Evolutionary origin and functional mechanism of Lhcx in the diatom photoprotection

Minoru Kumazawa[1,2], Seiji Akimoto[3], Atsushi Takabayashi[2], Ko Imaizumi[1], Shoko Tsuji[1], Hazuki Hasegawa[4], Atsushi Sakurai[4], Sousuke Imamura[4], Noriko Ishikawa[1], Natsuko Inoue-Kashino[5], Yasuhiro Kashino[5] & Kentaro Ifuku[1]

[1]Graduate School of Agriculture, Kyoto University, Japan

[2]Institute of Low Temperature Science, Hokkaido University, Japan

[3]Graduate School of Science, Kobe University, Japan

[4]Space Environment and Energy Laboratories, NTT, Inc., Japan

[5]Graduate School of Science, University of Hyogo, Japan

Author for correspondence: Kentaro Ifuku (ifuku.kentaro.2m@kyoto-u.ac.jp)

Summary

- Diatoms, a group of red-lineage algae that possess plastids derived from red algae through secondary endosymbiosis, utilize the light-harvesting complex (LHC) subfamily Lhcx for photoprotection; however, its evolutionary origin and functional mechanism remain unclear.
- To address these questions, we performed molecular phylogenetic analysis of Lhcx/Lhcsr subfamilies and produced the *Lhcx1* knockout mutants of the diatom *Chaetoceros gracilis*, which has low Lhcx redundancy.
- Phylogenetic analysis revealed that diatom Lhcxs and the green algal Lhcsrs evolved from a common ancestor, with green plants acquiring them through horizontal gene transfer. The *Lhcx1* mutants of *C. gracilis* mostly abolished non-photochemical quenching (NPQ), and the time-resolved fluorescence measurements clarified that Lhcx1-mediated quenching occurs in the energetically detached antenna complexes. Consistently, clear-native PAGE using Amphipol suggested that CgLhcx1 would interact with FCP L-dimer, functioning as a peripheral antenna for $C_2S_2M_2$ PSII–FCPII. Notably, the *lhcx1* mutant exhibited a higher PSII effective quantum yield than wild-type (WT) under high-light acclimation, attributed to reduced antenna size and enhanced carbon fixation capacity. The lack of NPQ accelerated high-light acclimation, with increased xanthophylls in *lhcx1*, indicating that compensatory mechanisms can lead to enhanced photosynthetic efficiency.
- Our findings provide novel insights into the origin and molecular mechanisms of photoprotection in diatoms.

Introduction

Photosynthesis consists of photochemical reactions in which two photosystems connected in series convert light energy into chemical energy. Among photosynthetic organisms, red-lineage secondary endosymbiotic algae, which possess plastids derived from red algae acquired through cellular endosymbiosis, are widely adapted to aquatic environments on Earth (Pierella Karlusich *et al.*, 2020, 2025). Light conditions in these environments are as dynamic as those in terrestrial environments, making it extremely important for photosynthetic organisms to protect their photosystems from excess light to suppress the generation of harmful reactive oxygen species caused by the absorption of excessive light energy. Non-Photochemical Quenching (NPQ), which dissipates excess light energy as heat, plays a crucial role in this protective mechanism (Niyogi, 2000).

Multiple NPQ mechanisms with different induction times exist, and these exhibit distinct characteristics depending on the taxonomic group (Goss & Lepetit, 2015; Malnoë, 2018). In particular, energy-dependent quenching (qE) is a major component of NPQ and is carried out by Light-Harvesting Complex (LHC) proteins (Peers *et al.*, 2009; Goss & Lepetit, 2015; Buck *et al.*, 2019). In red-lineage secondary endosymbiotic algae, qE, dependent on the Lhcx subfamily of the LHC family, serves as the primary NPQ mechanism (Buck *et al.*, 2019; Croteau *et al.*, 2025). Green algae, on the other hand, utilize qE dependent on the Lhcsr subfamily, which is closely related to Lhcx, as well as state transitions (qT), where LHCII detaches from photosystem II (PSII) and binds to photosystem I (PSI) through phosphorylation control to adjust the excitation balance between the two photosystems (Allorent *et al.*, 2013; Minagawa & Tokutsu, 2015). However, the evolutionary origin of Lhcx/Lhcsr has not been extensively discussed due to a lack of comprehensive phylogenetic analyses of LHC subfamilies.

The LHC of diatoms, representative of red-lineage algae, is called FCP (Fucoxanthin-Chlorophyll (Chl) *a*/*c*-binding Protein) because it binds pigments such as Chl *c*, fucoxanthin, and diadinoxanthin

59 in addition to Chl *a*. Diatom NPQ primarily consists of qE, which depends on the proton gradient
60 (ΔpH) across the thylakoid membrane formed by light (Lavaud & Kroth, 2006; Goss & Lepetit, 2015).
61 This qE in diatoms depends on Lhcx, which is homologous to the Lhcsr family essential for qE in
62 green algae (Koziol *et al.*, 2007; Peers *et al.*, 2009). However, PsbS, which is necessary for qE
63 induction in land plants, is absent in diatoms. While the homology between Lhcx and Lhcsr amino
64 acid sequences is well known, differences in the mechanisms that regulate gene expression have been
65 reported (Petroutsos *et al.*, 2016; Allorent *et al.*, 2016; Tokutsu *et al.*, 2019a,b; Im *et al.*, 2024; Zhang
66 *et al.*, 2024).

67 While Lhcx is required for qE induction, xanthophyll pigments are also important for qE induction.
68 In diatoms, the diadinoxanthin (Ddx) to diatoxanthin (Dtx) conversion cycle (Ddx-Dtx cycle) is the
69 major xanthophyll cycle, with a strong positive correlation between Dtx amount and qE (Lavaud &
70 Lepetit, 2013). Additionally, the violaxanthin-antheraxanthin-zeaxanthin (VAZ) cycle, known in land
71 plants, also functions in diatoms, complementing the Ddx-Dtx cycle (Lohr & Wilhelm, 1999; Giossi
72 *et al.*, 2025). Furthermore, it has been reported that mitochondria in diatoms function as electron sinks
73 for the photosynthetic electron transport chain, contributing to the maintenance of normal chloroplast
74 function (Bailleul *et al.*, 2015).

75 NPQ in red-lineage algae has been studied in diatoms and *Nannochloropsis* belonging to heterokont
76 algae (Bailleul *et al.*, 2010; Park *et al.*, 2019; Buck *et al.*, 2019, 2021; Giovagnetti *et al.*, 2022). In
77 model diatom species, multiple Lhcx isoforms are expressed under different conditions and show
78 functional redundancy. Therefore, almost complete Lhcx-deficient strains have not been created in
79 diatoms to date, and the physiological responses and details of excitation energy transfer in states with
80 significantly deficient NPQ have not been fully elucidated. In contrast, *Chaetoceros gracilis* has low
81 redundancy of Lhcxs, with primarily only CgLhcx1 being expressed (Kumazawa *et al.*, 2022),
82 providing the possibility that an almost complete Lhcx-deficient strain can be obtained by creating a

83 knockout of this CgLhcx1. Furthermore, the three-dimensional structures of PSII–FCPII (Wang *et al.*,
84 2019; Nagao *et al.*, 2019, 2022) and PSI–FCPI (Nagao *et al.*, 2020a; Xu *et al.*, 2020) supercomplexes
85 have already been elucidated in this species, establishing a foundation for structural biology research.
86 This study aims to investigate the origin of Lhcx phylogenetically and to elucidate the mechanism of
87 its excitation energy quenching for high light (HL) protection by creating Lhcx1-deficient mutants in
88 *C. gracilis* using CRISPR/Cas9 genome editing technology.

Results

Origin of LHCs in the Lhcx/Lhcsr Subfamily

To elucidate the origins of the diatom Lhcx and green algal Lhcsr subfamilies, which perform NPQ, we conducted a comprehensive molecular phylogenetic analysis of light-harvesting complexes (LHCs) from red algae, green algae, cryptophytes, and diatoms, and represented the estimated phylogenetic tree as an unrooted tree (Figure 1). Comprehensive retrieval of LHC sequences by BLAST confirmed that, as in our previous reports, red algae possess only the Lhcr subfamily, cryptophytes have Lhcr and Lhcz subfamilies, and green algae contain Lhca, Lhcb, and Lhcsr subfamilies, while diatoms possess Lhcr, Lhcz, Lhcq, CgLhcr9 homolog, Lhcx, and Lhcf subfamilies (Kumazawa *et al.*, 2022; Kumazawa & Ifuku, 2024). In Figure 1, each LHC subfamily in the red lineage is separated. However, within the Lhcr subfamily, there exists diversity among the Lhcr proteins from red algae, cryptophytes, and diatoms, encompassing gene groups with distant evolutionary relationships.

Lhcx and Lhcsr formed a monophyletic clade with an ultrafast bootstrap value of 100% (Figure 1, inset). Within this clade, *Thalassiosira pseudonana* Lhcx6_1 and *Cyclotella cryptica* Lhcx6_1, derived from diatoms belonging to the order Thalassiosirales, clearly branched at the most basal position and were evolutionarily distant from typical Lhcx and Lhcsr proteins. All Lhcx1–3 proteins from the diatom *Chaetoceros gracilis* were positioned within the typical Lhcx clade. Additionally, the Lhcx and Lhcsr clade showed monophyly with the Lhcq, CgLhcr9 homolog, and Lhcf subfamilies, supported by a high bootstrap value (100%). Within that clade, Lhcq was located at the most basal position, while the Lhcx/Lhcsr clade was positioned internally. In contrast, the Lhcx/Lhcsr clade did not show monophyly with the Lhca/Lhcb clade. Furthermore, the Lhcr of red algae and the Lhca and Lhcb subfamilies, which are responsible for light harvesting in red and green algae possessing primary plastids, did not form a monophyletic group; between them existed cryptophyte Lhcr, diatom Lhcr,

113 cryptophyte and diatom Lhcz, as well as other subfamilies of the red evolutionary lineage. These data
114 suggest that Lhcsr and Lhcx subfamilies are derived from the red-lineage-specific subfamilies.
115
116 **Creating *Lhcx1* Knockout Strains with Ribozyme-Based Genome Editing**
117 *C. gracilis* has three typical Lhcx genes (*CgLhcx1–3*): according to the genome database of *C.*
118 *gracilis* (ChaetoBase v1.1, https://chaetoceros.nibb.ac.jp/), *CgLhcx1* showed the highest transcription
119 level by a significant margin (Kumazawa *et al.*, 2022). Therefore, we decided to create an *Lhcx1*-
120 deficient mutant in *C. gracilis* using the CRISPR/Cas9 system. To facilitate the comprehensive
121 analysis of LHC genes using genome editing in *C. gracilis,* we have developed the inducible
122 CRISPR/Cas9 system, in which both Cas9 and guide RNA (gRNA) expression is regulated by the
123 inducible Pol-II-type promoter. First, we introduced the *Cas9* gene under the nitrate reductase (NR)
124 promoter into wild-type *C. gracilis* with a vector having the nourseothricin resistance gene (*nat*) via
125 electroporation (Supplemental Figure 1A). Colonies obtained through antibiotic-resistance selection
126 were further selected for high Cas9 expression in the presence of nitrate by immunoblotting
127 (Supplemental Figure 1B). The selected hC9_5 strain showed morphology and growth characteristics
128 equivalent to the wild-type strain (WT) and was established as the host strain for the genome editing
129 (Supplemental Figure 1C). Second, to express mature gRNA in the host hC9_5 under the NR promoter,
130 we employed a method utilizing ribozymes (Gao & Zhao, 2014; Zhang *et al.*, 2017). This method
131 employed a ribozyme-gRNA-ribozyme (RGR) construct in which a hammerhead ribozyme with self-
132 cleavage activity at the 3' end was attached to the 5' end of the gRNA, and an HDV ribozyme with
133 self-cleavage activity at the 5' end was attached to the 3' end of the gRNA. This RGR cassette for the
134 gene of interest can be easily constructed by overlap PCR and inserted into the pCgNRp^{ble} vector,
135 having an antibiotic-resistant gene against zeocin (*ble*) for second selection.
136 The transformation of the host hC9_5 strain with the linearized pCgNRp^{ble}-Lhcx1-RGR vector

137 produced 252 zeocin-resistant colonies, from which 23 strains were randomly selected, and 13 strains
138 had various mutations in the target sequence of the *CgLhcx1* gene. Among the strains confirmed for
139 target sequence editing, *lhcx1-7*, *15*, and *234* strains were used for subsequent analysis. When the
140 nucleotide sequences of the *Lhcx1* region were confirmed by Sanger sequencing, heterogeneity was
141 observed (Supplemental Figure 2A). Since *C. gracilis* is a diploid organism like other diatoms,
142 mutagenic events can occur in both alleles to completely inactivate the target gene. However,
143 simultaneous mutations in both alleles are rare, and in most cases, only one allele is mutated, or neutral
144 mutations such as in-frame deletions/insertions may occur, resulting in the formation of mixed cell
145 populations with different mutation status and types through subsequent cell divisions (Huang &
146 Daboussi, 2017). Therefore, we examined the accumulation of Lhcx1 protein in these strains by
147 western blotting using an anti-Lhcx1 antibody (Supplemental Figure 2B). Almost no accumulation of
148 Lhcx1 protein was observed in the *lhcx1-7* and *lhcx1-15* strains. On the other hand, the *lhcx1-234*
149 strain showed reduced accumulation of Lhcx1 compared to WT, but it was not a complete deficiency
150 since *lhcx1-234* may include neutral mutations such as in-frame deletions/insertions.
151
152 **Phenotypic Analysis of *Lhcx1* Deficient Strains**
153 WT and *lhcx1-7*, *15*, and *234* strains were cultured with air bubbling for 7–8 days under low light
154 (LL) at 30 μmol photons m^{-2} s^{-1} and high light (HL) at 300 μmol photons m^{-2} s^{-1}. No significant
155 differences in growth were observed between WT and genome-edited strains (Supplemental Figure 3).
156 To investigate the photosynthetic characteristics of WT and *lhcx1* strains, pulse-amplified modulated
157 (PAM) fluorescence measurements were performed on day 3 and day 7–8 using cells cultured under
158 LL and HL conditions (Figure 2, Supplemental Figure 4). As a result, *lhcx1-7*, *15*, and *234* were hardly
159 able to induce NPQ, an indicator of thermal dissipation of excess excitation energy around PSII, under
160 both LL and HL conditions (Figure 2A, Supplemental Figure 4A). The *lhcx1* strains cultured under

161 LL showed lower photochemical quenching qP under light exposure compared to WT, indicating a
162 more reduced state of the plastquinone pool. However, this trend was not clear in the *lhcx1* strains
163 cultured under HL, because non-regulated energy dissipation due to photoinhibition affects the qP
164 level (Figure 2B). The effective quantum yield of PSII, Y(II), showed higher values in all *lhcx1* strains
165 compared to the WT strain in cells cultured under HL for 7–8 days (Figure 2C). In cells cultured under
166 HL for 3 days, *lhcx1* strains showed a trend of higher Y(II) compared to WT, but the difference was
167 not significant (Supplemental Figure 4C). The *lhcx1-7, 15* strains exhibited significantly higher
168 oxygen-evolving activity compared to WT cells at the same total Chl *a* content of 5 μg (Supplemental
169 Figure 5). The *lhcx1-234* strains showed the tendency towards higher oxygen-evolving activity, though
170 it was not statistically significant ($p = 0.157$). This may be due to the mosaic genetic background of
171 *Lhcx1* mutations in *lhcx1-234*.
172
173 **Whole-Cell Absorption Spectra and Pigment Analysis**
174 The whole-cell absorption spectra showed no significant differences between WT and *lhcx1*
175 mutants in cells cultured under LL, while in the cells cultured under HL, the *lhcx1* mutants exhibited
176 greater absorption in the 400–500 nm region compared to the WT (Supplemental Figure 6). Since this
177 region corresponds to carotenoid absorption, pigment analysis using HPLC (High Performance Liquid
178 Chromatography) was performed to examine carotenoid contents. The amount of Chl *a* per cell was
179 decreased in the mutants compared to WT under LL conditions (Figure 3A). Additionally, under HL
180 conditions, WT also showed a decrease in Chl *a* content compared to LL conditions, but the decrease
181 tends to be greater in the *lhcx1* mutants. Since the total absorption spectra were normalized at the Qy
182 peak of Chl *a*, carotenoid content was quantified per Chl *a* amount. Fucoxanthin, the major carotenoid
183 in *C. gracilis* FCP, showed no significant differences between WT and mutants under both light
184 conditions, except for *lhcx1-15* under HL conditions (Figure 3B). The *lhcx1-15* under HL conditions

185 showed slightly lower fucoxanthin content compared to WT under the same conditions. On the other
186 hand, for diadinoxanthin (Ddx) and diatoxanthin (Dtx), the major pigments of the diatom xanthophyll
187 cycle, no significant differences in accumulation were observed between WT and *lhcx1* mutants under
188 LL conditions (Figure 3C, D). In contrast, under HL conditions, WT showed greater accumulation of
189 Dtx compared to LL conditions. The mutant strains accumulated significantly more Dtx compared to
190 the WT under the same conditions. Regarding Ddx, both WT and mutants showed greater
191 accumulation under HL conditions compared to LL conditions, but significant increases in
192 accumulation in mutants compared to WT were not observed, except for *lhcx1-15*. The de-epoxidation
193 state (DEPS = [Dtx]/([Dtx]+[Ddx])), an indicator of the xanthophyll cycle, was calculated from the
194 accumulation of Dtx and Ddx (Figure 3E). Under LL, there was no significant difference between WT
195 and mutants. Under HL, the WT showed a higher DEPS (0.34) compared to LL conditions. The
196 mutants showed significantly higher DEPS (0.47–0.52) compared to WT under the same conditions.
197

198 **Steady-State Low-Temperature Fluorescence and Antenna Size Measurements**

199 WT and *lhcx1* cultured under HL were subjected to 30-minute treatments under either dark
200 conditions (dark-treated) or HL conditions (HL-treated), then were immediately frozen in liquid
201 nitrogen, and used for the low-temperature (77K) fluorescence spectra measurements (Figure 4A, B).
202 WT was measured as two independent samples cultured separately (solid and dashed lines in Figure
203 4A, B). The 77K steady-state fluorescence spectra well represent the distribution of excitation energy
204 between photosystems and light-harvesting pigment proteins. The excitation wavelength was 459 nm,
205 which efficiently excites Chl *c* and fucoxanthin bound to the peripheral antenna FCP in diatoms. The
206 77K fluorescence spectra of *C. gracilis* showed two peaks derived from CP43 and CP47 of
207 photosystem II, observed around 686 nm and 694 nm, respectively (Andrizhiyevskaya *et al.*, 2005;
208 Nagao *et al.*, 2010). In all cells cultured under HL, the mutants showed lower CP43 peaks compared

209 to the WT in spectra normalized at the CP47 peak (Figure 4A). The fluorescence intensity around 680
210 nm, which is considered to be the fluorescence from free FCP, was smaller in *lhcx1* mutants under
211 dark conditions compared to the WT. These indicate that energy transfer from the peripheral antenna
212 FCP via CP43 is prominent in the dark, especially in WT. This also suggests that the effective antenna
213 size of PSII differed between WT and *lhcx1* in the dark. Therefore, the antenna size after dark
214 adaptation of cells cultured under HL was measured by measuring the induction curve of Chl
215 fluorescence with a Fluorometer FL 3500 (PSI, Czech Republic) (Nedbal *et al.*, 1999). As a result, it
216 was revealed that *lhcx1* had a significantly smaller effective antenna size of PSII compared to WT in
217 all strains (Supplemental Figure 7).
218
219 **Time-Resolved Low-Temperature Fluorescence Spectra**
220 To elucidate the process of excitation energy transfer from the peripheral antenna FCP to the
221 photosystems, time-resolved fluorescence (TRF) spectra were measured under the same conditions as
222 the samples subjected to low-temperature steady-state fluorescence spectrum measurements. TRF
223 fluorescence was measured on samples of whole cells that were frozen in liquid nitrogen immediately
224 after dark treatment and HL treatment (300 μmol photons m^{-2} s^{-1}). This allows observation of
225 excitation energy transfer and quenching in states where NPQ is not induced and where it is induced
226 (Chukhutsina *et al.*, 2017). Fluorescence decay-associated spectra (FDAS) were obtained by global
227 analysis that decomposed the obtained TRF spectra into multiple exponential components with time
228 as a variable (Figure 5, Supplemental Figure 8). In FDAS, positive peaks (decay) indicate the decay
229 of fluorescence, while negative peaks (rise) indicate the rise of fluorescence. From the TRF spectra of
230 the WT, FDAS were constructed with six-time constants. The assignment of each peak followed
231 previous literature (Andrizhiyevskaya *et al.*, 2005; Nagao *et al.*, 2010). In the dark-adapted WT, the
232 FDAS with a lifetime of 40 ps showed positive peaks of FCP at <680 nm and a large negative peak

233 around 687 nm corresponding to CP43 (black solid line in Figure 5 and Supplemental Figure 8). This
234 reflects excitation energy transfer from FCP to CP43. In the spectrum of the second lifetime of 140 ps,
235 positive peaks were observed not only at ~680 nm but also around 687 nm corresponding to CP43,
236 while the waveform around 694 nm corresponding to CP47 showed a convex downward shape. This
237 is thought to be due to the negative peak of CP47 overlapping with the shoulder of the nearby positive
238 peak, and is considered to indicate excitation energy transfer from FCP/CP43 to CP47. The third
239 lifetime of 640 ps continued to show the same trend as 140 ps, suggesting that excitation energy
240 transfer from FCP/CP43 to CP47 was ongoing. The fourth and fifth lifetimes, at 1.6 ns and 3.8 ns,
241 showed positive peaks corresponding to CP43 and CP47, indicating simple fluorescence decay. The
242 longest lifetime of 26 ns did not exhibit a prominent positive peak around 710 nm, which is
243 characteristic of PSI fluorescence, but instead showed positive peaks corresponding to CP43 and CP47,
244 associated with charge recombination in photosystem II. Therefore, excitation energy transfer from
245 PSII to PSI, so-called spillover, did not occur.

246 On the other hand, WT subjected to HL (black dashed line in Figure 5 and Supplemental Figure 8)
247 showed a similar waveform at 40 ps as the dark-adapted WT. At 140 ps, although the positive peak
248 around 687 nm was much larger compared to the dark-adapted WT, the value at the dip around 694
249 nm was larger than that of the dark-adapted WT. In other words, WT acclimated to HL had a smaller
250 fluorescence rise at CP47 compared to the dark-adapted sample, while the energy decay around 687
251 nm was larger. The *lhcx1* mutant, which does not induce NPQ, showed spectra similar to the dark-
252 adapted WT in the first and second time constants when dark-adapted (blue, green, and red solid lines
253 in Figure 5 and Supplemental Figure 8A, B). However, in the third time constant (640 ps), while the
254 dark-adapted WT showed excitation energy transfer from FCP/CP43 to CP47, *lhcx1* already showed
255 positive peaks at both 687 nm and 694 nm, indicating fluorescence decay of CP43 and CP47. In
256 subsequent time constants, 687 nm and 694 nm also showed only positive peaks. The HL-acclimated

257 *lhcx1* (blue, green, and red dashed lines in Figure 5 and Supplemental Figure 8A, B) also showed
258 spectra similar to the dark-adapted *lhcx1* at all time constants and did not show the large positive peak
259 around 687 nm at 140 ps observed in the HL-acclimated WT. Furthermore, since energy transfer from
260 FCPII to CP43 and CP47, and from CP43 to CP47, was completed by 140 ps, similar to the HL-treated
261 WT and dark-treated *lhcx1*, it was concluded that *lhcx1* has a smaller antenna size compared to the
262 WT even under HL conditions. From the above, while the decay of fluorescence around 687 nm is
263 normally attributed to energy transfer from CP43, when looking at the energy transfer of *lhcx1* at the
264 same time constant, such decay is not observed. Therefore, it is concluded that the decay of
265 fluorescence around 687 nm in the HL-acclimated WT is not decay of CP43 but quenching attributed
266 to Lhcx1.
267
268 **Molecular interaction of Lhcx1 in thylakoid membranes**
269 To determine the localization of Lhcx, we performed native-PAGE followed by second-dimension
270 SDS-PAGE and subsequent immunoblotting using anti-Lhcx1 antibody. Since conventional Clear
271 Native (CN)-PAGE with deoxycholate made it difficult to separate diatom $C_2S_2M_2$ (Core$_2$–S-
272 tetramer$_2$–M-tetramer$_2$) PSII–FCPII while keeping FCP-A (which contains CgLhcf1 as a major
273 component) bound to PSII (Nagao *et al.*, 2012), we applied a modified CN-PAGE using Amphipol
274 A8-35 (Kameo *et al.*, 2021), which enhances the stability of solubilized protein complexes, instead of
275 deoxycholate as the surfactant that imparts negative charge to proteins (Figure 6A). In the diatom *C.*
276 *gracilis*, a dense band was observed at a higher molecular weight position than spinach $C_2S_2M_2$
277 (Core$_2$–S-trimer$_2$–M-trimer$_2$) PSII–LHCII. Additionally, a broad brown band, which was observed at
278 a higher molecular weight position than the LHC/FCP monomer, extended to a lower molecular weight
279 position than the spinach LHCII trimer. This CN-PAGE gel strip was subjected to SDS-PAGE with a
280 gel containing 6 mol L^{-1} urea to separate each subunit (Figure 6B). Annotation of each spot was based

281 on previous reports (Nagao *et al.*, 2007; Zhou *et al.*, 2024). The amino-acid sequence of
282 HGRIAQLAFLGN shared with CgLhcf2 and CgLhcf12 was identified from the major FCP fraction
283 (Ishihara *et al.*, 2015). These reports indicate that FCP-B/C is one of the major FCPs in *C. gracilis*.
284 Furthermore, since the FCP-B/C heterodimer is presumed to be loosely bound to $C_2S_2M_2$ PSII–FCPII
285 and transfer excitation energy, it was designated as L (Loose)-dimer (Figure 6A, B). Additionally,
286 immunoblotting was performed to identify the Lhcx1 spot from the spots in the two-dimensional SDS-
287 PAGE (Figure 6C). The Lhcx1 spot, indicated by the blue arrowhead in Figure 6C, was detected at a
288 higher molecular weight than FCP-B/C in 2D-SDS-PAGE. In the 1D-CN-PAGE direction, the spot of
289 Lhcx1 was detected across a wide range of multimers from monomers to higher-order multimeric
290 complexes, co-migrating with FCP-B/C (Figure 6B). This result suggests that Lhcx1 may interact with
291 the FCP-B/C dimers, functioning as the loosely associated antenna of $C_2S_2M_2$ PSII–FCPII.
292
293 **Phenotype of *lhcx1* Mutants in High CO_2 Environment**
294 We investigated whether the *lhcx1* mutants, which showed a higher effective quantum yield of PSII
295 when acclimated to HL under normal atmospheric aeration conditions, would further improve their
296 PSII effective quantum yield by enhancing downstream electron sink capacity with 3% CO_2 addition.
297 First, even when cultured with CO_2 aeration, *lhcx1* showed almost no NPQ regardless of light intensity
298 (Figure 7A). The effective quantum yield of PSII, Y(II), increased in *lhcx1* when cultured with CO_2
299 aeration even under LL conditions, where no difference was observed with atmospheric aeration (solid
300 line in Figure 7C). This Y(II) under light irradiation increased with time during the initial irradiation
301 period and reached a constant value after about 2 minutes, which is thought to be due to the activation
302 of the carbon assimilation pathway. Additionally, an increase in Y(II) was also observed under HL
303 conditions. Photochemical quenching qP, which indicates the redox state of plastoquinone, was similar
304 between WT and *lhcx1* under LL conditions. As observed in Figure 2B, qP of *lhcx1* varied among

305 strains, with strains having lower Y(II) showing lower qP under HL conditions (Figure 7B).

306 To understand the general physiological response of *lhcx1*, we performed RNA sequencing analysis
307 under LL/HL and ambient/3%-CO_2-supplemented air conditions. Under ambient air conditions, HL
308 treatment induced 1,494 genes in WT and 1,530 genes in *lhcx1* mutant, with 458 genes commonly
309 upregulated in both genotypes. Under CO_2-enriched conditions, HL induced substantially more genes:
310 2,411 in WT and 2,402 in *lhcx1* mutant. To identify responses compensating the absence of NPQ in
311 *lhcx1*, we compared the magnitude of HL response (fold-change from LL to HL) between genotypes.
312 *lhcx1* showed enhanced HL response compared to WT in 978 genes under ambient air and 311 genes
313 under CO_2-enriched conditions, with 105 genes showing this enhanced response under both ambient
314 and CO_2 conditions. KEGG pathway enrichment analysis reveals that while both WT and *lhcx1*
315 respond to HL through porphyrin metabolism and photosynthesis pathways (Figure 8A), *lhcx1* mutant
316 specifically enhances carbon fixation and central carbon metabolism pathways under both ambient air
317 and CO_2 conditions (Figure 8B).

318 **Discussions**

319 The deficiency of Lhcx, while Lhcx belongs to the LHC family primarily responsible for light
320 harvesting, almost completely suppresses the induction of non-photochemical quenching (NPQ). The
321 NPQ-deficient *lhcx1* strains acclimated by inducing various HL defense systems possessed by *C.*
322 *gracilis* and showed higher PSII activity compared to the WT. Since the increase in PSII activity was
323 also induced in *lhcx1* with the addition of CO_2, it was revealed that acclimation of downstream electron
324 sinks in the electron transport chain, along with acclimation of the upstream light-harvesting system,
325 plays an important role in HL acclimation of *C. gracilis*. This study suggests that the multiple
326 photoprotection mechanisms can be interrelated and flexibly regulated in response to environmental
327 changes in the marine diatom *C. gracilis*.

328

329 **Origin of Lhcx**

330 The Lhcx subfamily, which is responsible for non-photochemical quenching (NPQ) in red-lineage
331 secondary endosymbiotic algae such as heterokonts including diatoms, and presumably haptophytes,
332 and the Lhcsr subfamily, which similarly mediates NPQ in green algae, streptophyte algae, and
333 bryophytes, have been considered to share a common ancestral gene as they form a monophyletic
334 group (Koziol *et al.*, 2007). Although the Lhcx/Lhcsr subfamily is suggested to be derived from the
335 red lineage, detailed molecular phylogenetic relationships with other subfamilies have not been
336 thoroughly investigated, likely due to insufficient sequence information (Dittami *et al.*, 2010). In this
337 study as well, the diatom Lhcx subfamily and the basal streptophyte algae and core green algae Lhcsr
338 subfamily formed a monophyletic group (Figure 1). The Lhcx/Lhcsr clade formed by these showed
339 strong support for monophyly with the diatom Lhcq, CgLhcr9 homolog, and Lhcf subfamilies, and
340 was in a different clade from the green algae Lhca/Lhcb subfamily, clearly indicating that it is derived
341 from red-lineage algae. In particular, since the Lhcx/Lhcsr clade showed monophyly with the CgLhcr9

342 homolog and Lhcf subfamily, it was demonstrated that the Lhcx subfamily is one of the new LHC
343 subfamily groups that arose in heterokont algae along with Lhcq, CgLhcr9 homolog, and Lhcf
344 subfamilies. Furthermore, since Lhcsr is conserved in both core green algae and streptophyte algae,
345 this supports the hypothesis proposed in Dittami et al. that the ancestor of extant green plants acquired
346 the Lhcx/Lhcsr subfamily through horizontal gene transfer from a red-lineage alga, presumably a
347 heterokont alga or haptophyte, before the last common ancestor of extant green plants. The estimated
348 divergence time of extant green plants into Streptophyta and Chlorophyta overlaps with the emergence
349 of heterokonts (Strassert *et al.*, 2021), indicating that gene transfer of Lhcx/Lhcsr from heterokonts to
350 ancestral green plants could have occurred prior to this divergence event. It is difficult to determine
351 whether this gene transfer was merely a horizontal gene transfer or involved the acquisition of a
352 temporary red-lineage plastid, such as a kleptoplast. Additionally, looking at Figure 1, the
353 monophyletic clade of the green algae Lhca and Lhcb subfamilies does not form a monophyletic group
354 with the red algae Lhcr subfamily, and between them exist cryptophyte Lhcr and diatom Lhcr. This
355 does not negate the possibility that not only Lhcsr but also the Lhca and Lhcb subfamilies were
356 acquired through horizontal gene transfer. This result calls for reconsideration of the formation and
357 establishment of early chloroplasts.

358 Within Lhcx, Lhcx6_1, which exists at the most basal position, is present in *T. pseudonana* and
359 *Cyclotella cryptica* but not in *C. gracilis* or *Phaeodactylum tricornutum* (Figure 1, inset). This
360 Lhcx6_1 not only has an evolutionary distance from other typical Lhcx in the molecular phylogenetic
361 tree but also does not show a low NPQ phenotype when deficient (Nakayasu *et al.*, 2024). Furthermore,
362 Lhcx6_1 directly binds to PSII and has a completely different assembly from *C. gracilis* PSII–FCPII
363 (Feng *et al.*, 2023; Zhao *et al.*, 2023). Based on these observations, it is considered that Lhcx6_1 is an
364 atypical Lhcx and does not mediate NPQ, and we argue that its classification within the Lhcx
365 subfamily should also be reconsidered.

366

367 **Spectroscopic Characteristics of Lhcx Quenching**

368 In this study, we successfully obtained NPQ signals at 140 ps around 687 nm by measuring time-
369 resolved fluorescence comparing frozen whole cells immediately after dark adaptation and HL
370 irradiation, reflecting the *in vivo* state (Figure 5, Supplemental Figure 7). By using the *lhcx1* mutant
371 that shows a clear phenotype with almost no NPQ induction, and by optimizing the dark adaptation
372 and HL exposure time to 30 minutes, we were able to clearly distinguish signals attributable to Lhcx-
373 mediated NPQ.

374 Our results contrast with several previous studies on diatoms. Similar analyses were performed for
375 the diatoms *C. meneghinia* and *P. tricornutum* (Miloslavina *et al.*, 2009). A different approach was
376 taken regarding DAS fitting compared to this study, with strong assumptions placed on the
377 fluorescence signals of PSII and PSI, and quenching signals like those observed in this study were not
378 obtained. More detailed observations of quenching using time-resolved fluorescence have been
379 attempted in the diatom *C. meneghiniana*, but the ~200 ps signal around 685 nm observed in this study
380 was not observed; instead, a ~600 ps quenching signal at 695 nm was observed and attributed to PSII
381 (Chukhutsina *et al.*, 2014). In a similar approach using WT, a 214 ps signal at 685 nm in
382 *Nannochloropsis gaditana* was attributed to LHCII quenching, but it was not clear whether this was a
383 quenching signal due to the presence of Lhcx (Chukhutsina *et al.*, 2017). A similar approach has also
384 been used with the *lhcx1* knockdown mutant of *P. tricornutum*, but this mutant showed a limited
385 decrease in NPQ, and no clear NPQ signal was observed in time-resolved fluorescence (Taddei *et al.*,
386 2018).

387 The spectroscopic characteristics of NPQ in diatoms differ notably from those in green algae. In
388 green alga *Chlamydomonas reinhardtii stt7-9* state-transition-less mutants, NPQ signals were
389 observed at 150 ps in time-resolved fluorescence and ~100 ps in transient absorption at physiological

390 temperature, similar to our time constant, but the NPQ signal from fluorescence was detected at 682
391 nm and attributed to LHCII fluorescence (Tian *et al.*, 2019; Zheng *et al.*, 2024). This wavelength
392 difference (682 nm vs. 687 nm) likely reflects differences in measurement temperatures and distinct
393 pigment compositions between the two algal groups, rather than fundamental mechanistic differences.
394 Time-resolved fluorescence measurements in whole cells revealed longer excitation energy transfer
395 times compared to isolated complexes. While excitation energy transfer from FCPII to the PSII core
396 in $C_2S_2M_2$ PSII–FCPII occurs up to a time constant of 210 ps (Nagao *et al.*, 2022), our measurements
397 in dark-adapted whole cells showed a time constant of 640 ps (Figure 5, Supplemental Figure 8). This
398 threefold increase suggests that PSII–FCPII complexes *in vivo* possess larger antenna systems than the
399 isolated $C_2S_2M_2$ form, indicating the presence of additional FCPs connected to the excitation energy
400 network. These peripheral antenna complexes, likely including L-FCPII dimers, extend the effective
401 antenna size and modify energy transfer kinetics.
402
403 **The mechanism of Lhcx1-mediated NPQ**
404 The localization of diatom Lhcx has been debated for many years. Except for Lhcx6_1 associated
405 with PSII core (Feng *et al.*, 2023; Zhao *et al.*, 2023), the localization of Lhcx has been detected in free
406 FCP or bulk FCP (Lepetit *et al.*, 2010; Grouneva *et al.*, 2011), except for reports that *P. tricornutum*
407 Lhcx1/Lhcx2 bind to PSI (Grouneva *et al.*, 2011). Additionally, in SDS-PAGE following CN-PAGE,
408 *T. pseudonana* Lhcx1 was detected as a free monomer (Zhou *et al.*, 2024). In the 2D-CN/SDS-PAGE
409 of the diatom *C. gracilis* by Zhou et al. (2024), Lhcx1 was not detected. In this study, we utilized
410 Amphipol A8-35, a surfactant that enhances the stability of membrane proteins, to enable the
411 separation of diatom $C_2S_2M_2$ PSII–FCPII, which had been previously challenging with CN-PAGE
412 methods (Figure 6A). Under this condition, the dimer (L-dimer) consisting of CgLhcf12/CgLhcf2
413 (FCP-B/C) showed a broad band in the first dimension (Figure 6B), and a similar band pattern was

414 observed in the immunoblot for CgLhcx1 (Figure 6C). These findings suggest that CgLhcx1 may
415 interact with the FCP L-dimer.

416 Assuming the interaction between L-FCPII dimers and Lhcx1, we propose that Lhcx1-mediated
417 NPQ operates through two possible mechanisms: either Lhcx1 directly functions as the quenching
418 center, or quenching occurs through the L-dimer via its interaction with Lhcx1. This mechanism can
419 be classified within the framework of Q1 and Q2 quenching, where Q1 occurs in detached antenna
420 complexes and Q2 occurs in minor LHCs bound to the PSII core (Miloslavina *et al.*, 2009; Holzwarth
421 *et al.*, 2009; Chukhutsina *et al.*, 2014). Given that L-FCPII dimers bind loosely to S/M-FCPII, Lhcx1-
422 mediated quenching likely represents Q1-type quenching occurring in the peripheral light-harvesting
423 antenna, although it is unclear whether the L-FCPII dimer detaches from PSII-S/M-FCPII.

424 The 140 ps quenching time constant observed in HL-treated whole cells is notably faster than
425 previously reported values for isolated FCP complexes. Isolated FCP-A tetramers and FCP-B/C dimers
426 have been reported to show characteristic quenching with time constants of 270 ps and 200 ps,
427 respectively (Nagao *et al.*, 2013). This acceleration in whole cells may result from two factors: first,
428 isolated FCP-B/C complexes lack Lhcx proteins, which appear crucial for efficient quenching; second,
429 isolated complexes lack the native membrane environment, including the ΔpH gradient and three-
430 dimensional protein organization necessary for optimal quenching function. These results highlight
431 the importance of studying NPQ mechanisms in intact cellular systems rather than relying solely on
432 isolated protein complexes.

433

434 **Compensatory mechanisms to the Lhcx1 Deficiency**

435 The *lhcx1* strain, which lacks the most highly expressed Lhcx1 in *C. gracilis* created using genome
436 editing, induced almost no NPQ and showed higher PSII effective quantum yield and oxygen-evolving
437 activity compared to the WT when acclimated to HL. Since NPQ mutants in land plants show

438 decreased PSII electron transport activity, i.e., decreased effective quantum yield, under HL
439 acclimation (Yang *et al.*, 2022), it is unusual that *lhcx1*, an NPQ mutant in this study, shows a higher
440 PSII effective quantum yield. Since *lhcx1* has a smaller antenna size even when dark-adapted after HL
441 acclimation (Figure 5, Supplemental Figure 8), the decrease in excitation energy acquisition per PSII
442 core is thought to contribute to a higher PSII effective quantum yield. Additionally, since the PSII
443 effective quantum yield increased even under LL conditions with the addition of CO_2 (Figure 7), and
444 transcriptomic analysis revealed that *lhcx1* specifically enhances carbon fixation and central carbon
445 metabolism pathways under both ambient air and CO_2 conditions (Figure 8B), it is likely that the
446 increase in carbon fixation capacity in the downstream electron transport chain also contributed to the
447 improvement of PSII effective quantum yield. The activation of downstream electron sinks can be
448 observed under HL conditions (Figure 2C).

449 HL acclimation in *lhcx1* was also suggested by the pigment analysis (Figure 3C–E). Chl content per
450 cell decreased in *lhcx1*, and while the amount of Fx per Chl did not change, the amounts of
451 xanthophylls Dtx and Ddx increased. The de-epoxidation state (DEPS) calculated from Dtx and Ddx
452 also increased in *lhcx1* under HL, indicating that acclimation responses under HL are enhanced from
453 the perspective of pigment composition as well. Generally, in WT diatoms, the amount of Dtx is very
454 well correlated with NPQ intensity (Lavaud & Lepetit, 2013). However, since *lhcx1* strains show
455 almost no induction of NPQ despite increased Dtx amounts, the correlation between Dtx amount and
456 NPQ requires the presence of Lhcx. Although it is not clarified whether the increased xanthophylls in
457 *lhcx1* are incorporated in LHC/FCP, they may function as antioxidants in thylakoid membranes and
458 contribute to photoprotection. The above multiple photoprotection mechanisms that became apparent
459 in the absence of Lhcx1 would contribute to the adaptation in variable marine environments.

460 **Materials and Methods**

461 **Phylogenetic Analysis of LHCs Across Various Photosynthetic Eukaryotic Lineages**

462 LHC sequences were collected from genomes and transcriptomes included in the Phycocosm and
463 EukProt databases, in addition to our previous dataset (Kumazawa *et al.*, 2022; Kumazawa & Ifuku,
464 2024), following the previous method (Kumazawa & Ifuku, 2024). The LHCs in the dataset were
465 obtained from red algae, early-diverging streptophyte green algae (Leliaert *et al.*, 2012): *Mesostigma*
466 *viride* NIES-296 (Liang et al., 2020) and *Klebsormidium nitens* NIES-2285 (Hori *et al.*, 2014),
467 diatoms: *C. gracilis*, *Thalassiosira pseudonana*, *Cyclotella cryptica*, *Phaeodactylum tricornutum*, and
468 cryptophytes: *Guillardia theta* CCMP2712 (Curtis *et al.*, 2012) and Cryptophyceae sp. (*Baffinella* sp.)
469 CCMP2293 (Dorrell *et al.*, 2023). As reference sequences for Lhcsr, *Chlamydomonas reinhardtii*
470 Lhcsr1 (UniProt: P93664) and Lhcsr3.1 (UniProt: P0DO19), which is identical to Lhcsr3.2 (UniProt:
471 P0DO18), were added. Redundant sequences were removed from the LHC protein sequences using
472 CD-HIT (Fu *et al.*, 2012), and alignment was performed with mafft-einsi v7.525, followed by manual
473 curation (Katoh & Standley, 2013). The curated dataset was realigned, and trimming was performed
474 using ClipKit v2.3.0 in smart-gap mode to remove unaligned sites (Steenwyk *et al.*, 2020). The
475 phylogenetic tree was estimated using IQ-TREE v2.3.6 with the Q.pfam+I+R7 model selected by
476 ModelFinder based on LG and Q.pfam models (Kalyaanamoorthy *et al.*, 2017; Minh *et al.*, 2020,
477 2021), and visualized with iTOL v6 (Letunic & Bork, 2024). Ultra-fast bootstrap approximation
478 (UFBoot) and SH-aLRT support (%) tests were performed with 1000 replicates each (Guindon *et al.*,
479 2010; Hoang *et al.*, 2018). The aBayes support test was also conducted (Anisimova *et al.*, 2011). The
480 annotation of LHC subfamilies in the phylogenetic tree was based on our previous studies (Kumazawa
481 *et al.*, 2022; Kumazawa & Ifuku, 2024) and drawn on the tree using InkScape software.

482

483 **Cultivation**

484 The marine centric diatom *C. gracilis* (UTEX LB 2658) was used for all analyses. Pre-cultivation
485 was performed in IMK medium —4% (w/v) sea salt (SIGMA S9883, US) or artificial seawater made
486 with Daigo's Artificial Seawater SP (Nihon Pharmaceutical MS, Japan) supplemented with 1x Daigo's
487 IMK medium (Nihon Pharmaceutical MS, Japan) and 0.2 mmol L^{-1} Na_2SiO_3— at 20ºC under
488 fluorescent lights at approximately 30 µmol photons m^{-2} s^{-1} with continuous shaking at 100 rpm.
489 Seven–eight-day or three-day cultivation was performed in IMK medium at 25ºC under warm white
490 LEDs at 30 µmol photons m^{-2} s^{-1} (LL) or 300 µmol photons m^{-2} s^{-1} (HL), with continuous aeration
491 of air or 3% (v/v) CO_2.

492

493 **Preparation of Thylakoid Membranes from *C. gracilis* and Spinach**

494 Diatom WT cells cultured under HL at 25ºC, 300 µmol photons m^{-2} s^{-1} for three days were
495 centrifuged at 1,200×*g*, 20ºC for 5 minutes to collect the cells, then resuspended in artificial seawater,
496 washed, and centrifuged again. After removing the supernatant, the pellet was resuspended in 1 mL of
497 artificial seawater, transferred to a 1.5 mL microcentrifuge tube, and centrifuged at 21,500×*g*, 4ºC for
498 5 minutes to obtain a pellet. The pellet was resuspended in 1 mL of cell disruption buffer (20 mM
499 Tricine-KOH (pH 7.6), 0.4 mol L^{-1} sorbitol, 10 (w/v) % PEG-6000, 10 mM EDTA), mixed with 0.5
500 g of 0.1 mm diameter glass beads, and processed with a bead beater for 10 seconds. The sample was
501 then well-chilled on ice, and the disruption process was repeated a total of three times. The supernatant
502 was centrifuged at 21,500×*g*, 4ºC for 5 minutes to obtain a pellet. The pellet was resuspended in 1 mL
503 of washing buffer (20 mmol L^{-1} Tricine-KOH (pH 7.6), 0.4 mol L^{-1} Sorbitol, 5 mmol L^{-1} $MgCl_2$, 2.5
504 mmol L^{-1} EDTA) and centrifuged at 50×*g*, 4ºC for 5 minutes. The portion containing thylakoid
505 membranes was carefully collected, avoiding the upper and lower contaminants, and centrifuged at
506 21,500×*g*, 4ºC for 5 minutes to obtain a thylakoid membrane pellet. The pellet was resuspended in

507 solubilization buffer (50 m mol L^{-1} imidazole-HCl (pH 7.0 at 4°C), 20 (v/v) % glycerol), and a portion
508 was taken to measure the concentrations of Chl *a* and *c* (Jeffrey & Humphrey, 1975). The thylakoid
509 membrane of spinach (*Spinacia oleracea*) was isolated from market spinach based on previously
510 reported methods (Yamamoto *et al.*, 2011) using the modified buffer (50 mmol L^{-1} tricine-NaOH (pH
511 7.8 at room temperature), 0.4 L^{-1} sucrose, 10 mmol L^{-1} NaCl, 5 mmol L^{-1} $MgCl_2$). The thylakoid
512 membranes were adjusted to 1 μg μL^{-1} of Chl *a*+*c* with solubilization buffer (50 mM imidazole-HCl
513 (pH 7.0 at 4°C), 20 (v/v) % glycerol).
514
515 **2D-CN/SDS-PAGE**
516 The 4–13% gradient gel used for CN-PAGE was prepared according to existing methods (Järvi et
517 al., 2011; Kameo et al., 2021), and the entire CN-PAGE procedure followed the modified CN-PAGE
518 method using Amphipol A8-35 (Kameo et al., 2021). However, 75 mmol L^{-1} imidazole-HCl (pH 7.0
519 at 4ºC) was used in the 3x gel buffer, and 40 (w/v) % acrylamide/bis solution (37.5:1) (019-25655,
520 Fujifilm Wako, Japan) was used for acrylamide. The anode buffer contained 25 mmol L^{-1} imidazole-
521 HCl (pH 7.0 at 4ºC), and the cathode buffer contained 50 mmol L^{-1} Tricine, 7.5 mmol L^{-1} imidazole.
522 The Chl concentrations in resuspended thylakoids were measured (Jeffrey & Humphrey, 1975), and
523 thylakoid membranes were adjusted to 1 μg μL^{-1} of Chl *a*+*c* and Chl *a*+*b* with solubilization buffer
524 (50 mmol L^{-1} imidazole-HCl (pH 7.0 at 4°C), 20 (v/v) % glycerol) and resuspended. An equal volume
525 of 2 (w/v) % dodecyl-α-D-maltoside (α-DM, Anatrace, OH, USA) was added for solubilization, placed
526 on ice for 2 minutes, then centrifuged at 21,500×*g* for 2 minutes to remove insoluble residues. The
527 supernatant was mixed with half the volume of 2 (w/v) % amphipol A8-35 (Anatrace, OH, USA), and
528 5 μg of Chl in 15 μL was applied to each well. Electrophoresis was performed using MyPowerII 300
529 power supply (AE-8135, ATTO, Japan) at 40 V, 1 mA for 660 minutes at 4ºC with two mini gels
530 simultaneously. After electrophoresis, the gels were scanned with the scanner ES2200 (EPSON, Japan).

531 The CN-PAGE gels were immersed in 50 (v/v) % glycerol on ice, cut into gel strips, and stored at −
532 80ºC in 50 (v/v) % glycerol until use.

533 Gel strips were solubilized by immersing them in solubilization buffer (1% SDS and 50 mM
534 dithiothreitol (DTT) for 30 minutes at room temperature, and then applied to gels with a stacking gel
535 (pH 6.8) and separation gel (14% acrylamide (37.5:1), 6 mol L^{-1} Urea, pH 8.6). The Laemmli system
536 was used for electrophoresis. When the dye had electrophoresed to the bottom of the marker well at
537 250 V, 100 mA, the process was temporarily stopped, and Precision Plus Protein Dual Color Standards
538 (#1610374, BioRad, CA, USA) were applied. Electrophoresis was then continued at 250 V for 90
539 minutes. After electrophoresis, gels not used for immunoblotting were fluorescently stained with
540 Oriole fluorescent gel stain (#1610496, BioRad, CA, USA) for 100 minutes, gently rinsed with
541 distilled water, and then fluorescence was detected using the ChemiDoc Touch Imaging System (Bio-
542 Rad).

543

544 **Construction of hCas9 Vector**

545 The hCas9 sequence was amplified from the hCas9 vector (Addgene, plasmid #41815) using the
546 InF-hCas9-fw1 primer (TTTATAAAGCGGATCATGGACAAGAAGTACTCCATTGG) and InF-
547 hCas9-rv1 primer (AACAGCTTGCATGCCTCACACCTTCCTCTTCTTCTTG). The PCR product
548 was cloned into the linearized pCgNRp vector digested with BamHI and PstI using In-Fusion reaction
549 according to the manufacturer's protocol. The obtained clone vector was named pCgNRphCas9.

550

551 **Construction of pCgNRpble-Lhcx1-RGR Vector**

552 A 20-mer genome editing target sequence containing the PAM (protospacer adjacent motif)
553 sequence (NGG) for gRNA in the CRISPR/Cas9 system, GTTGCTATGCTTGCCGTTAT**TGG** (bold
554 indicates the PAM sequence), was selected. A system that expresses complete gRNA in cells via a

555 mRNA-type Pol II promoter by connecting self-processing ribozymes (Gao & Zhao, 2014; Zhang *et*
556 *al.*, 2017) was used. This method was adapted to the genome editing system of *C. gracilis*. The genome
557 editing vector containing gRNA was constructed using three primers: Lhcx1-RGR-F1;
558 GACGAAACGAGTAAGCTCGTC**GTTGCTATGCTTGCCGTTAT**GTTTTAGAGCTAGAAATA
559 GCAAG (bold indicates the 20-mer sequence for gRNA target recognition excluding the PAM
560 sequence), Lhcx1-RGR-F2;
561 GTCAGGATCC**AGCAAC**CTGATGAGTCCGTGAGGACGAAACGAGTAAGCTCGTC (BamHI
562 restriction enzyme site is underlined, and bold indicates the complementary sequence of the first six
563 bases of the 20-mer for gRNA target recognition including the PAM sequence), and PstI-RGR-Rv;
564 ATATCTGCAGGTCCCATTCGCCATGCCGAAGC (PstI restriction enzyme site is underlined). The
565 PstI-RGR-Rv primer is a universal primer in this protocol. The first fragment (199 bp) was amplified
566 by PCR from pRS316-RGR-GFP (Addgene, plasmid #51056) using the Lhcx1-RGR-F1 and PstI-
567 RGR-Rv primers. The second fragment (231 bp) was amplified by PCR from the first purified PCR
568 product using the Lhcx1-RGR-F2 and PstI-RGR-Rv primers. The purified second product was
569 digested with BamHI (NEB) and PstI (NEB), and ligated with the pCgNRpble vector linearized with
570 BamHI and PstI. The final product was named pCgNRp^{ble}-Lhcx1-RGR.
571

572 **Transformation**

573 The electroporation transformation procedure using the Super electroporator NEPA21 Type-II
574 (NEPAGENE, Japan) followed previous reports (Miyahara *et al.*, 2013; Ifuku *et al.*, 2015; Ifuku &
575 Yan, 2020). The initial transformation of the *hCas9* gene into *C. gracilis* was performed using the
576 pCgNRphCas9 vector linearized with DraIII and SapI. Colonies were selected on IMK plates
577 containing the antibiotic Nourseothricin at 400 μg mL^{-1} (clonNAT; Hans Knoell Institute, Jena,
578 Germany). The introduced *hCas9* gene was confirmed by PCR using the first set: InF-hCas9-fw1

579 primer and hCas9-2088-Rv primer: CCTTAAAGGTGAGAGAGTCATCATGGATC, and another set:
580 hCas9-1965n-Fw primer: ATGAAACAGCTCAAGAGGCGCCGA and InF-hCas9-rv1 primer. The
581 hCas9 strain (hC9 strain) used for further genome editing was selected from eight transformants
582 screened by PCR and immunoblotting using anti-Cas9 antibody (described later).

583 The linearized pCgNRpble-Lhcx1-RGR was introduced into the hC9 strain by electroporation. This
584 was essentially the same as the transformation of hCas9. Before introducing the RGR vector, the vector
585 was linearized with HindIII (NEB) and concentrated to 850–1000 ng μL^{-1} by ethanol precipitation.
586 The amount of linearized vector was 5 μg per electroporation. Colonies were selected on IMK plates
587 containing the antibiotic Zeocin at 100 μg mL^{-1} (InvivoGen #ant-an-1, US CA). Mutations were
588 confirmed by Sanger sequencing of the products amplified by PCR using Lhcx1_seq_FW primer:
589 AGAGACTGTAAGGATTCGGA and Lhcx1_seq_RV primer: GAAGGAAGATCCCTCTACCT.

590

591 **Immunoblotting**

592 Diatom cells were collected and solubilized in sample buffer containing 62.5 mmol L^{-1} Tris–HCl
593 (pH 6.8), 2.5% (w/v) SDS, 10% (w/v) glycerol, 2.5% (v/v) 2-mercaptoethanol, and a trace amount of
594 bromophenol blue (BPB). Solubilization was performed by incubating at 98ºC for 6 minutes, followed
595 by 30 minutes at room temperature, and then centrifugation. The protein concentration was adjusted
596 to 30 μg/20 μL. Thirty micrograms of protein was applied to each lane of an 8% acrylamide separation
597 gel with a 3% acrylamide stacking gel and separated by SDS-PAGE. Proteins were transferred to an
598 Immobilon-P PVDF membrane (Millipore, MA, USA) and incubated with anti-Cas9 polyclonal
599 antibody (632607, Takara, Japan) at a 1:2000 dilution.

600 For the detection of Lhcx1, SDS-PAGE gels were prepared with a 15% acrylamide separation gel
601 and a 4.5% acrylamide stacking gel. Two μg of protein was applied to each lane and separated by SDS-
602 PAGE. The anti-Lhcx1 peptide antibody was produced against the synthetic peptide

603 "KEEYAPGDLRFDPFGLMP". After electrophoresis, the gel was desalted and equilibrated in
604 Electron transfer buffer (25 mmol L^{-1} Tris, 192 mmol L^{-1} glycine, 20 (v/v) % methanol, 0.05 (w/v) %
605 SDS). Subsequently, proteins were transferred from the gel to an Immobilon-P PVDF membrane
606 (Millipore, Billerica, MA, USA) and incubated with anti-Lhcx1 antibody at a 1:5000 dilution.
607 Immunoreactive proteins were detected using the Amersham ECL Prime Western Blotting Detection
608 Reagent (Cytiva RPN2236) with the ChemiDoc Touch Imaging System (Bio-Rad).
609
610 **PAM Fluorescence Measurement**
611 Chl fluorescence by pulse-amplitude modulation was detected using an actinic light of 450 or 455
612 nm (blue) at 300 µmol photons m^{-2} s^{-1} with the NPQ2 protocol (200 seconds of light period and 390
613 seconds of dark recovery, with the first saturating pulse light in the light period at 10 seconds after the
614 start of illumination) of AquaPen AP-100 or AP-110 (PSI, Czech Republic) using a bandpass filter
615 (667 to 750 nm). The measuring pulse was set to 10%, and the super pulse (saturating pulse light) was
616 set to 30% (equivalent to 900 µmol photons m^{-2} s^{-1}). Cells used in the NPQ2 protocol were dark-
617 adapted for more than 10 minutes before measurement.
618
619 **Oxygen-Evolving Activity Measurements**
620 The diatom cells grown under HL conditions (200 µmol photons m^{-2} s^{-1} with rotational shaking at
621 ~22ºC) were adjusted to a concentration of 5 µg Chl *a* mL^{-1} (Jeffrey & Humphrey, 1975), following
622 the methods described in Nakayasu et al. (Nakayasu *et al.*, 2024). Using a Clark-type oxygen electrode
623 (Hansatech, UK) at 25°C, oxygen evolution was measured in the medium supplemented with 10 mM
624 $NaHCO_3$. Cells were illuminated with white light at 300 µmol photons m^{-2} s^{-1}, and the cell
625 suspension in the reaction chamber was continuously stirred with a magnetic stirrer during
626 measurements.

627

628 **HPLC Pigment Analysis**

629 Cells cultured for 7–8 days under LL or HL conditions were collected by centrifugation at 1,500 x
630 *g* for 5 minutes, and the supernatant was removed. The pellet was immediately frozen in liquid nitrogen
631 and stored at −80 ºC until extraction. The frozen pellet was extracted using 100% acetone (Wako,
632 Japan) with sonication for 6 minutes in an ice-water bath. Subsequently, the sample was centrifuged
633 at 15,000 rpm for 5 minutes at 4 ºC. The resulting supernatant was filtered through a 0.45 nm
634 hydrophilic PTFE filter (SLPT0445NL, Hawach Scientific, IL, USA) and subjected to HPLC. The
635 HPLC method is based on Nagao, Ueno, et al. (2020) with modifications derived from Manuel Zapata
636 et al. (2000). The 0.25 mol L^{-1} pyridine contained in solution A was adjusted to pH 5.0 using acetic
637 acid. A Shimadzu HPLC system equipped with LC-20AD pumps and an SPD-M20A detector was
638 used for pigment composition analysis. The system was equipped with an Inertsil C8 reverse-phase
639 column (5020-01228, GL Sciences, Japan). The sample injection volume was 20 µL for each analysis.
640 Standard pigments of Chl *a*, fucoxanthin, diadinoxanthin, and diatoxanthin were purchased from DHI
641 (Denmark), analyzed similarly, and used to create calibration curves.

642

643 **77K Steady-State Fluorescence Spectra**

644 Cells for measurement were cultured under HL at 300 µmol photons m^{-2} s^{-1} at 25ºC for 3 days.
645 Samples were diluted with IMK medium and adjusted to 10 µg Chl *a*+*c*/150 µL. Chl concentration
646 was measured after extraction with 90% acetone (Jeffrey & Humphrey, 1975). The diluted samples
647 were treated in the dark or under HL at 300 µmol photons m^{-2} s^{-1} for 30 minutes and immediately
648 frozen in liquid nitrogen. Low-temperature steady-state fluorescence emission spectra of whole cells
649 were measured at 77K with excitation at 459 nm using a JASCO FP-8500 equipped with a PMU-130
650 liquid nitrogen cooling unit. The spectra were integrated from 5 consecutive measurements with a 2.5

651 nm sampling pitch. The spectra were normalized at the PSII CP47 peak (693–695 nm)
652 (Andrizhiyevskaya *et al.*, 2005; Nagao *et al.*, 2010).
653
654 **Time-Resolved Fluorescence and Fluorescence Decay-Associated Spectra**
655 Samples were prepared for time-resolved fluorescence (TRF) spectral measurements using the same
656 method as for steady-state measurements. TRF spectra were measured using a time-correlated single-
657 photon counting system in the range of 670 nm to 750 nm with wavelength intervals of 1 nm and time
658 intervals of 18.3 ps (Hamada *et al.*, 2017). A picosecond pulsed diode laser (PiL047X; Advanced Laser
659 Diode Systems) operating at 459 nm with a repetition frequency of 3 MHz was used as the excitation
660 light source to excite Chl *c* and fucoxanthin, which are the main light-harvesting pigments in diatoms.
661 Detailed conditions for TRF measurements were described in Nagao et al. (2020). Fluorescence decay-
662 associated spectrum (spectra) (FDAS) was constructed by global analysis (Ye *et al.*, 2025). The
663 analysis of fluorescence decay $I_F(\lambda, t)$ at wavelength λ with the instrument response function $I_{IRF}(t)$
664 followed the function:

665 $$\boldsymbol{I_F(\lambda, t) = C\sum_{i=1}^{6} A_i(\lambda)\int_{-\infty}^{t} I_{IRF}\left(t'\right) exp\left(-\frac{t-t'}{\tau_i}\right)dt'}$$

666 The time constants were set to τ_{1-6} = 40 ps, 140 ps, 640 ps, 1.6 ns, 3.8 ns, and 26 ns based on the
667 analysis results of WT cells cultured under white LED light. $A_i(\lambda)$ represents each FDAS. The rise and
668 decay components of fluorescence are indicated by negative and positive peaks, respectively. Positive
669 and negative amplitudes in the same FDAS reflect energy transfer from pigments with positive
670 amplitude to pigments with negative amplitude (Yokono *et al.*, 2008; Akimoto *et al.*, 2012). Positive
671 amplitude without accompanying negative amplitude signifies energy quenching. For Figure 5 and
672 Figure S8, FDAS of each sample were normalized as,

673 $$\frac{A_i(\lambda)}{\int \sum_{i=1}^{6} A_i(\lambda)\, d\lambda} \times 100.$$

674 For easy viewing, a magnification factor 100 was used.

675

676 **Measurement of PSII Antenna Size**

677 Chl fluorescence induction curves of wild-type (WT) and three *lhcx1* mutants (*lhcx1-7*, *lhcx1-15*,
678 *lhcx1-234*) cultured under HL at 300 μmol photons m^{-2} s^{-1} at 25ºC for 3 days were measured using
679 a Fluorometer FL 3500 (PSI, Czech Republic) in the presence of 40 μM DCMU after 10 minutes of
680 dark adaptation (Nedbal et al., 1999). The cultures were grown in two independent batches. In the
681 measurement, Flash intensity was set to 50%, and Flash duration was set to 100 μs. The relative
682 antenna size was determined by measuring the initial slope of fluorescence rise during the first 1 μs
683 relative to the maximum fluorescence intensity. Four replicate measurements were performed for each
684 strain.

685 For data analysis, R statistical software (R version 4.4.0 (2024-04-24)) was used, utilizing the
686 tidyverse, readr, lme4, emmeans, and ggplot2 packages. To account for batch-specific variations,
687 antenna size values were normalized by the WT mean value within each batch. A linear mixed model
688 (LMM) was applied to the normalized data. Estimated marginal means were calculated, and multiple
689 comparisons were performed using the Tukey method. The threshold for statistical significance was
690 set at adjusted $p < 0.05$.

691

692 **RNA sequencing**

693 Cells were harvested during the exponential growth phase after three days of treatment exposure to
694 LL/HL and ambient/3%-CO_2-supplemented air aeration, frozen with liquid nitrogen, and stored at −
695 80ºC until use. The frozen cells were crushed by bead-beating with 3-mm diameter zirconia beads.
696 Total RNA was extracted using the RNeasy Plant Mini Kit (Qiagen). RNA-seq libraries were
697 constructed using the NEBNext Ultra II RNA Library Prep Kit for Illumina (Cat No. 7770) according

to the manufacturer's instructions. Paired-end sequencing (2 × 150 bp) was performed on an Illumina NovaSeq 6000 platform. Adapter sequences (Read 1: AGATCGGAAGAGCACACGTCTGAACTCCAGTCA, Read 2: AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT) were trimmed from raw sequencing reads. Prior to mapping, the gene structure annotation provided in ChaetoBase v1.1 (predicted_all.gff3) was modified using the following procedures: (1) replacement of geneID attributes with Parent attributes for transcripts lacking gene correspondence, (2) correction of transcripts incorrectly defined as genes, and (3) addition of missing gene definitions using AGAT v1.0.0.

Clean reads were aligned to the *C. gracilis* reference genome (ChaetoBase v1.1) using STAR v2.7.11b with the following parameters: --alignIntronMax 10000. Gene expression levels were quantified using RSEM v1.3.3 with --strandedness reverse to account for Illumina stranded RNA-seq protocol, and expected count matrices were generated for downstream analysis.

Differential Gene Expression Analysis

Statistical analysis of gene expression was performed using the edgeR package (v4.6.2) in R (v4.5.0). Genes with low expression levels were filtered out by retaining only those with CPM (counts per million) values exceeding the 20th percentile in at least one condition. A DGEList object was created, and normalization factors were calculated using the TMM (trimmed mean of M-values) method. Dispersion parameters were estimated using the estimateDisp function, and generalized linear models (GLM) were fitted using glmQLFit.

Statistical Contrasts and Hypothesis Testing

To address the research objectives, the following statistical contrasts were performed:

1) Primary Analysis: HL Response Comparison

WT HL response: Air.HL.WT vs Air.LL.WT

lhcx1 HL response: Air.HL.lhcx1 vs Air.LL.lhcx1

Enhanced HL response in lhcx1: (Air.HL.lhcx1 - Air.LL.lhcx1) - (Air.HL.WT - Air.LL.WT)

2) Secondary Analysis: CO_2 Effect Validation

WT HL response under CO_2: CO_2.HL.WT vs CO_2.LL.WT

lhcx1 HL response under CO_2: CO_2.HL.lhcx1 vs CO_2.LL.lhcx1

Enhanced HL response in *lhcx1* under CO2: (CO2.HL.lhcx$_1$ - CO_2.LL.lhcx1) - (CO_2.HL.WT - CO_2.LL.WT)

Differential expression was determined using quasi-likelihood F-tests implemented in glmQLFTest. P-values were adjusted for multiple testing using the Benjamini-Hochberg method, with an FDR threshold of 0.05.

Functional Enrichment Analysis

KEGG pathway enrichment analysis was conducted using the enrichKEGG function from the clusterProfiler package (v4.16.0). KEGG Orthology (KO) annotations for *C. gracilis* genes were obtained from BLASTP searches against the KEGG database using BLASTKOALA. Only pathways with q-value < 0.05 were considered significantly enriched.

Acknowledgment

This work was supported in part by JSPS KAKENHI grant nos. JP22KJ2017 (M.K.), JP24H02086 (S.A.), JP23KJ1361 (Ko.I.), JP23H02347 (Ke.I. and S.A.), JP24H02081 (Ke.I.), and a grant no. L-2023-3-007 from the Institute for Fermentation, Osaka, Japan (Ke.I.).

Competing Interests

Authors HH, AS, and SI are employed by the NTT, Inc.

Author Contributions

M.K. and Ke.I. conceived the project; M.K. performed phylogenetic analysis, 2D-CN/SDS-PAGE, growth experiments, PAM analyses, absorbance spectrum analyses, 77K emission fluorescence analysis, antenna size measurements, and RNA-seq *in silico* analysis, and prepared samples for ime-resolved fluorescence analyses; S.A. performed time-resolved fluorescence analyses; A.T. conceived the methodology of 2D-CN/SDS-PAGE; Ko.I. measured and analyzed oxygen-evolving activity; S.T. measured pigments; H.H., A.S., and S.I. performed RNA-seq and provided funding; N.I. conceived methodology of SDS-PAGE and immunoblotting; N.I.-K. and Y.K. provided anti-Lhcx1 antibody. M.K. drafted the original manuscript, and Ke.I. revised and edited the final manuscript. All authors contributed to the discussion of the results and improvement of the manuscript.

References

Akimoto S, Yokono M, Hamada F, Teshigahara A, Aikawa S, Kondo A. **2012**. Adaptation of light-harvesting systems of Arthrospira platensis to light conditions, probed by time-resolved fluorescence spectroscopy. In: *Biochimica et Biophysica Acta - Bioenergetics*. 1483–1489.

Allorent G, Lefebvre-Legendre L, Chappuis R, Kuntz M, Truong TB, Niyogi KK, Ulm R, Goldschmidt-Clermont M. **2016**. UV-B photoreceptor-mediated protection of the photosynthetic machinery in *Chlamydomonas reinhardtii. Proceedings of the National Academy of Sciences of the United States of America* **113**: 14864–14869.

Allorent G, Tokutsu R, Roach T, Peers G, Cardol P, Girard-Bascou J, Seigneurin-Berny D, Petroutsos D, Kuntz M, Breyton C, *et al.* 2013. A Dual Strategy to Cope with High Light in *Chlamydomonas reinhardtii*. *The Plant Cell* **25**: 545–557.

Andrizhiyevskaya EG, Chojnicka A, Bautista JA, Diner BA, Van Grondelle R, Dekker JP. **2005**. Origin of the F685 and F695 fluorescence in Photosystem II. In: *Photosynthesis Research*. 173–180.

Anisimova M, Gil M, Dufayard JF, Dessimoz C, Gascuel O. **2011**. Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. *Systematic Biology* **60**: 685–699.

Bailleul B, Berne N, Murik O, Petroutsos D, Prihoda J, Tanaka A, Villanova V, Bligny R, Flori S, Falconet D, *et al.* 2015. Energetic coupling between plastids and mitochondria drives CO_2 assimilation in diatoms. *Nature* **524**: 366–369.

Bailleul B, Rogato A, De Martino A, Coesel S, Cardol P, Bowler C, Falciatore A, Finazzi G. **2010**. An atypical member of the light-harvesting complex stress-related protein family modulates diatom responses to light. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 18214–18219.

Buck JM, Kroth PG, Lepetit B. **2021**. Identification of sequence motifs in Lhcx proteins that confer qE-based photoprotection in the diatom *Phaeodactylum tricornutum*. *Plant Journal* **108**: 1721–1734.

Buck JM, Sherman J, Bártulos CR, Serif M, Halder M, Henkel J, Falciatore A, Lavaud J, Gorbunov MY, Kroth PG, *et al.* 2019. Lhcx proteins provide photoprotection via thermal dissipation of absorbed light in the diatom *Phaeodactylum tricornutum*. *Nature Communications* **10**.

Chukhutsina VU, Büchel C, Van Amerongen H. **2014**. Disentangling two non-photochemical quenching processes in *Cyclotella meneghiniana* by spectrally-resolved picosecond fluorescence at 77 K. *Biochimica et Biophysica Acta - Bioenergetics* **1837**: 899–907.

Chukhutsina VU, Fristedt R, Morosinotto T, Croce R. **2017**. Photoprotection strategies of the alga *Nannochloropsis gaditana*. *Biochimica et Biophysica Acta - Bioenergetics* **1858**: 544–552.

Croteau D, Jaubert M, Falciatore A, Bailleul B. **2025**. Pennate diatoms make non-photochemical quenching as simple as possible but not simpler. *Nature Communications* **16**.

Curtis BA, Tanifuji G, Maruyama S, Gile GH, Hopkins JF, Eveleigh RJM, Nakayama T, Malik SB, Onodera NT, Slamovits CH, *et al.* 2012. Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs. *Nature* **492**: 59–65.

Dittami SM, Michel G, Collén J, Boyen C, Tonon T. **2010**. Chlorophyll-binding proteins revisited - a multigenic family of light-harvesting and stress proteins from a brown algal perspective. *BMC Evolutionary Biology* **10**: 365.

Dorrell RG, Kuo A, Füssy Z, Richardson EH, Salamov A, Zarevski N, Freyria NJ, Ibarbalz FM, Jenkins J, Karlusich JJP, *et al.* 2023. Convergent evolution and horizontal gene transfer in Arctic Ocean microalgae. *Life Science Alliance* **6**.

Feng Y, Li Z, Li X, Shen L, Liu X, Zhou C, Zhang J, Sang M, Han G, Yang W, *et al.* 2023. Structure of a diatom photosystem II supercomplex containing a member of Lhcx family and dimeric

FCPII. *Science Advances* **9**.

Fu L, Niu B, Zhu Z, Wu S, Li W. **2012**. CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**: 3150–3152.

Gao Y, Zhao Y. **2014**. Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. *Journal of Integrative Plant Biology* **56**: 343–349.

Giossi CE, Wünsch MA, Dautermann O, Schober AF, Buck JM, Kroth PG, Lohr M, Lepetit B. **2025**. Both major xanthophyll cycles present in nature promote non-photochemical quenching in a model diatom. *Plant Physiology*: kiaf371.

Giovagnetti V, Jaubert M, Shukla MK, Ungerer P, Bouly JP, Falciatore A, Ruban AV. 2022. Biochemical and molecular properties of LHCX1, the essential regulator of dynamic photoprotection in diatoms. *Plant Physiology* **188**: 509–525.

Goss R, Lepetit B. **2015**. Biodiversity of NPQ. *Journal of Plant Physiology* **172**: 13–32.

Grouneva I, Rokka A, Aro EM. **2011**. The thylakoid membrane proteome of two marine diatoms outlines both diatom-specific and species-specific features of the photosynthetic machinery. *Journal of Proteome Research* **10**: 5338–5353.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. **2010**. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Systematic Biology* **59**: 307–321.

Hamada F, Murakami A, Akimoto S. **2017**. Adaptation of Divinyl Chlorophyll *a*/*b*-Containing Cyanobacterium to Different Light Conditions: Three Strains of Prochlorococcus marinus. *Journal of Physical Chemistry B* **121**: 9081–9090.

Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. **2018**. UFBoot2: Improving the Ultrafast Bootstrap Approximation. Molecular biology and evolution. *Molecular Biology and Evolution* **35**: 518–522.

Holzwarth AR, Miloslavina Y, Nilkens M, Jahns P. **2009**. Identification of two quenching sites active in the regulation of photosynthetic light-harvesting studied by time-resolved fluorescence. *Chemical Physics Letters* **483**: 262–267.

Hori K, Maruyama F, Fujisawa T, Togashi T, Yamamoto N, Seo M, Sato S, Yamada T, Mori H, Tajima N, *et al.* 2014. Klebsormidium flaccidum genome reveals primary factors for plant terrestrial adaptation. *Nature Communications* **5**.

Huang W, Daboussi F. **2017**. Genetic and metabolic engineering in diatoms. *Philosophical Transactions of the Royal Society B: Biological Sciences* **372**.

Ifuku K, Yan D. **2020**. *Efficient Transformation of the Diatoms Phaeodactylum tricornutum by Multipulse Electroporation*. Humana, New York, NY.

Ifuku K, Yan D, Miyahara M, Inoue-Kashino N, Yamamoto YY, Kashino Y. **2015**. A stable and efficient nuclear transformation system for the diatom *Chaetoceros gracilis*. *Photosynthesis Research* **123**: 203–211.

Im SH, Lepetit B, Mosesso N, Shrestha S, Weiss L, Nymark M, Roellig R, Wilhelm C, Isono E, Kroth PG. **2024**. Identification of promoter targets by Aureochrome 1a in the diatom *Phaeodactylum tricornutum*. *Journal of Experimental Botany* **75**: 1834–1851.

Ishihara T, Ifuku K, Yamashita E, Fukunaga Y, Nishino Y, Miyazawa A, Kashino Y, Inoue-Kashino N. **2015**. Utilization of light by fucoxanthin–chlorophyll-binding protein in a marine centric diatom, *Chaetoceros gracilis*. *Photosynthesis Research* **126**: 437–447.

Järvi S, Suorsa M, Paakkarinen V, Aro EM. **2011**. Optimized native gel systems for separation of thylakoid protein complexes: Novel super- and mega-complexes. *Biochemical Journal* **439**: 207–214.

Jeffrey SW, Humphrey GF. **1975**. New spectrophotometric equations for determining chlorophylls *a*, *b*, c_1 and c_2 in higher plants, algae and natural phytoplankton. *Biochemie und*

Physiologie der Pflanzen **167**: 191–194.

Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. **2017**. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods* **14**: 587–589.

Kameo S, Aso M, Furukawa R, Matsumae R, Yokono M, Fujita T, Tanaka A, Tanaka R, Takabayashi A. **2021**. Substitution of Deoxycholate with the Amphiphilic Polymer Amphipol A8-35 Improves the Stability of Large Protein Complexes during Native Electrophoresis. *Plant and Cell Physiology* **62**: 348–355.

Katoh K, Standley DM. **2013**. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* **30**: 772–780.

Koziol AG, Borza T, Ishida KI, Keeling P, Lee RW, Durnford DG. **2007**. Tracing the evolution of the light-harvesting antennae in chlorophyll *a*/*b*-containing organisms. *Plant Physiology* **143**: 1802–1816.

Kumazawa M, Ifuku K. **2024**. Unraveling the evolutionary trajectory of LHCI in red-lineage algae: Conservation, diversification, and neolocalization. *iScience* **27**: 110897.

Kumazawa M, Nishide H, Nagao R, Inoue-Kashino N, Shen J, Nakano T, Uchiyama I, Kashino Y, Ifuku K. **2022**. Molecular phylogeny of fucoxanthin-chlorophyll *a* / *c* proteins from *Chaetoceros gracilis* and Lhcq/Lhcf diversity. *Physiologia Plantarum* **174**: e13598.

Lavaud J, Kroth PG. **2006**. In diatoms, the transthylakoid proton gradient regulates the photoprotective non-photochemical fluorescence quenching beyond its control on the xanthophyll cycle. *Plant and Cell Physiology* **47**: 1010–1016.

Lavaud J, Lepetit B. **2013**. An explanation for the inter-species variability of the photoprotective non-photochemical chlorophyll fluorescence quenching in diatoms. *Biochimica et Biophysica Acta - Bioenergetics* **1827**: 294–302.

Leliaert F, Smith DR, Moreau H, Herron MD, Verbruggen H, Delwiche CF, De Clerck O. **2012**.

Phylogeny and Molecular Evolution of the Green Algae. *Critical Reviews in Plant Sciences* **31**: 1–46.

Lepetit B, Volke D, Gilbert M, Wilhelm C, Goss R. **2010**. Evidence for the existence of one antenna-associated, lipid-dissolved and two protein-bound pools of diadinoxanthin cycle pigments in diatoms. *Plant Physiology* **154**: 1905–1920.

Letunic I, Bork P. **2024**. Interactive Tree of Life (iTOL) v6: Recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Research* **52**: W78–W82.

Lohr M, Wilhelm C. **1999**. *Algae displaying the diadinoxanthin cycle also possess the violaxanthin cycle*.

Malnoë A. **2018**. Photoinhibition or photoprotection of photosynthesis? Update on the (newly termed) sustained quenching component qH. *Environmental and Experimental Botany* **154**: 123–133.

Manuel Zapata, Francisco Rodríguez, José L. Garrido. **2000**. Separation of chlorophylls and carotenoids from marine phytoplankton: a new HPLC method using a reversed phase C8 column and pyridine-containing mobile phases. *Marine Ecology Progress Series* **195**: 29–45.

Miloslavina Y, Grouneva I, Lambrev PH, Lepetit B, Goss R, Wilhelm C, Holzwarth AR. **2009**. Ultrafast fluorescence study on the location and mechanism of non-photochemical quenching in diatoms. *Biochimica et Biophysica Acta - Bioenergetics* **1787**: 1189–1197.

Minagawa J, Tokutsu R. **2015**. Dynamic regulation of photosynthesis in *Chlamydomonas reinhardtii*. *The Plant Journal* **82**: 413–428.

Minh BQ, Dang CC, Vinh LS, Lanfear R. **2021**. QMaker: Fast and Accurate Method to Estimate Empirical Models of Protein Evolution. *Systematic Biology* **70**: 1046–1060.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. **2020**. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era (E Teeling, Ed.). *Molecular Biology and Evolution* **37**: 1530–1534.

Miyahara M, Aoi M, Inoue-Kashino N, Kashino Y, Ifuku K. **2013**. Highly efficient

901 transformation of the diatom *Phaeodactylum tricornutum* by multi-pulse electroporation. *Bioscience,*
902 *Biotechnology and Biochemistry* **77**: 874–876.

903 **Nagao R, Ishii A, Tada O, Suzuki T, Dohmae N, Okumura A, Iwai M, Takahashi T, Kashino**
904 **Y, Enami I**. **2007**. Isolation and characterization of oxygen-evolving thylakoid membranes and
905 Photosystem II particles from a marine diatom *Chaetoceros gracilis*. *Biochimica et Biophysica Acta -*
906 *Bioenergetics* **1767**: 1353–1362.

907 **Nagao R, Kato K, Ifuku K, Suzuki T, Kumazawa M, Uchiyama I, Kashino Y, Dohmae N,**
908 **Akimoto S, Shen J-R, *et al.* 2020a**. Structural basis for assembly and function of a diatom
909 photosystem I-light-harvesting supercomplex. *Nature Communications* **11**: 2481.

910 **Nagao R, Kato K, Kumazawa M, Ifuku K, Yokono M, Suzuki T, Dohmae N, Akita F, Akimoto**
911 **S, Miyazaki N, *et al.* 2022**. Structural basis for different types of hetero-tetrameric light-harvesting
912 complexes in a diatom PSII-FCPII supercomplex. *Nature Communications* **13**: 1764.

913 **Nagao R, Kato K, Suzuki T, Ifuku K, Uchiyama I, Kashino Y, Dohmae N, Akimoto S, Shen J-**
914 **R, Miyazaki N, *et al.* 2019**. Structural basis for energy harvesting and dissipation in a diatom PSII–
915 FCPII supercomplex. *Nature Plants* **5**: 890–901.

916 **Nagao R, Tomo T, Noguchi E, Nakajima S, Suzuki T, Okumura A, Kashino Y, Mimuro M,**
917 **Ikeuchi M, Enami I**. **2010**. Purification and characterization of a stable oxygen-evolving Photosystem
918 II complex from a marine centric diatom, *Chaetoceros gracilis*. *Biochimica et Biophysica Acta -*
919 *Bioenergetics* **1797**: 160–166.

920 **Nagao R, Tomo T, Noguchi E, Suzuki T, Okumura A, Narikawa R, Enami I, Ikeuchi M**. **2012**.
921 Proteases are associated with a minor fucoxanthin chlorophyll *a*/*c*-binding protein from the diatom,
922 *Chaetoceros gracilis*. *Biochimica et Biophysica Acta - Bioenergetics* **1817**: 2110–2117.

923 **Nagao R, Ueno Y, Akimoto S, Shen JR**. **2020b**. Effects of CO2 and temperature on photosynthetic
924 performance in the diatom *Chaetoceros gracilis*. *Photosynthesis Research* **146**: 189–195.

925 **Nagao R, Yokono M, Akimoto S, Tomo T**. **2013**. High excitation energy quenching in fucoxanthin
926 chlorophyll *a*/*c*-binding protein complexes from the diatom *Chaetoceros gracilis*. *Journal of Physical*
927 *Chemistry B* **117**: 6888–6895.
928 **Nagao R, Yokono M, Ueno Y, Shen JR, Akimoto S**. **2020c**. Excitation-Energy Transfer and
929 Quenching in Diatom PSI-FCPI upon P700 Cation Formation. *Journal of Physical Chemistry B* **124**:
930 1481–1486.
931 **Nakayasu M, Akimoto S, Yoneda K, Ikuta S, Shimakawa G, Matsuda Y**. **2024**. Different
932 functions of Lhcx isoforms in photoprotective mechanism in the marine diatom *Thalassiosira*
933 *pseudonana*. *BioRxiv*.
934 **Nedbal L, Trtflek M, Kaftan D**. **1999**. *Flash fluorescence induction: a novel method to study*
935 *regulation of Photosystem II*.
936 **Niyogi KK**. **2000**. Safety valves for photosynthesis. *Current Opinion in Plant Biology* **3**: 455–460.
937 **Park S, Steen CJ, Lyska D, Fischer AL, Endelman B, Iwai M, Niyogi KK, Fleming GR**. **2019**.
938 Chlorophyll–carotenoid excitation energy transfer and charge transfer in *Nannochloropsis oceanica*
939 for the regulation of photosynthesis. *Proceedings of the National Academy of Sciences of the United*
940 *States of America* **116**: 3385–3390.
941 **Peers G, Truong TB, Ostendorf E, Busch A, Elrad D, Grossman AR, Hippler M, Niyogi KK**.
942 **2009**. An ancient light-harvesting protein is critical for the regulation of algal photosynthesis. *Nature*
943 **462**: 518–521.
944 **Petroutsos D, Tokutsu R, Maruyama S, Flori S, Greiner A, Magneschi L, Cusant L, Kottke T,**
945 **Mittag M, Hegemann P, *et al.* 2016**. A blue-light photoreceptor mediates the feedback regulation of
946 photosynthesis. *Nature* **537**: 563–566.
947 **Pierella Karlusich JJ, Cosnier K, Zinger L, Henry N, Nef C, Bernard G, Scalco E, Dvorak E,**
948 **Wincker P, Sunagawa S, *et al.* 2025**. Patterns and drivers of diatom diversity and abundance in the

global ocean. *Nature Communications* **16**.

Pierella Karlusich JJ, Ibarbalz FM, Bowler C. **2020**. Phytoplankton in the *Tara* Ocean. *Annual Review of Marine Science* **12**: 233–265.

Steenwyk JL, Buida TJ, Li Y, Shen X-X, Rokas A. **2020**. ClipKIT: A multiple sequence alignment trimming software for accurate phylogenomic inference (A Hejnol, Ed.). *PLOS Biology* **18**: e3001007.

Strassert JFH, Irisarri I, Williams TA, Burki F. **2021**. A molecular timescale for eukaryote evolution with implications for the origin of red algal-derived plastids. *Nature Communications* **12**: 1879.

Taddei L, Chukhutsina VU, Lepetit B, Stella GR, Bassi R, van Amerongen H, Bouly JP, Jaubert M, Finazzi G, Falciatore A. **2018**. Dynamic changes between two LHCX-related energy quenching sites control diatom photoacclimation. *Plant Physiology* **177**: 953–965.

Tian L, Nawrocki WJ, Liu X, Polukhina I, Van Stokkum IHM, Croce R. **2019**. PH dependence, kinetics and light-harvesting regulation of nonphotochemical quenching in Chlamydomonas. *Proceedings of the National Academy of Sciences of the United States of America* **116**: 8320–8325.

Tokutsu R, Fujimura-Kamada K, Matsuo T, Yamasaki T, Minagawa J. **2019a**. The CONSTANS flowering complex controls the protective response of photosynthesis in the green alga Chlamydomonas. *Nature Communications* **10**.

Tokutsu R, Fujimura-Kamada K, Yamasaki T, Matsuo T, Minagawa J. **2019b**. Isolation of photoprotective signal transduction mutants by systematic bioluminescence screening in Chlamydomonas reinhardtii. *Scientific Reports* **9**.

Wang W, Yu L-J, Xu C, Tomizaki T, Zhao S, Umena Y, Chen X, Qin X, Xin Y, Suga M, *et al.* 2019. Structural basis for blue-green light harvesting and energy dissipation in diatoms. *Science* **363**.

Xu C, Pi X, Huang Y, Han G, Chen X, Qin X, Huang G, Zhao S, Yang Y, Kuang T, *et al.* 2020. Structural basis for energy transfer in a huge diatom PSI-FCPI supercomplex. *Nature Communications*

11: 5081.

Yamamoto Y, Leng J, Shen J-R. **2011**. Isolation of Photosystem II-Enriched Membranes and the Oxygen-Evolving Complex Subunit Proteins from Higher Plants. In: Carpentier R, ed. Photosynthesis Research Protocols. Totowa, NJ: Humana Press, 1–10.

Yang YN, Le TTL, Hwang JH, Zulfugarov IS, Kim EH, Kim HU, Jeon JS, Lee DH, Lee CH. **2022**. High Light Acclimation Mechanisms Deficient in a PsbS-Knockout Arabidopsis Mutant. *International Journal of Molecular Sciences* **23**.

Ye Z, Sawada M, Iwasa M, Moriyama R, Dey D, Furutani M, Kitao M, Hara T, Tanaka A, Kishimoto J, *et al.* 2025. Revisiting the early light-induced protein hypothesis in the sustained thermal dissipation mechanism in yew leaves. *Journal of Experimental Botany* **76**: 513–531.

Yokono M, Akimoto S, Tanaka A. **2008**. Seasonal changes of excitation energy transfer and thylakoid stacking in the evergreen tree *Taxus cuspidata*: How does it divert excess energy from photosynthetic reaction center? *Biochimica et Biophysica Acta - Bioenergetics* **1777**: 379–387.

Zhang T, Gao Y, Wang R, Zhao Y. **2017**. Production of Guide RNAs in vitro and in vivo for CRISPR Using Ribozymes and RNA Polymerase II Promoters. *Bio Protoc.* **7**.

Zhang H, Xiong X, Guo K, Zheng M, Cao T, Yang Y, Song J, Cen J, Zhang J, Jiang Y, *et al.* 2024. A rapid aureochrome opto-switch enables diatom acclimation to dynamic light. *Nature Communications* **15**: 5578.

Zhao S, Shen L, Li X, Tao Q, Li Z, Xu C, Zhou C, Yang Y, Sang M, Han G, *et al.* 2023. Structural insights into photosystem II supercomplex and trimeric FCP antennae of a centric diatom *Cyclotella meneghiniana*. *Nature Communications* **14**: 8164.

Zheng M, Pang X, Chen M, Tian L. **2024**. Ultrafast energy quenching mechanism of LHCSR3-dependent photoprotection in Chlamydomonas. *Nature communications* **15**: 4437.

Zhou C, Feng Y, Li Z, Shen L, Li X, Wang Y, Han G, Kuang T, Liu C, Shen J-R, *et al.* 2024.

997 Structural and spectroscopic insights into fucoxanthin chlorophyll a/c-binding proteins of diatoms in
998 diverse oligomeric states. *Plant Communications* **5**: 101041.
999
1000

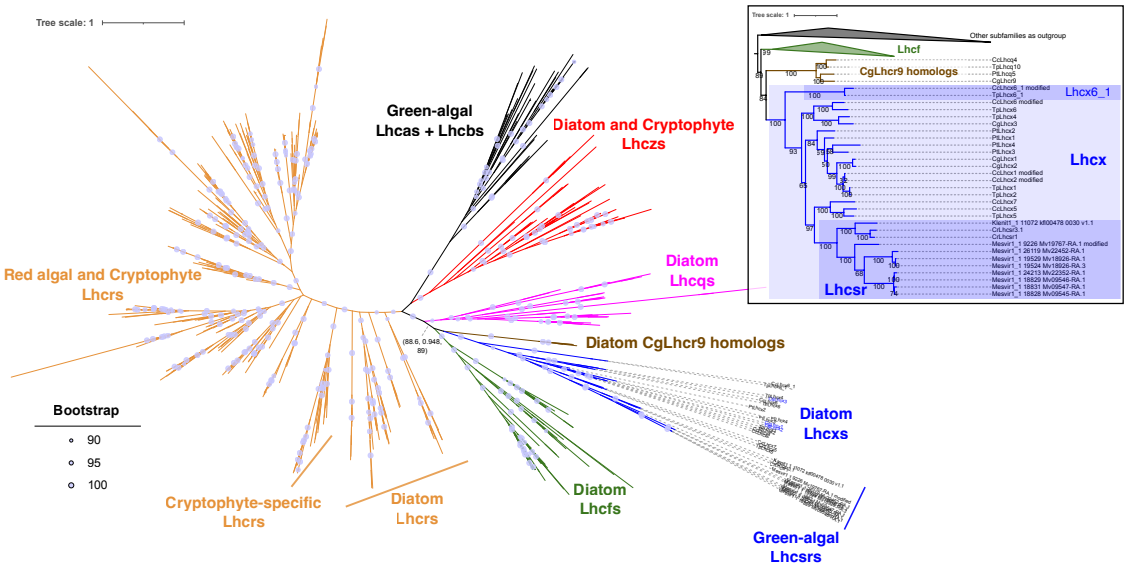
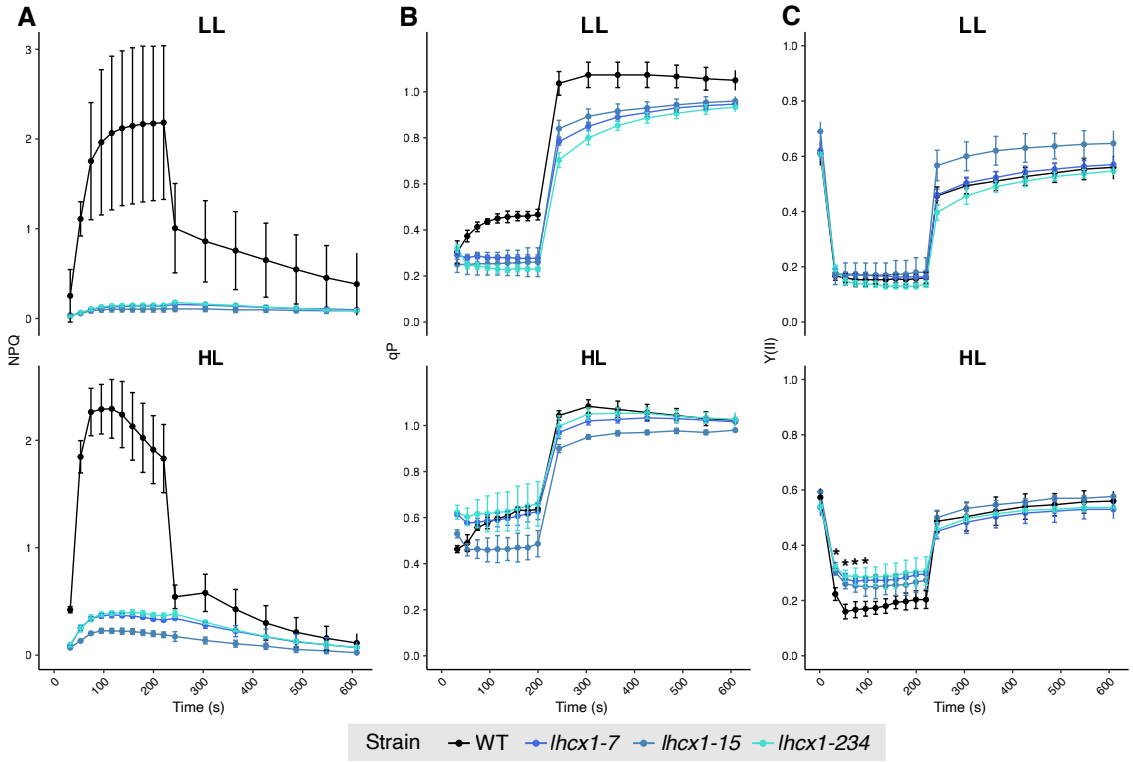


Figure 1. Molecular Phylogenetic Tree of Light-Harvesting Complexes (LHCs) from Red and Green Evolutionary Lineages

LHC sequences were obtained from red algae, green algae *Mesostigma viride* NIES-296 (Mesvir1_1) and *Klebsormidium nitens* NIES-2285 (Klepit1_1), diatoms: *Chaetoceros gracilis* (Cg), *Thalassiosira pseudonana* (Tp), *Cyclotella cryptica* (Cc), *Phaeodactylum tricornutum* (Pt), and cryptophytes: *Guillardia theta* CCMP2712 and Cryptophyceae sp. (*Baffinella* sp.) CCMP2293. These were combined with reference sequences for Lhcsr from *Chlamydomonas reinhardtii*: Lhcsr1 (UniProt: P93664) and Lhcsr3.1 (UniProt: P0DO19), which is identical to Lhcsr3.2 (UniProt: P0DO18), and a molecular phylogenetic tree was constructed using IQ-TREE2. The phylogenetic tree is displayed as an unrooted tree, and the inset tree is rerooted with the other LHC family members except for Lhcx, Lhcf, and CgLhcr9 homologs. Detailed annotations are shown only for the Lhcx/Lhcsr subfamily. Circles on nodes indicate ultra-fast bootstrap support values of 90% or higher. Numbers in parentheses indicate SH-aLRT support values (%) / aBayes support values / ultra-fast bootstrap support values (%). Numbers in the inset tree are ultra-fast bootstrap support values (%).



1017

1018 **Figure 2. PAM Fluorescence Measurements of *C. gracilis* Wild-Type and *lhcx1* Strains Cultured**

1019 **with Aeration for 7–8 Days Under Low Light (LL) and High Light (HL)**

1020 A) NPQ of LL and HL cultures. B) qP of LL and HL cultures. C) PSII effective quantum yield Y(II)

1021 of LL and HL cultures. * indicates that adjusted $p < 0.05$ by Tukey's HSD test was satisfied between

1022 the wild-type and all *lhcx1* strains. The points in each figure represent mean ± standard deviation with

1023 biological replicates of $n = 3$.

1024

1025

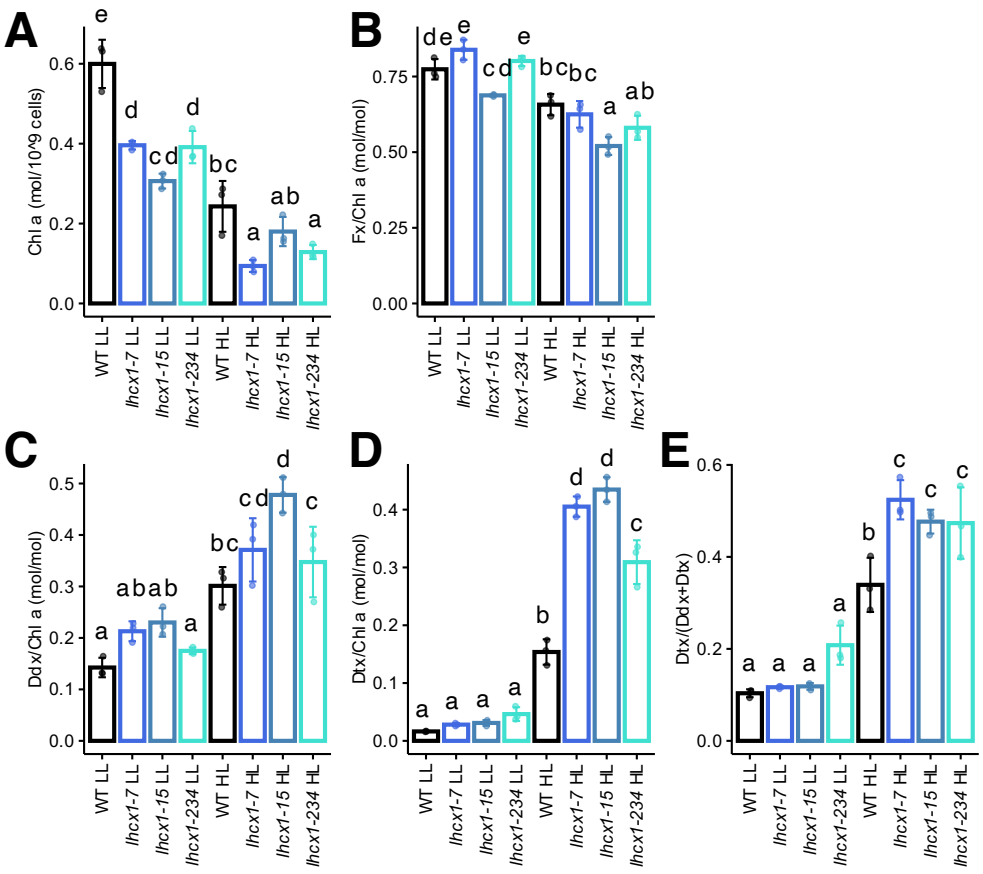


Figure 3. Pigment Analysis of Low Light (LL) and High Light (HL) Cultures

A) Chlorophyll (Chl) *a* content per 10^9 cells. B) Fucoxanthin (Fx) content per Chl *a*. C) Diatoxanthin (Dtx) content per Chl *a*. D) Diadinoxanthin (Ddx) content per Chl *a*. E) De-epoxidation state (DEPS = [Dtx]/([Dtx]+[Ddx])). Data are presented as mean ± standard deviation (n = 3). Different alphabets indicate adjusted $p < 0.05$ by Tukey's HSD test.

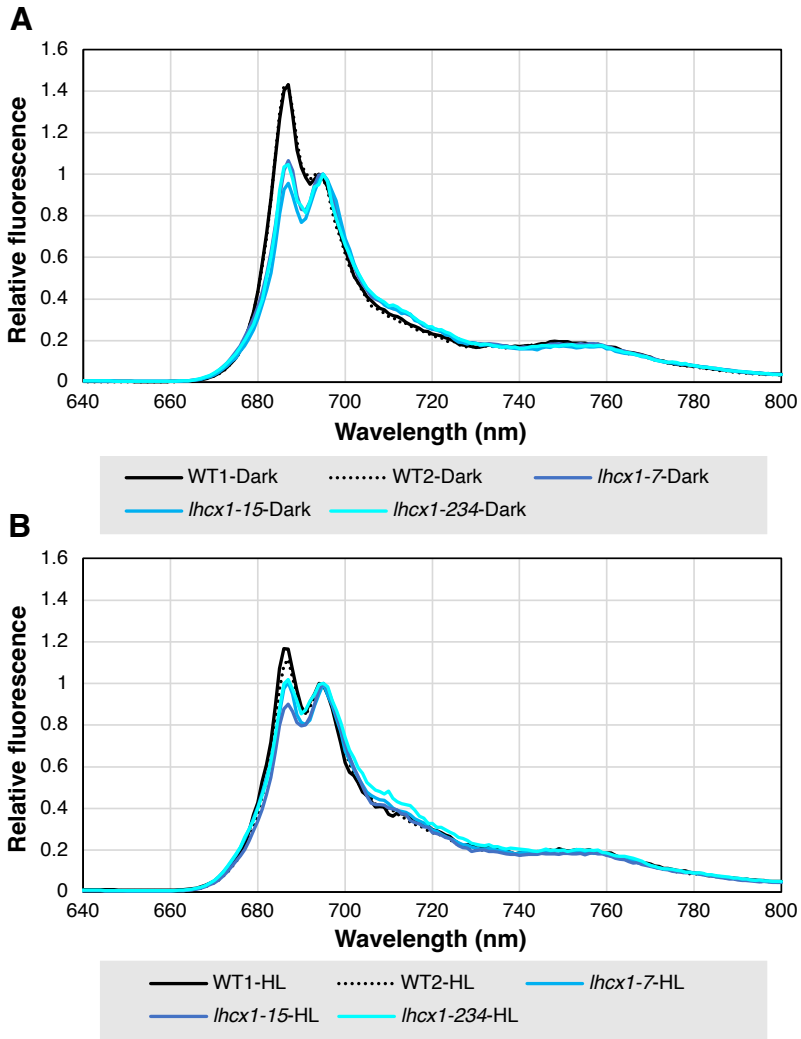
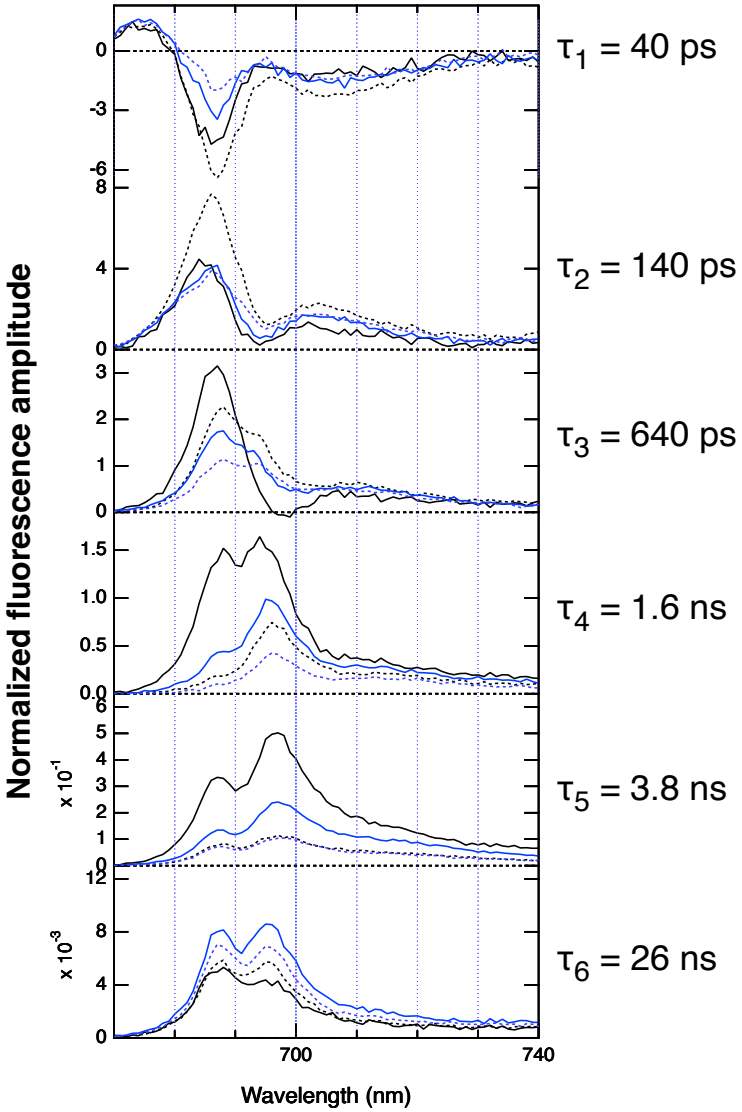


Figure 4. Low-Temperature Fluorescence Emission Spectra of Wild-Type (WT) and *lhcx1* Mutants Cultured Under High Light

A) Low-temperature fluorescence emission spectra measured after 30 minutes of dark treatment followed by rapid freezing in liquid nitrogen. B) Low-temperature fluorescence emission spectra measured after 30 minutes of HL treatment followed by rapid freezing in liquid nitrogen. Each spectrum was normalized at the peak corresponding to CP47 (693–695 nm). The wild type was measured in two separately cultured samples (black solid line and black dashed line).



1039 **Figure 5. Fluorescence Decay-Associated Spectra of Time-Resolved Fluorescence**

1040 Black lines indicate WT, blue lines indicate *lhcx1-7*, solid lines represent measurements after 30

1041 minutes of dark treatment, and dashed lines represent measurements after 30 minutes of treatment at

1042 300 µmol photons m^{-2} s^{-1}.

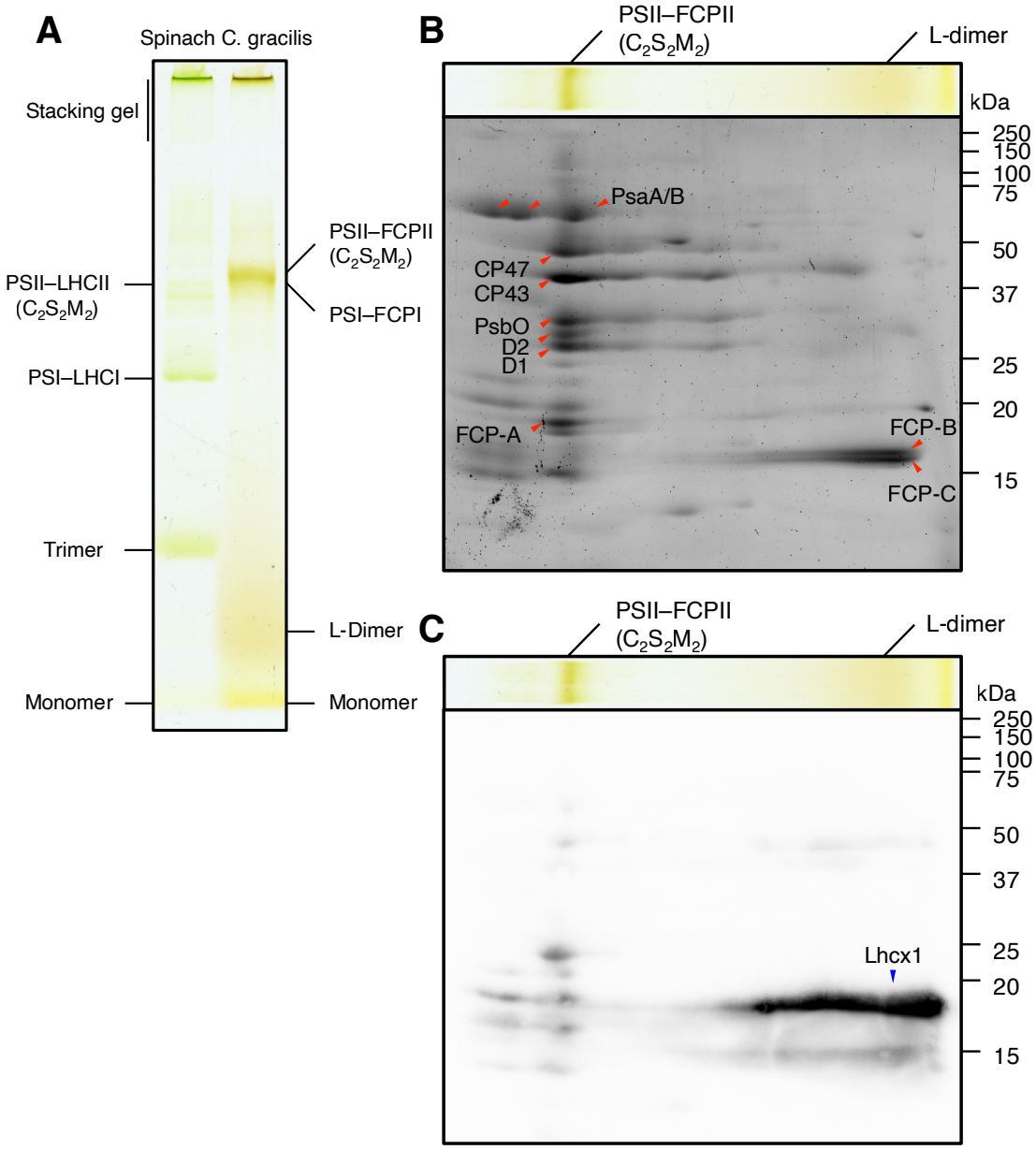


Figure 6. 2D-Amphipol-based-CN/SDS-PAGE of *C. gracilis* Thylakoid Membranes

A) Amphipol-based CN-PAGE of *C. gracilis* and spinach thylakoid membranes. Annotation of each band for spinach follows previous reports (Järvi *et al.*, 2011; Kameo *et al.*, 2021). B) Protein fluorescence staining image of two-dimensional electrophoresis of Amphipol-based CN-PAGE. Annotation of each spot indicated by red arrowheads is based on previous reports (Nagao *et al.*, 2007).

1049 FCP-A, FCP-B, and FCP-C correspond to CgLhcf1, CgLhcf12, and CgLhcf2, respectively (Nagao *et*

1050 *al.*, 2019, 2022; Zhou *et al.*, 2024). C) Identification of the Lhcx1 spot by immunoblotting using anti-

1051 Lhcx1 antibody. The blue arrowhead indicates the detected spot.

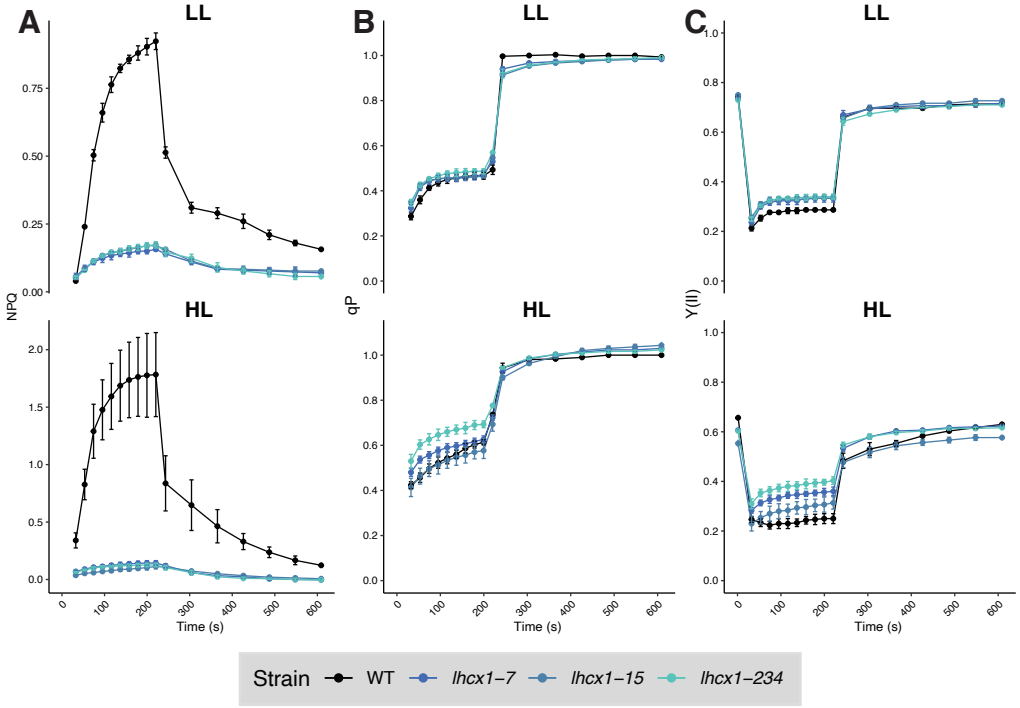


Figure 7. PAM Fluorescence Measurements of *C. gracilis* Wild-Type and *lhcx1* Strains Cultured with 3% CO_2 Aeration for three Days Under Low Light (LL) and High Light (HL)

A) NPQ of LL and HL cultures. B) qP of LL and HL cultures. C) PSII effective quantum yield Y(II) of LL and HL cultures. The points in each figure represent mean ± standard deviation with biological replicates of $n = 3$.

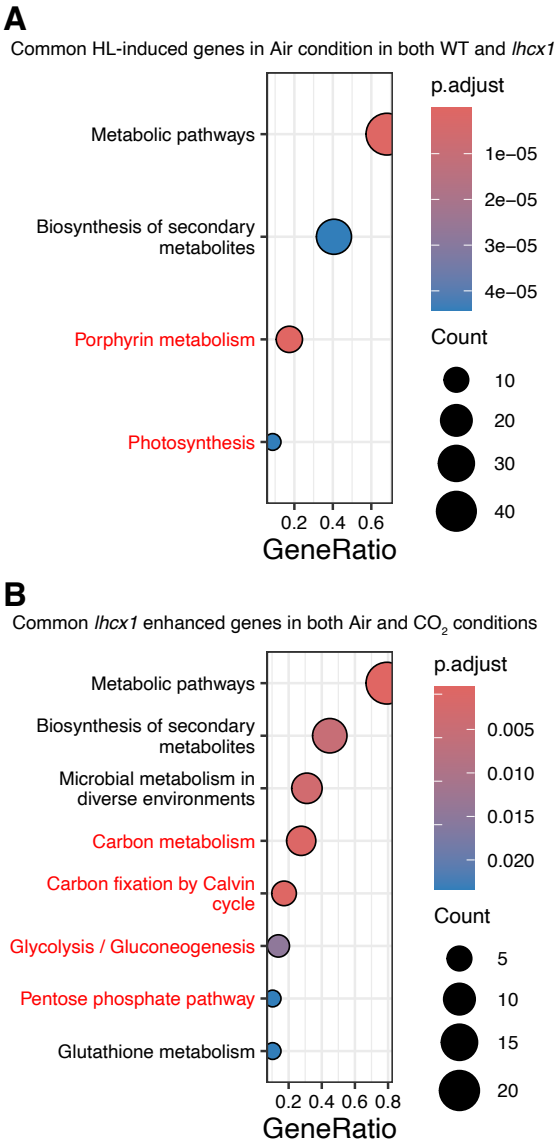


Figure 8. Gene Expression Analysis under CO_2 Enrichment

KEGG pathway enrichment analysis of differentially expressed genes in the *lhcx1* mutant under LL and HL conditions. A**)** Pathways enriched in genes commonly induced by HL in both WT and the *lhcx1* mutant under ambient air conditions. B**)** Pathways enriched in genes showing enhanced HL response in *lhcx1* mutant under both ambient air and CO_2-enriched conditions. Dot size represents gene count, and color intensity indicates statistical significance (q-value shown as p.adjust). Pathways related to typical HL response and photosynthesis are indicated with red letters.