub>2</sub>O + 5 ppm 8-apo-carotenal) was added to lyophilized algal cells. Samples were incubated in brown glass vials at room temperature for 1 h. For pigment extraction from *Nicotiana benthamiana*, leaves expressing chlorophyll *c* synthase genes were extracted with 100% acetone. After extraction, the acetone was evaporated and the pigments were resuspended in 90% MeOH with 5 ppm 8-apocarotenal as an internal standard. All samples were filtered through a 96-well filter plate with 0.2 μm pore size polyvinylidene fluoride membrane (PVDF) membrane (Agilent, 203980-100) prior to analysis.

### UHPLC-APCI-qTOF-MS analysis

Pigment extracts were analyzed using an Ultimate 3000 UHPLC+ Focused system (Dionex Corporation, Sunnyvale, CA, USA) coupled to a Bruker Compact APCI-QTOF-MS (Bruker, Billerica, MA, USA) system. Method for analysis was adopted as previously described.<sup>71</sup> Briefly, samples were separated on a ACQUITY UPLC HSS C18 SB Column (100 Å, 1.8 μm, 2.1 mm × 100 mm; Phenomenex, Torrance, CA, USA) maintained at 35°C with a flow rate of 0.5 mL min<sup>-1</sup>. Injection volume was 10 μL. Mobile phase consisted of A: 50:22.5:22.5:5 5 mM ammonium acetate dissolved in water:methanol:acetonitrile:ethyl acetate and B: 50:50 acetonitrile:ethyl acetate. The liquid chromatography (LC) procedure was outlined as following: 0–0.1 min, 10% B; 0.1–0.8 min, linear increase from 10 to 30% B; 0.8–20 min, increase 30% to 91% B; 20–20.1 min, increase from 91% to 100% B; 20.1–20.4 min isocratic; 20.4–20.5 min linear decrease from 100% to 10% B; 20.5–23 min isocratic. Mass spectra were acquired in positive ion mode over a scan range of 50–1200 mass-to-charge ratios (*m/z*) and 2 Hz sample rate, with the following settings for MS and atmospheric pressure chemical ionization (APCI): Capillary voltage, 4000 V; endplate offset, 500 V; corona, 4 μA; Vaporizer temperature, 400°C; dry gas temperature, 250°C; dry gas flow, 5 L min<sup>-1</sup>; and nebulizer pressure, 3 bar. Acquired chromatogram data were calibrated using an internal standard, APCI-L (low concentration tuning mix, Agilent technologies). Data analysis was performed with DataAnalysis 4.3 (Bruker, Billerica, MA, USA).

### Bioinformatic identification of CHLCS

Total RNAs from *B. minutum* WT strain SSB01 and *lbr1* were isolated as previously described.<sup>72</sup> RNA Sequencing was subsequently carried out by Novogene using an Illumina NovaSeq 6000 Sequencing System. Reads were deposited at the NCBI SRA in the Bio-project PRJNA1054151 under Biosample: SAMN38909378. Reads were mapped to the *B. minutum* WT strain SSB01 Symb6 transcriptome assembly (GenBank: GICE00000000)<sup>73,74</sup> using BWA.<sup>66</sup> Variants were called using bcftools mpileup (–max-depth 250 –adjust-MQ 50) and bcftools call (–m –ploidy 1).<sup>67</sup> The functional effects of the variants identified were called with SnpEff<sup>68</sup> and filtered with SnpSift<sup>69</sup> and the Pandas python package.<sup>75</sup> Two wild type samples (denoted as WT and WT2) were included in the analysis to help with variant identification in *lbr1*. Variants were filtered to only include those that met the following criteria: (1) *lbr1* called the alternative genotype; (2) WT and WT2 called the reference genotype; (3) WT and WT2 had greater than 10 read depth of coverage for the reference genotype; (4) WT and WT2 had less than 10 read depth of coverage for the alternative genotype; (5) *lbr1* had greater than 10 read depth of coverage for the alternative genotype; and (6) *lbr1* had less than 10 read depth of coverage for the reference genotype. 79 variants from 64 transcripts (Figure S2A) met all criteria and were manually reviewed. Transcript s6\_3623 (GenBank: GICE01003545.1), which had a 1-bp deletion resulting in a frameshift, was identified as a candidate for further investigation. The sequence of CHLOROPHYLL C SYNTHASE (CHLCS) is deposited at GenBank: OR978340.

### Heterologous expression of CHLCS

The function of different candidate chlorophyll *c* synthase genes was characterized by transient expression in *Nicotiana benthamiana* using established protocols.<sup>8</sup> CHLCS coding sequences from *Breviolum minutum* SSB01, *Emiliania huxleyi* (CCMP1516, XP\_005758961), *Karenia brevis* (CCMP2229, CAMPEP\_0173950170<sup>22</sup>), *Diacronema lutheri* (KAG8460781), *Guillardia theta* (CCMP2712 XP\_005836611), *Thalassiosira pseudonana* (CCMP1335 XP\_002294744), *Phaeodactylum tricorutum* (CCAP1055/1 XP\_002177807), *Amphidinium carterae* (CCMP1314, CAMPEP\_0176533696<sup>22</sup>) and *Biecheleriopsis adriatica* (BATY0608, transcript assembled from SRR11947552) were used for transient expression. Peptide sequences and sequences codon optimized for *N. benthamiana* expression can be found in Table S2. The transformation was performed with *Agrobacterium tumefaciens* AGL1<sup>63</sup> harboring new\_pLIFE and pEAQ-HT vectors.<sup>65,64</sup> Native genes, or those with plant chloroplast targeting signals, were cloned into the vectors. The PGK (phosphoglycerate kinase) promoter drove expression in the New pLIFE vector, while the CaMV (Cauliflower Mosaic Virus) 35S promoter was utilized for the pEAQ-HT vector. The N-terminal chloroplast targeting sequence from *N. tabacum* VDE was used to target the stroma side. The N-terminal signal from *Arabidopsis thaliana* ZEP was used for chloroplast lumen localization.<sup>21,76</sup> Tobacco transient expression was performed on 4 to 6-week-old *N. benthamiana* plants. After infiltration of *Agrobacterium* into tobacco leaves and expressing the desired genes for 4 days, the leaves were harvested for LC-MS analysis. For experiments involving δ-aminolevulinic acid (ALA) feeding,<sup>77</sup> leaf discs (diameter Ø = 2 cm) were extracted on the third day post-infiltration. Leaf discs were then incubated overnight in the dark, submerged in a buffer solution containing 10 mM δ-aminolevulinic acid, 5 mM MgCl<sub>2</sub> and 10 mM potassium phosphate pH 7.0.

### Localization of BmCHLCS in *N. benthamiana*

To determine if chloroplast transit peptide exists in BmCHLCS, HECTAR<sup>17</sup> was employed which identified a 22 amino acid signal peptide sequence. Using  $\Delta 22$ BmCHLCS as a query, the subcellular localization prediction method DeepLoc 2.0<sup>18</sup> predicted the presence of a plastid targeting sequence and a ASAFAP motif was observed at the alanine in position 67. To visualize CHLCS proteins, mCitrine was tagged to the C-terminal of the full-length or truncated forms of the protein using the new\_pLIFE vector for expression.<sup>8</sup> Constructs were generated via USER cloning, see primers Table S2. *N. benthamiana* leaf discs after 4 days of agroinfiltration were mounted on a microscope slide using perfluorodecalin (Sigma, P9900) and visualized with a Leica SP5 X confocal microscope. Citrine was excited at 510 nm using a white laser and its emission was captured at 520–560 nm. Chlorophyll autofluorescence was excited using a 458 nm argon laser and detected at 650–700 nm.

### Chloroplast isolation from tobacco leaves

Tobacco leaves were washed with distilled water and sectioned into fragments measuring approximately 1–3 cm in size. For every 5 mg of tissue, 5 mL of cold, sterile-filtered HS buffer (50 mM HEPES/KOH pH 8.0, 0.33 M sorbitol) was added. The tissue was homogenized in a blender using two 2-s pulses, separated by a 5-s break. The resulting homogenate was filtered through a two-layer nylon mesh into a chilled beaker. This homogenate was centrifuged at 1,000 g at 4°C for 7 min. After discarding the supernatant, the pellet was slowly resuspended in 4 mL of HS buffer. This was achieved by rolling the tube gently on ice, and a portion was aliquoted and set aside as the total protein. To isolate chloroplasts from the homogenate, a Percoll gradient was prepared in 15 mL round-bottom tubes by first adding 2.5 mL of an 80% Percoll solution in the HS buffer. This was carefully overlaid with 5 mL of a 40% Percoll solution in the HS buffer. The required percentages of Percoll in 1x HS buffer were achieved using 5x HS buffer stock and water. The homogenate was then introduced to this gradient using wide orifice tips and centrifuged at 3,200 g for 15 min at 4°C. The interface between 40% and 80% Percoll was collected as the intact chloroplast fraction. Subsequently, 3 volumes of 1x HS buffer were added, and the intact chloroplasts were centrifuged for 1,700 g for 2 min at 4°C. Protein concentration for each fraction was determined via Bradford assay. The isolated chloroplast fraction was aliquoted, snap frozen, and stored at -80°C<sup>78</sup>

### Signal peptide and transit peptide analysis

Protein sequence of BmCHLCS was analyzed using the HECTAR ver 1.3 for identification of heterokont type signal peptide.<sup>17</sup> DeepLoc ver. 2.0<sup>18</sup> was used for chloroplast transit peptide prediction using  $\Delta 22$ BmCHLCS, peptide sequence without the predicted signal peptide, as a query. TMHMM ver. 2.0<sup>79</sup> was used for prediction of transmembrane domains.

### Western blot of isolated chloroplasts

The total or chloroplast fractions were individually centrifuged at 20000 g for 10 min at 4°C. The pellet was resuspended in 1X SDS loading buffer (2% [w/v] SDS, 10% [v/v] glycerol, 0.02% [w/v] bromophenol blue, 62.5 mM Tris-HCl, pH 6.8, 1% [v/v]  $\beta$ -mercaptoethanol) and incubated at 95°C for 5 min. Samples were analyzed by SDS-PAGE 12% and immunodetection using a monoclonal  $\alpha$ -FLAG antibody (F3165 Sigma-Aldrich) in a dilution of 1:1000 in Tris-Buffered Saline (TBS) 3% milk and a secondary anti-mouse HRP conjugated antibody (DAKO P0161).

### Phylogenetic analysis

For chlorophyll c synthase, the query file consisted of the sequences identified in the present study. Using these query files, BLASTp (v2.9)<sup>80</sup> searches were performed (E-value threshold  $\leq 1e^{-25}$ ) against the Marine Microbial Eukaryotic Transcriptome Sequencing Project (MMETSP) transcriptome database.<sup>22</sup> The retrieved sequences were aligned with the original query sequences using MAFFT (v7.481)<sup>81</sup> and trimmed using trimAl (v1.4)<sup>82</sup> with a gap threshold of 0.8. Phylogenetic trees were then constructed from the trimmed alignments using FastTree (v2.1.11),<sup>83</sup> followed by a thorough inspection to eliminate any contaminant sequences. Subsequently, the purified gene collections were used to initiate another round of BLASTp searches (E-value threshold  $\leq 1e^{-25}$ ) against a curated database of stramenopiles, which include select ochrophytes and major stramenopile groups that are non-photosynthetic, compiled from Eukprot (v3),<sup>84</sup> PhyloFisher (v1.1.2),<sup>85</sup> and additional published resources.<sup>86–89</sup> For each gene, ochrophyte sequences were combined with their query file, and then aligned and trimmed as described above. The final phylogenetic trees were deduced using IQ-TREE2 (v2.1.0)<sup>90</sup> employing the ultrafast bootstrap approximation (UFBoot)<sup>91</sup> and ModelFinder<sup>92</sup> to determine the most suitable model for each alignment. For *ZEP1* and *VDL2*, query files of known amino acid sequence homologs from diverse eukaryotes were curated.

## QUANTIFICATION AND STATISTICAL ANALYSIS

In Figure 4C, the graph shows the mean with error bars representing the standard deviation of the mean. Sample sizes are provided in the figure legends and refer to the number of plant and algal samples (details are described in the [method details](#)). Pearson correlation analysis, two-tailed t-test, one-way ANOVA analysis were used to compare the ratio of chlorophyll  $c_1$  to  $c_2$  between *in planta* and *in alga* (details are described in the [results](#)). All statistical analysis was performed with Microsoft Excel (Version 16.80).