1   **SHORT TITLE:** Effects of loss of ALB3b insertase in diatoms

2

3   **CORRESPONDING AUTHOR:** Marianne Nymark (marianne.nymark@ntnu.no)

4

5   **Loss of ALBINO3b insertase results in a truncated light-harvesting antenna in diatoms**

6   Marianne Nymark*#[1], Charlotte Volpe*[2], Marthe Caroline Grønbech Hafskjold[1], Henning
7   Kirst[3], Manuel Serif[1], Olav Vadstein[2], Atle Magnar Bones[1], Anastasios Melis[3], Per Winge[1].

8   *Joint first author

9   [1]Department of Biology, Norwegian University of Science and Technology, N-7491, Trondheim,
10   Norway

11   [2]Department of Biotechnology and Food Science, Norwegian University of Science and
12   Technology, N-7491, Trondheim, Norway

13   [3]Department of Plant and Microbial Biology, UC Berkeley, Berkeley, California 94720, USA

14

15   **ONE SENTENCE SUMMARY:**

16   Diatom ALB3b is required for insertion of Fx-Chl binding proteins in thylakoid membranes, and

17   has a novel conserved domain implying different interaction partners from those in plants/green

18   algae.

19

20   **AUTHOR CONTRIBUTIONS:**

21   M.N., A.M.B., O.V., A.M. and P.W. conceived the research plans. M.N., A.M.B., A.M. and

22   P.W. supervised and designed the experiments. M.N., C.V., M.C.G.H., H.K and M.S. performed

23   the experiments. M.N., C.V., M.C.G.H., H.K, A.M. and P.W. analyzed the data. M.N. and C.V.

24   wrote the article with contributions of all the authors. M.N. agrees to serve as the author

25   responsible for contact and ensures communication.

26

27

36

37 **AUTHOR FOR CONTACT:** marianne.nymark@ntnu.no (Marianne Nymark)

38

39 **ABSTRACT**

40 The family of chloroplast ALBINO3 proteins function in the insertion and assembly of thylakoid

41 membrane protein complexes. Loss of ALB3b in the marine diatom *Phaeodactylum tricornutum*

42 leads to a striking change of cell color from the normal brown to green. A 75% decrease of the

43 main fucoxanthin-chlorophyll *a/c*-binding proteins was identified in the *alb3b* strains as the

44 cause of changes in the spectral properties of the mutant cells. The *alb3b* lines exhibit a truncated

45 light-harvesting antenna phenotype with reduced amounts of light harvesting pigments and

46 require a higher light intensity for saturation of photosynthesis. Accumulation of photoprotective

47 pigments and LHCX proteins were not negatively affected in the mutant strains, but still the

48 capacity for non-photochemical quenching was lower compared to wild type. In plants and green

49 algae, ALB3 proteins interact with members of the chloroplast signal recognition particle

50 pathway through a lysine-rich C-terminal domain. A novel conserved C-terminal domain was

51 identified in diatoms and other stramenopiles, questioning if ALB3b proteins have the same

52 interaction partners as its plant/green algae homologs.

53

54 **INTRODUCTION**

55 Diatoms (*Bacillariophyceae*) are a major group of eukaryotic phytoplankton belonging to the

56 phylum Heterokont that evolved through a secondary endosymbiotic event around 200 to 180
57 million years ago (Brown and Sorhannus, 2010). Diatoms are key primary producers in the
58 marine food chain. They account for 40% of the total carbon fixation in oceans and 25% of the
59 total global oxygen production (Falkowski et al., 1998). Diatom plastids differ significantly from
60 the ones in green algae and higher plants due to their peculiar inheritance and evolution (Oudot-
61 Le Secq et al., 2007). Because of secondary endosymbiotic events, four membranes surround the
62 diatom chloroplast. The outer envelope, known as chloroplast endoplasmic reticulum, is a
63 continuum with the nuclear envelope. The diatom thylakoids are organized in stacked bands of
64 three membranes, also known as girdle lamellae, spanning along the entire length of the plastid.
65 This configuration differs significantly from the classic grana stacks and interconnecting stroma-
66 exposed thylakoids organization found in higher plant chloroplasts (Austin and Staehelin, 2011).
67 Light harvesting complexes (LHCs) are embedded in the thylakoid membrane of the chloroplast
68 and surround the photosynthetic reaction centers of the photosystems.

69 In contrast to higher plants, where specific LHCs serve either PSI or PSII, diatoms are
70 characterized by a peripheral fucoxanthin (Fx)-chlorophyll (Chl) *a/c* antenna complex believed
71 to deliver excitation energy to both photosystems, in addition to having a PSI associated antenna
72 (Lepetit et al., 2010; Büchel, 2015). Proteins of the peripheral Fx-Chl *a/c* antenna complex in
73 diatoms belong to the LHC superfamily (Durnford et al., 1996), but are often referred to as Fx-
74 Chl *a/c* binding proteins (FCPs) in order to distinguish them from the LHCs of the green lineages
75 (Falkowski and Raven, 2007). In addition to the light harvesting pigments, FCPs also bind
76 diadinoxanthin (Ddx) and diatoxanthin (Dtx), photoprotective pigments essential during light
77 stress conditions (Wang et al., 2019). The FCPs belong to three major LHC classes: the LHCF,
78 including the main Fx-Chl *a/c* binding proteins, the red algal-like LHCRs, and the LHCXs,
79 related to the LhcSRs in *Chlamydomonas reinhardtii* (Büchel, 2015). The latter has been shown
80 to play a central role in dissipating excessively absorbed energy through non-photochemical
81 quenching (NPQ) in cooperation with photoprotective pigments (Bailleul et al., 2010; Taddei et
82 al., 2016; Lepetit et al., 2017; Taddei et al., 2018).

83

84 LHC proteins and certain photosystem core proteins are known to be integrated into the
85 thylakoid membrane of higher plants and green microalgae through the post-translational or co-

86　translational part of the chloroplast signal recognition particle (CpSRP) assembly pathway
87　(Sundberg et al., 1997; Schuenemann et al., 1998; Bellafiore et al., 2002; Gerdes et al., 2006;
88　Kirst et al., 2012; Kirst et al., 2012; Kirst and Melis, 2014). The plant/green algae CpSRP
89　pathway includes the LHC specific chaperon CpSRP43, the GTPase CpSRP54, the signal
90　recognition receptor CpFTSY and the ALBINO3 insertase (ALB3) (Bellafiore et al., 2002; Kirst
91　and Melis, 2014). Homologues of CpSRP54, CpFTSY and ALB3 can be identified in diatom
92　genomes (Armbrust et al., 2004; Bowler et al., 2008; Mock et al., 2017), whereas no homologue
93　for the molecular chaperon CpSRP43 have been identified (Träger et al., 2012). CpSRP43
94　orthologues appear to be restricted to plants and green algae, however distantly related ankyrin
95　repeat proteins can be found in Haptophyceae. Diatom CpSRP54 knockout mutants have been
96　shown to be light sensitive (Nymark et al., 2016), but no further information exists about
97　CpSRP54's role, or the role of any other members of the CpSRP pathway, in integration and
98　assembly of thylakoid membrane proteins in diatoms. It has been shown, however, that efficient
99　integration of FCPs depend on stromal factors and on the presence of GTP (Lang and Kroth,
100　2001).

101

102　In higher plants and green microalgae, members of the CpSRP pathway guide certain chloroplast
103　proteins to the thylakoid membranes where ALB3 mediates protein insertion in the developing
104　thylakoids. ALB3 belongs to the YidC/Oxa1/Alb3 family of proteins functioning in folding,
105　insertion and assembly of membrane protein complexes in bacteria and in certain eukarytotic
106　organelles, such as mitochondria and chloroplasts (Hennon et al., 2015). The homologs within
107　each subfamily have different C-terminal domains that are crucial for their function and protein-
108　protein interaction. Two homologs belonging to this protein family are found in the chloroplasts
109　of *Arabidopsis thaliana*, ALB3 and ALB4 (Sundberg et al., 1997; Gerdes et al., 2006) and *C.*
110　*reinhardtii*, ALB3.1 and ALB3.2 (Bellafiore et al., 2002). ALB3 mutants of *A. thaliana* have a
111　severe phenotype. They are characterized by white/pale-yellow leaves, are defective in thylakoid
112　membrane development, have strongly decreased pigment content and are unable to survive
113　phototrophically beyond the seedling stage when grown on soil (Sundberg et al., 1997). The *A.*
114　*thaliana* ALB3 insertase is essential for insertion of LHC proteins through the post-translational
115　CpSRP pathway and seems to be involved in co-translational assembly of certain chloroplast
116　encoded membrane proteins (Sundberg et al., 1997; Moore et al., 2000; Kugelmann et al., 2013).

117 Functional data exist also for the two *C. reinhardtii* ALB3 homologs, ALB3.1 and ALB3.2
118 (Bellafiore et al., 2002; Ossenbühl et al., 2004; Göhre et al., 2006). The ALB3.1 of *C. reinhardtii*
119 has been shown to be crucial for insertion of LHC proteins into the developing thylakoid
120 membrane and to play a role in the assembly of D1 reaction center protein into PSII (Bellafiore
121 et al., 2002; Ossenbühl et al., 2004). In contrast to the *A. thaliana* ALB3 mutants, *C. reinhardtii*
122 cells lacking ALB3.1 are still capable of phototrophic growth. The other *C. reinhardtii* ALB3
123 homolog, ALB3.2, is however essential for cell survival and is believed to be associated with the
124 assembly and maintenance of the photosystems (Göhre et al., 2006).

125

126 Important differences have been identified between the function of the ALB3 homologs of
127 organisms within the green lineage. We therefore hypothesized that characterization of diatom
128 ALB3 insertases have the potential to uncover new and unique functional features connected to
129 this protein family. Using a reverse genetics approach, we applied the CRISPR/Cas9 technology
130 to knock out *ALB3b*, encoding one of the two ALB3 proteins present in the diatom
131 *Phaeodactylum tricornutum*. We demonstrate here that ALB3b's primary functional role pertains
132 to insertion of light-harvesting antenna proteins in the developing thylakoid membrane. This,
133 however, does not include antenna proteins functioning in photoprotection. Reduced levels of
134 light-harvesting antenna proteins resulted in changes in the spectral properties, pigment content,
135 growth rate and photosynthetic performance of the cells.

136

137 **RESULTS**

138 Two homologs of the ALB3 insertase were identified in *P. tricornutum* and in all other
139 stramenopiles where sequence data are available (Supplementary Figure S1). Phylogenetic
140 analyses showed that ALB3 proteins in plants/green algae and ALB3 proteins from
141 stramenopiles were clearly divided into two distinct groups (Supplementary Figure S1).
142 Sequence similarity with the two ALB3 proteins with known functions in the green algae *C.*
143 *reinhardtii* could therefore not be used to predict the individual function of the two *P.*
144 *tricornutum* ALB3 proteins (ALB3a and ALB3b). The ALB3a paralog has a basic lysine-rich C-
145 terminal domain (CTD) with similarities to CTD domains in ALB3 proteins in plants and green

146   algae (Supplementary Figure S2). In *A. thaliana* this domain has been reported to interact

147   directly with CpSRP43 and CpSRP54·CpFTSY complexes (Falk et al., 2010; Falk and Sinning,

148   2010; Lewis et al., 2010; Dünschede et al., 2011; Chandrasekar and Shan, 2017). ALB3b

149   proteins in stramenopiles, however, do not contain the lysine-rich CTD but have instead a unique

150   conserved domain (Figure 1). Both *P. tricornutum ALB3* genes (*ALB3a* (Phatr2_43657) and

151   *ALB3b* (Phatr2_46411)) were targeted for CRISPR/Cas9-mediated disruption, but we were only

152   able to generate viable KO-lines for the *ALB3b* gene. Three independent *alb3b* knockout lines

153   (*alb3b-14*, *alb3b-16*, *alb3b-19*) with large insertions of different sizes toward the 5' end of the

154   gene (Supplementary Figure S3) were identified and cultured from single cells. All insertions

155   consisted of fragments of the vectors used for transformation and caused premature stop codons

156   at the N-terminal part of the protein (Figure 1B). To verify that both alleles were mutated and

157   that no WT sequence was present, allele-specific PCR was performed. Both alleles could be

158   amplified in the WT whereas only one allele could be amplified in the mutant strains, indicating

159   larger insertion or deletion events which prevent amplification of the other mutated allele by

160   PCR (Supplementary Figure S4). Complementation of all three *alb3b* KO mutants with a codon

161   modified *ALB3b* (to avoid gene editing) was performed to confirm that the phenotype described

162   below was the result of a lack of a functional ALB3B insertase.

163

164   **Spectral properties of WT and *alb3b* mutants**

165   Previous studies on green algae and plants showed that mutations causing a reduction in the size

166   of the light harvesting antennae result in a pale green color of the chloroplasts (Sundberg et al.,

167   1997; Bellafiore et al., 2002; Polle et al., 2003; Kirst et al., 2012; Kirst et al.; Oey et al., 2013;

168   Gu et al., 2017). The diatom FCP complexes contain, in addition to Chl *a* and *c*, high amounts of

169   Fx responsible for the golden-brown coloration of the diatom cells (Gundermann and Büchel,

170   2014; Büchel, 2015; Wang et al., 2019). The absorption properties of Fx are strongly dependent

171   on the protein environment, and undergo extreme bathochromic shifts upon protein binding,

172   dividing the different Fx molecules into more red, green and blue absorbing complexes

173   (Premvardhan et al., 2009; Premvardhan et al., 2010; Gundermann and Büchel, 2014; Wang et

174   al., 2019). We therefore hypothesized that a distortion of the normal antenna size/structure of *P.*

175   *tricornutum* could result in a visible change in cell coloration. Disruption of the gene encoding

176   the ALB3B insertase did indeed cause a change in coloration from the normal golden brown of

177   the WT cells, to a green coloration, suggesting structural changes of the light harvesting antenna

178   in the *alb3b* KO mutants (Figure 2A).

179

180   To further explore the visual changes in spectral properties in the *alb3b* mutants compared to

181   WT cultures, we recorded the *in vivo* absorbance (Figure 2B) and fluorescence excitation spectra

182   (Figure 2C) for medium light (ML) acclimated cultures. The spectra showed that less light

183   energy in the blue-green region is absorbed and available for photosynthesis in cultures lacking

184   the ALB3b insertase. *In vivo* fluorescence excitation spectra were used to indicate the pigments'

185   relative energy transfer efficiency (ETE) to Chl *a* in the reaction center of PSII (RCII). The

186   differences in the *in vivo* fluorescence excitation spectra between WT and *alb3b* mutants (Figure

187   2C, inset) strongly resembled the absorption characteristics of Chl *c* (peak at 462 nm) and Fx (peak

188   at 520 nm) (Bricaud et al., 2004; Premvardhan et al., 2009; Gundermann and Büchel, 2014),

189   implying a significantly lower contribution in energy transfer from Chl *c* and Fx to RCII in the *alb3b*

190   KO mutants. Smaller differences between WT and mutant strains are expected for the absorption

191   spectra, as these spectra will also include pigments associated with PSI and non-protein bound

192   carotenoids dissolved in the thylakoid membrane that do not transfer absorbed energy to PSII

193   (Lepetit et al., 2010). Even so, the difference in the peak profile for the absorption spectra (Figure

194   2B, inset) matches the difference in the *in vivo* fluorescence excitation spectra confirming a

195   reduction of Chl *c* and Fx in the mutants.

196

197   Low temperature (77 K) fluorescence measurements were performed to clarify the distribution of

198   excitation energy between PSII and PSI in WT compared to *alb3b* mutant cultures (Figure 3). The

199   same samples were excited with either 435 nm (targeting Chl *a* absorption maxima; Figure 3A) or

200   470 nm (targeting antenna pigments (Chl *c* and carotenoids; Figure 3B)). 77 K emission spectra

201   recorded from ML acclimated samples revealed fluorescence emission maxima at 688 nm and 710

202   nm, which are traditionally attributed to PSII and PSI, respectively (Ikeda et al., 2008; Yamagishi et

203   al., 2010; Juhas and Buchel, 2012). In addition, an increase in fluorescence at 710 nm (F_{710})

204   emission at the expense of F_{687} was observed in *P. tricornutum* cells that were in a state of high NPQ

205   (Lavaud and Lepetit, 2013). In WT samples the chosen excitation wavelengths caused a preferential

206   energy transfer to PSII, displaying a relative amplitude of PSII fluorescence emission that was 2.5-

207 fold (435 nm) or 3.3-fold (470 nm) higher than the PSI emission (F_{687}/F_{710}). In contrast, the average

208 F_{687}/F_{710} observed in the *alb3b* mutants were F_{687}/F_{710}= 1.3 (435 nm) or 1.4 (470 nm), implying that

209 excitation energy transfer to PSII was relatively more affected than energy transfer to PSI.

210

211 Effect of lack of ALB3b insertase on the organization of photochemical apparatus

212 The green color of the *alb3b* KO mutants and the combined results from the absorbance,

213 fluorescence excitation and emission spectra suggested that these mutants have an altered

214 functional light-harvesting antenna size. To investigate this in more detail the WT and the *alb3b*

215 KO lines were analyzed using an absorbance difference spectrophotometer (Melis, 1989). The

216 rate of light absorption per second by PSII and PSI was measured by using low intensity actinic

217 light selected by cut-off and interference filters to selectively excite Fx (533 nm) or Chl *a* (670

218 nm), respectively (Table 1). When exiting Fx, the rate of light utilization by the photosystems

219 revealed a severe decrease in the absorption cross section both for PSII and for PSI in the *alb3b*

220 mutant lines compared to WT (Table 1). The functional Chl *a* antenna size of PSII and PSI in

221 the mutants were less affected because of the Chl *a* molecules bound to the photosystem core

222 subunits (Ben-Shem et al., 2003; Nelson and Yocum, 2006; Ago et al., 2016) (Table 1). In

223 accordance with the 77 K data, these data also suggest a more severe decrease of the antenna

224 size of the PSII compared to the PSI (Table 1).

225

226 Organization of the photochemical apparatus was further studied by quantification of PSI (P700)

227 relative to the Chl *a* content of the cells. P700 content was measured from the light induced

228 ΔA_{700} absorbance change at 700 nm attributed to photooxidation of P700. On a P700 basis, there

229 was a substantially lower number of Chl *a* molecules in the *alb3b*, i.e., from 663 Chl *a*/P700 in

230 the WT, down to an average of 425 Chl *a*/P700 in the mutants (Table 1). This directly reflects

231 the lowering of Chl *a* pigments per electron-transport chain (i.e., per P700) in the *alb3b* mutants

232 relative to the WT.

233

234 Western blot was used for examination of the role of the ALB3b insertase in incorporating

235 proteins in the thylakoid membrane. Antibodies specific for antenna proteins (LHCFs and

236 LHCXs) and photosystem subunits (D1, D2 and PsaC) were used, and an antibody against AtpB

237 was employed as loading control. The level of LHCF proteins in the *alb3b* mutants was assessed

238   by an antibody binding to a highly conserved epitope of the LHCF1 to LHCF11 proteins (Juhas
239   et al., 2014), and found to be lowered to about 25% of WT levels in cells grown under both LL
240   and ML conditions (Figure 4A). The relative decline of LHCF proteins is in good agreement
241   with the smaller functional antenna size of PSII, as estimated from the kinetic
242   spectrophotometric measurements using Fx excitation (Table 1). The relative gene expression
243   levels of four *LHCF* genes (*LHCF1*, *LHCF2*, *LHCF5* and *LHCF8*) were examined to determine
244   if the low content of LHCF proteins in the *alb3b* lines could be explained by a strong
245   downregulation of the expression of these genes. Our data showed high gene expression levels
246   (low Ct-values) of the examined *LHCF*s in all lines (Supplementary Table S1). Of the examined
247   LHCF genes, only *LHCF8* was significantly, but moderately, down-regulated in all *alb3b* lines
248   (Supplementary Figure S5). No antibodies are available for detection of LHCR proteins
249   constituting the main LHC protein fraction associated the PSI antenna (Lepetit et al., 2010;
250   Grouneva et al., 2011; Gundermann and Büchel, 2014). However, the smaller functional PSI
251   antenna size in the mutant lines implied that ALB3b plays a vital role also in insertion of LHCR
252   proteins. An antibody (anti-FCP6) against an LHCX (FCP6) of *Cyclotella meneghiniana*, which
253   also cross-react with the *P. tricornutum* LHCX proteins (Juhas et al., 2014), was used for
254   comparison of the relative content of these photoprotective proteins. LHCX1 is crucial for NPQ
255   to take place, whereas LHCX2-3 can provide additional NPQ capacity during high light stress
256   (Bailleul et al., 2010; Lepetit et al., 2017; Taddei et al., 2016; Taddei et al., 2018). LHCX1 and
257   LHCX3 are of highly similar size (21.9 kDa and 22.8 kDa, respectively), therefore complete
258   separation by western blot analysis is challenging. Based on the expression pattern of the LHCX
259   isoforms known from literature, we interpret the proteins detected under both LL and ML
260   conditions to be a mix of LHCX1 and LHCX3 with the major contribution coming from LHCX1
261   under these conditions (Taddei et al., 2016; Taddei et al., 2018). The relative content of the
262   LHCX1+3 proteins in the mutants compared to WT seemed to be unaffected (slightly reduced
263   levels of LHCX1+3 in *alb3b-14*) in both light conditions (Figure 4A). The LHCX2 protein (24.7
264   kDa) was detected at similar levels in WT and *alb3b* lines after 6 h of ML exposure
265   (Supplementary Figure S8B), but it was not detectable in LL or ML-acclimated samples (Figure
266   4A). The strong band of ~22 kDa detected in WT and *alb3b* lines 6 h after the shift from LL to
267   ML (Supplementary Figure S8B) is likely to contain large amounts of LHCX3 in addition to
268   LHCX1 (Taddei et al., 2016; Taddei et al., 2018). Based on Western blot analyses performed on

269   PSI/II core proteins, the lack of a functional ALB3b insertase does not seem to have a negative

270   impact on the incorporation of chloroplast-encoded photosystem subunits (Figure 4B).

271   Preliminary analysis with transmittance electron microscopy (TEM) showed a lower number of

272   thylakoid membranes per chloroplast, but no obvious difference in the thylakoid architecture

273   could be observed in the *alb3b-14* mutant line acclimated to LL (Supplementary Figure S6).

274

Functional properties of the *alb3b* KO mutants

276   To study the capability of the *alb3b* mutant to respond to a shift in light conditions, LL

277   acclimated cells (0 h) were shifted to ML conditions and sampled after 0.5, 6, 24, 48 and 168 h.

278   The pigment content (Figure 5) and photosynthetic performance (Figure 6 and Figure 7) of the

279   acclimating cells were analyzed.

280

Capacity for photoacclimation and photoprotection

282   As expected from the changed coloration and spectroscopic analyses, the *alb3b* KO mutants had

283   a significantly lower content of light-harvesting pigments (LHPs) per cell compared to WT

284   (Figure 5). Even though the content of LHPs in LL-acclimated *alb3b* mutants was already lower

285   than in ML-acclimated WT cells, the LHP concentration in the mutants decreased further as a

286   response to the ML-treatment (Figure 5A-B). This observation implies that the mechanisms

287   controlling the downregulation of the LHPs in response to an increase in available light are

288   independent of the actual pigment concentration in the cells. The *alb3b* mutant lines contained

289   ~40-60 % less Chl *a* and ~60-65 % less Fx in response to the light treatment (Figure 5A-B,

290   Supplementary Table S2).

291   The smaller antennae size of the mutant lines had no negative impact on the cell content of the

292   xanthophyll cycle carotenoids Ddx and Dtx (Figure 5C-D). Both WT and *alb3b* mutant lines

293   showed the expected photoprotective response to a shift to a higher light intensity (Nymark et al.,

294   2009), which could be observed as an immediate rise in Dtx concentration inversely to a

295   decrease in Ddx concentration. The conversion of Ddx to Dtx peaked at the 0.5 h time point as

296   evident by the de-epoxidation state (DES) index (Figure 6A). The DES index decreased and

297   stabilized at a lower level after prolonged exposure to ML, indicating that the algae were

298   acclimating to the new light condition. Although changes in DES index for both WT and mutants

299   followed the same pattern after the shift to higher light intensities, the DES index were higher in

300   the mutants than in WT cultures at all time points. The NPQ capacity of the *alb3b* mutants was

301   initially (approx. two months after isolation of mutated single cells) found to be lowered to

302   around half of that in the WT levels at irradiance levels > 400 µmol m^{-2} s^{-1} (Figure 6B), but when

303   the same experiment was repeated after the cells had been maintained in culture for one more

304   year (approx. 100-150 generations) the differences between WT and mutants had declined for all

305   lines (Figure 6C). Measurements of time-dependent NPQ development in *alb3b* mutants and WT

306   produced highly similar results as when calculating NPQ from rapid light curves (Supplementary

307   Figure S7). The NPQ of *alb3b-16* was closer to WT-levels whereas a lower NPQ was observed

308   in the two other *alb3b* lines. The smaller differences in NPQ capacity between *alb3b* lines and

309   WT led us to also re-analyze the relative LHCF protein content, pigment levels and

310   photosynthetic parameters in LL-acclimated *alb3b* and WT cultures after one more year of

311   growth (Supplementary Figure S8A, Supplementary Figure S9 and Supplementary Figure S10).

312   No major changes were observed for the *alb3b* lines relative to WT cells compared to the initial

313   analyses of these parameters.

314

Photosynthetic performance

316   Variable Chl *a* fluorescence (PAM) was used to calculate the photosynthetic (PSII) efficiency

317   (F_v/F_m,) of WT and mutant lines during the light experiment. In LL acclimated cells, the F_v/F_m

318   were ~0.7 for all lines (Figure 7A), which is around the maximum value expected for algal cells

319   under optimal growth conditions (Falkowski and Raven, 2007). After 0.5 h of ML exposure, both

320   WT and mutant cells showed a modest decrease in F_v/F_m (Figure 7A). The F_v/F_m in the mutant

321   cultures stabilized close to ~0.6 in ML, whereas F_v/F_m in WT cultures increased after prolonged

322   exposure to ML. The maximum relative electron transport rate ($rETR_{max}$) and light saturation

323   index (E_k) values increased as a function of ML exposure time in all cultures (Figure 7C-D), as

324   the photoacclimation mechanisms enabled the cells to utilize the increased amount of light

325   energy available for photosynthesis (Nymark et al., 2009). However, the *alb3b* mutants

326   displayed, on average, a ~30-40% higher $rETR_{max}$ and E_k compared to WT cultures, showing the

327   largest differences during the first part of the light experiment before the cells had been able to
328   downsize the photosynthetic apparatus in response to the increased light intensities. Less
329   pronounced differences in rETR$_{max}$ and E$_k$ were found between WT and *alb3b* cultures at the 24
330   h time point due to a more rapid change in photoacclimation status in WT cells, probably
331   because of a higher cell division rate as described below (Table 3). To further investigate the
332   apparent increased photosynthetic performance of the *alb3b* KO lines indicated by the PAM
333   measurement, light-saturation curves of photosynthesis (P-E curves) based on oxygen evolution,
334   were measured for WT and *alb3b* KO lines acclimated to either LL (Figure 7E) or ML (Figure 7F).
335   The maximum photosynthetic rate (P$_{max}$ (µmol O$_2$/ mol Chl/s), the maximum light utilization
336   coefficient (α) and the saturation intensity (E$_s$) of photosynthesis (P$_{max}$/α (µmol photons m^{-2} s^{-1}))
337   were calculated from the P-E curves (Table 2) (Powles and Critchley, 1980). When normalized to
338   Chl *a*, the mutant lines showed a typical truncated light-harvesting antenna (TLA) - mutant
339   phenotype with higher P$_{max}$ and E$_s$ and slightly lower α compared to WT due to lower functional
340   absorption cross-section caused by the smaller antenna (Kirst et al., 2014). Thus, it should be
341   noted that these results do not indicate a higher photosynthetic performance per cell. In fact, when
342   oxygen evolution was normalized per cell, the mutant lines showed a P$_{max}$ similar to WT
343   (Supplementary Figure S11). Also, the light saturation curves of the *alb3b* KO lines acclimated to
344   LL showed a tendency of declining photosynthetic activity at light intensity > 1000 µmol photons m^{-}
345   ^{2} s^{-1} (Figure 7E).

346

347   **Effect of light intensity on cell growth**

348   Growth parameters were calculated from the exponential phase in batch cultures of LL and ML
349   acclimated cultures (Table 3; Supplementary Figure S12) to investigate how the changes in
350   antennae size and composition affected the cell division rate. The results showed that WT cells
351   grew faster than *alb3b* KO mutants at both light conditions, but a shift from LL- to ML-
352   intensities diminished that growth rate gap between the *alb3b* KO mutants and WT (Table 3), as
353   recently observed in other TLA mutants (Kirst et al., 2014; Formighieri and Melis, 2017). At ML
354   conditions the WT cells already divided at a maximum rate slightly above two cell divisions per
355   day (Fawley, 1984). We hypothesized that if the slower growth rate of the *alb3b* mutants were
356   caused by a lower ability to capture light energy, increasing the light intensities should have a

357 positive effect on growth of the mutant cells. To investigate if a further increase in light intensity
358 could close the growth rate gap, mutants and WT cells were acclimated to HL conditions (480
359 μmol photons m^{-2} s^{-1}). The growth temperature was set to 23°C which supports the highest cell
360 division rate in *P. tricornutum* (Fawley, 1984). During the HL acclimation period (two weeks),
361 the majority of the cells in one of the *alb3b* lines (*alb3b-16*) changed from the fusiform
362 morphotype to a rounded phenotype. The rounded cells showed a tendency for aggregation,
363 making accurate counting necessary for growth rate calculations difficult. The attempt to
364 acclimate *alb3b-16* to HL was repeated after the discovery of the strongly increased NPQ
365 capacity in cells that had been maintained in culture for one year after isolation of single cells,
366 but the HL treatment induced the same change in morphotype as previously observed. The two
367 other *alb3b* lines did not show a change in morphotype during the HL acclimation period or
368 during the following growth rate experiments, but prolonged HL treatment (months) including
369 periods in stationary phase, induced the formation of the rounded cell type also in the two other
370 *alb3b* lines. The same treatment did not provoke the formation of round cells in WT cultures.
371 Growth curves are included in Supplementary Figure S13A. The growth rate calculations from
372 the exponential part of the curve, showed that the WT cells still divided twice per day in HL,
373 whereas the average maximal growth rate of the *alb3b* mutants dropped from 1.2 in ML to 0.8
374 divisions per day under HL (Table 3). The physiological status of the cells, measured as F_v/F_m,
375 was monitored during the length of the growth experiment (Supplementary Figure S13B). The
376 average F_v/F_m in WT cultures during the period of maximal growth, was found to be 0.63. In
377 contrast, the corresponding F_v/F_m value in the *alb3b* mutants were 0.41, pointing to a higher
378 degree of photodamage. In order to investigate presence of oxidative damage, levels of lipid
379 peroxidation were measured for HL acclimated WT and mutant cells (*alb3b-14*, *alb3b-19*). The
380 mutant lines did not show higher levels of lipid peroxidation compared to the WT
381 (Supplementary Figure S14). Similar levels of xanthophyll pigments in the mutant compared to
382 the WT could explain these results, considering their role in the stabilization and protection of
383 the thylakoid membrane lipids from peroxidation (Hauvaux et al., 2007).

384 **Complementation studies of *alb3b* mutants**

385 A plasmid containing the codon modified *ALB3b* under control of its native promoter was
386 introduced to the three *alb3b* lines by biolistic bombardment. As a result, 70 of in total 75

387　transformed colonies regained their brown coloration. Six brown colonies (two colonies derived

388　from each of the three complemented lines) were randomly picked and subjected to PCR analysis

389　followed by sequencing. The introduction of the modified *ALB3b* gene and the absence of WT

390　sequence were confirmed (Supplementary Figure S15). Three brown colonies (representing each

391　of the three complemented mutant lines) were cultured for analyses of pigment and LHCF

392　content. The results showed that the WT phenotype was recovered by introduction of the

393　modified *ALB3b* gene (Figure 8).

DISCUSSION

Effects of loss of the *P. tricornutum* ALB3b insertase

The significantly lower level of antenna proteins belonging to the LHCF group (Figure 4A) indicate that the primary role of the *P. tricornutum* ALB3b insertase is the efficient integration of the main LHC proteins into the thylakoid membrane. However, a small functional antenna size is still assembled, implying a phenotype where some LHC proteins can be inserted through other thylakoid membrane insertion pathways, or that some functional redundancy exists between ALB3b and the uncharacterized diatom homolog ALB3a. The mainly unaffected levels of photoprotective LHCX proteins found in *alb3b* mutants (Figure 4A) clearly indicate presence of other integration pathway(s) for antenna proteins. The lower level of LHPs and smaller functional antenna size, the changed spectral properties and the increased light saturation level, can be seen as effects of the lower amount of antenna proteins causing a truncated light harvesting antenna. The phenotypic traits listed above are characteristic for TLA-phenotype mutants, previously generated in cyanobacteria, green microalgae and higher plants (Polle et al., 2003; Kirst et al., 2012; Kirst et al., 2012; Kirst et al., 2014; Formighieri and Melis, 2017; Gu et al., 2017; Kirst et al., 2017; Kirst et al., 2018). TLA mutants have been shown to grow at relatively similar rates as WT when enough light energy is available (Bellafiore et al., 2002; Polle et al., 2003; Kirst et al., 2014; Gu et al., 2017).

The slow growth of the *alb3b* mutants compared to WT cells might be partially explained by a reduced ability to capture light energy, since an increase in light intensity from 35 (LL) to 200 µmol photons m^{-2} s^{-1} (ML) diminished the difference in growth rate between WT and mutant by a factor 2. If the smaller antenna size of the mutants were the sole reason for the slow growth rate, a further increase in irradiance should further diminish the difference in growth between WT and mutant. Instead, analyses of algae cultures acclimated to HL (~480 µmol photons m^{-2} s^{-1}) revealed a negative effect on cell division rate, photodamage of the *alb3b* mutants and induction of a round cell phenotype. The round or oval cell shape has previously been reported to be associated with prolonged exposure to abiotic stress (De Martino et al., 2007; De Martino et al., 2011; Herbstova et al., 2017). The apparent increased photosynthetic capacity estimated for *alb3b* mutants at both LL and ML light conditions seems counter intuitive if the *alb3b* mutants are high light sensitive. However, these data are calculated from light-response curves where the algae are subjected to high

424 light intensities for relative short periods of time (minutes). The high light experienced by the algae

425 during the generation of light-response curves might be too short for extensive photodamage to

426 occur. However, mutants acclimated to LL conditions did show signs of photoinhibition observed

427 as a decrease in oxygen production when exposed to light intensities > 1000 µmol photons m^{-2} s^{-1}

428 (Figure 7E).

429

430 NPQ is an important photoprotective mechanism providing the ability to dissipate excessively

431 absorbed energy harmlessly as heat during high light exposure. In the *alb3b* mutants the NPQ

432 capacity was reduced compared to WT levels (Figure 6B-C and Supplementary Figure S7),

433 suggesting a reduced capability to handle prolonged high light exposure. Several studies show a

434 convincing relationship between the amount of both LHCX and Dtx and the capacity for NPQ, and

435 the presence of LHCX proteins and the conversion of protein bound Ddx to Dtx has been found to

436 be essential for NPQ to take place (Lavaud et al., 2002; Bailleul et al., 2010; Lepetit et al., 2012;

437 Lepetit et al., 2013; Lepetit et al., 2017; Taddei et al., 2018). The level of LHCX proteins and the

438 content of the xanthophyll cycle pigments (Ddx+Dtx) were not negatively affected by the lack of

439 ALB3b insertase. However, Ddx and Dtx are found in three different pools in diatoms, one located

440 in a lipid shield around the FCPs, and two that are bound to antennae proteins connected to PSI or

441 the peripheral FCP antenna, respectively (Lepetit et al., 2010). Only the protein bound fraction of

442 the peripheral antennae contribute to NPQ after conversion of Ddx to Dtx (Lepetit et al., 2010).

443 Because of the potential to store xanthophyll cycle pigments in the lipid phase of the thylakoid

444 membrane, the amount of accumulated Ddx+Dtx that are protein bound might still be reduced even

445 though the cell concentrations in the *alb3b* lines are similar or higher than in WT. The molecular

446 role of LHCX and Dtx in NPQ is still elusive, and no data exists about the precise localization of

447 FCPs or the LHCX proteins. The latest models for NPQ in diatoms suggest that there are two

448 quenching sites (Q1 and Q2) present in the diatom thylakoids (Miloslavina et al., 2009; Büchel,

449 2014; Lavaud and Goss, 2014; Goss and Lepetit, 2015; Giovagnetti and Ruban, 2017). NPQ at Q1

450 is believed to involve physical detachment of FCP oligomers from PSII that in *P. tricornutum* can

451 be measured as an increase in 77 K emission at 710 nm and as a decrease of PSII cross section

452 (Lavaud and Lepetit, 2013; Giovagnetti and Ruban, 2017), whereas Q2 seems to take place in FCPs

453 functionally connected to PSII, and involve antennae reorganization and aggregation of LHC

454 trimers (Miloslavina et al., 2009; Büchel, 2014; Lavaud and Goss, 2014; Giovagnetti and Ruban,

16

455 2017). Q2 is suggested to be dependent on the presence of protein bound Dtx and provides a much

456 higher level of NPQ compared to Q1 (Giovagnetti and Ruban, 2017). Despite the comparable

457 content of photoprotective antenna proteins and pigments in WT and *alb3b* mutants, the strong

458 decrease in *alb3b* antennae size might disturb crucial protein-pigment or protein-protein (e.g LHCF-

459 LHCX) interactions potentially necessary for effective antenna aggregation (Q2) and lower the pool

460 of detachable antenna involved in Q1. This might lead to the lower NPQ capacity observed in the

461 *alb3b* mutants. However, the difference in NPQ capacity between *alb3b* lines and WT decreased

462 after the *alb3b* lines had been maintained in culture for one additional year (approx. 100-150

463 generations). The increase in NPQ compared to WT was especially prominent for *alb3b-16*. No

464 major differences in pigment or LHCF content between the individual *alb3b* lines or changes in the

465 pigment or LHCF ratios between *alb3b* and WT were observed that could explain the changes in

466 NPQ capacity over time. The different NPQ levels in the mutants and the general increase in NPQ

467 over time in the *alb3b* lines compared to WT levels can therefore not be explained by changes in

468 antenna size over time. Giovagnetti and Ruban (Giovagnetti and Ruban, 2017) showed that the

469 amount of antenna detached are not proportional to the level of NPQ, and that the NPQ can

470 continue to increase without a further reduction of the PSII cross-section. We therefore suggest that

471 the increase in NPQ over time is caused not by a larger pool of detachable antenna, but that the

472 *alb3b* lines, over many generations, have been able to increase their capacity for NPQ at Q2 through

473 an unknown mechanism.

474

475 **Role of diatom ALB3b in integration of nucleus and plastid encoded proteins compared to**
476 **ALB3 in green alge/plants**

477 The *P. tricornutum* ALB3b showed functional similarities with the *C. reinhardtii* homolog ALB3.1

478 (Bellafiore et al., 2002; Ossenbühl et al., 2004). Both the diatom ALB3b and the green algae

479 ALB3.1 play a role in insertion of LHC proteins into the thylakoid membrane (Bellafiore et al.,

480 2002; Kirst and Melis, 2014), and loss of the insertase causes a notably smaller antenna size

481 (Bellafiore et al., 2002). In addition, *C. reinhardtii* cells lacking ALB3.1 contain a significantly

482 increased fraction of highly stable membrane inserted, but unassembled D1 protein (Ossenbühl et

483 al., 2004). The D1 content in *C. reinhardtii alb3.1* mutants were half of that of WT cells. Based on

484 the above described findings, an additional role in assembly of D1 into PSII was identified in green

485 microalgae (Bellafiore et al., 2002; Ossenbühl et al., 2004). Subunits of PSI (PsaC), PSII (D1, D2)

486   and ATP synthase complex (AtpB) were not negatively affected by the absence of the ALB3b

487   insertase in diatom cells (Figure 4B), but our analyses does not discriminate between unassembled

488   proteins in the thylakoid membrane and proteins that are incorporated into photosynthetic

489   complexes. More extensive protein analyses would be necessary to rule out a role of the diatom

490   ALB3b insertase in integration/assembly of chloroplast encoded thylakoid membrane proteins.

491   Assembled PSII complexes are fully functional in both *C. reinhardtii* (Ossenbühl et al., 2004) and

492   *P. tricornutum alb3b* mutants (Figure 7A). We detected no differences in photosynthetic efficiency

493   in LL acclimated cells between WT and mutants. This implies that even though the *alb3b* KO lines

494   have a truncated antenna size, there is no difference in the probability of the trapped excitation

495   energy to be used for photochemistry between WT and mutants. However, a less efficient repair of

496   PSII from photodamage (Guenther and Melis, 1990) and an associated slower replacement of

497   damaged D1, could explain the on average ~12-14% lower F_v/F_m measured in *alb3b* mutants during

498   prolonged ML exposure, and the on average ~36% lower F_v/F_m observed in HL acclimated mutant

499   cells. An efficient PSII repair mechanism including a more frequent replacement of photodamaged

500   D1 is required during such conditions (Baroli and Melis, 1996; Theis and Schroda, 2016).

501   Alternatively (or additionally), the PSII of the *alb3b* mutants might be more susceptible to

502   photodamage because of the altered light harvesting antennae disturbing the normally efficient NPQ

503   mechanism (Figure 6B-C) functioning in this algae (Lavaud and Goss, 2014). However, the

504   transformation of the normally fusiform *alb3b-16* line into the rounded morphotype in HL

505   regardless of having a lower (Figure 6B) or more similar (Figure 6C and Supplementary Figure S7)

506   NPQ capacity as WT indicates that there are other reasons for why *alb3b* mutants are sensitive to

507   HL.

508

509   The *Arabidopsis thaliana alb3p* mutant has also been reported to be photosensitive. *Alb3p* require

510   very low light intensities (12 μmol photons $m^{-2} s^{-1}$) to produce detectable levels of photosynthetic

511   complexes like LHC trimers and PSII monomers and dimers (Kugelmann et al., 2013). To explain

512   the severe phenotype of the *Alb3p* mutants, additional functions beyond the CpSRP pathway have

513   been suggested for ALB3p (Kugelmann et al., 2013). Based on phenotypic similarities between the

514   *Alb3p* and mutants defective in carotenoid synthesis it has been speculated that the ALB3p has a

515   role in integration and assembly of carotenoids into photosynthetic complexes (Kugelmann et al.,

516   2013). The slow growth of the *P. tricornutum alb3b* mutants that cannot be compensated by

517   increased light intensities, and the susceptibility to prolonged high light exposure, suggest

18

518   additional roles for the ALB3b insertase. A future comparison with other types of *P. tricornutum*

519   TLA mutants will be valuable for dissecting primary effects of the absence of ALB3b, from the

520   secondary effects of having a truncated light harvesting antenna size.

521

522   **CONCLUSION**

523   Our results show that ALB3b is essential for assembly of a full-size light harvesting antenna in

524   diatoms. In higher plants and green algae, ALB3 insertases are a part of the CpSRP pathway and the

525   basic lysine-rich CTD is necessary for the interaction with other members of the pathway

526   (Bellafiore et al., 2002; Chandrasekar and Shan, 2017). We also identified this domain within the

527   ALB3a proteins of the stramenopiles, but not in the ALB3b proteins which has a unique CTD

528   domain. The LHC specific chaperon CpSRP43 is one of the ALB3's known interaction partners

529   through its lysine-rich CTD domain, but neither we nor others (Träger et al., 2012) could identify

530   this chaperon in diatoms or other stramenopiles. Also, the *P. tricornutum* CpSRP54 mutant was not

531   reported to have a changed coloration, only to be light sensitive (Nymark et al., 2016). The different

532   CTD domain in ALB3b proteins, the absence of CpSRP43 and the unchanged coloration of the

533   diatom CpSRP54 mutant, imply that the ALB3b proteins have other interaction partners than

534   ALB3a and ALB3 of plants/green algae. A hypothetical model for the role of diatom ALB3

535   insertases is presented in Figure 9. For verification of the model, a more thorough investigation of

536   the *P. tricornutum* CpSRP54 mutant and characterization of diatom FTSY mutants should be

537   performed. This will clarify if ALB3b is part of the post-translational CpSRP pathway, or if diatom

538   LHC proteins are guided to ALB3b through other mechanisms.

539

540   **MATERIALS AND METHODS**

541   An axenic *P. tricornutum* culture originating from the sequenced clone Pt1 8.6 (CCMP2561) was

542   obtained from the culture collection of the Provasoli-Guillard National Center for Marine Algae and

543   Microbiota (NCMA), Bigelow Laboratory for Ocean Sciences.

544

545   **Experimental conditions**

546   Axenic culturing of *P. tricornutum* WT cells and the three *alb3b* KO lines (*alb3b-14*, *alb3b-16* and

19

547   *alb3b-19*) were performed as described previously unless otherwise stated (Nymark et al., 2009).

548   Cell cultures were grown at 15°C under continuous cool white fluorescent light at scalar irradiance

549   (E_{PAR}) of ~35 µmol photons $m^{-2} s^{-1}$ (LL), or ~200 µmol photons $m^{-2} s^{-1}$ (ML). For the high light

550   (HL) experiment the WT and the three independent *alb3b* KO lines were acclimated to 480

551   µmol $m^{-2} s^{-1}$ and grown at 23°C in a Vötsch VB 1514 plant growth chamber (Vötsch

552   Industrietechnik GmbH, Germany) equipped with metal halide lamps (Powerstar HQI-BT 400

553   W/D). The cultures were kept in the exponential growth phase for at least three weeks under

554   these conditions to ensure that all cells were fully acclimated prior to conducting measurements.

555

556   For the spectrophotometric and kinetic analysis cells were grown in F/2 enriched artificial seawater

557   media (Guillard and Ryther, 1962). To avoid carbon limitation during growth the media were

558   supplemented with $NaHCO_3$ (final concentration of 23.5 mM, pH=7.4). Cultures were grown at 25

559   °C in 2 L glass bottles constantly stirred to ensure homogenous growth. Continuous illumination

560   was provided by white fluorescent LED light tubes at ML. For the measurements 80-85% of the

561   total culture volume was harvested during the mid-exponential growth phase.

562

Growth rates

563

564   Growth rates were estimated in batch cultures of WT and *alb3b* KO lines (three biological

565   replicates) acclimated to LL, ML or HL using a starting concentration of 100.000 (ML, HL) or

566   200.000 (LL) cells/ml. Counting was performed either manually using a Bürker-Türk counting

567   chamber after fixation with Lugol`s solution (LL samples) or with a BD Accuri C6 Flow

568   Cytometer (BD Bioscience; ML and HL samples). For the latter, glutaraldehyde (2% final solution)

569   was used for fixation of cells. Samples were excited by a 20 mW 488 nm Solid State Blue laser

570   and chlorophyll fluorescence was measured by a >670 nm optical filter (FL3). The average

571   maximum growth rates (cell division/day) were calculated by using a mean of the growth rates

572   from the three biological replicates during the exponential phase.

573

Phylogenetic analyses

574

575   ALBINO3 proteins in the NCBI (National Center for Biotechnology Information) protein database

576   and from the iMicrobe transcriptome database (https://www.imicrobe.us/) were selected for

577   phylogenetic analyses. Accession numbers for the protein sequences used in the analysis are listed

578 in Supplementary Table S4. The analysis involved 47 ALB3 proteins from plants and algae, each

579 species was represented with two ALB3 paralogs (ALB3.1/ALB3.2 or ALB3a/ALB3b). The protein

580 alignment was generated by using the ClustalX program (Thompson et al., 1997) and manually

581 refined in GeneDoc 2.7.000 (Nicholas et al., 1997). The evolutionary relationships were estimated

582 using Maximum Likelihood (ML) method based on the Le-Gascuel model (Le and Gascuel, 2008)

583 and the Neighbor-Joining method (Saitou and Nei, 1987). The initial tree for both ML and NJ

584 analyses was obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix

585 of pairwise distances, estimated using a JTT model and the trees with best topology were selected.

586 For the ML-analyses a discrete Gamma distribution was used to model evolutionary rate differences

587 among sites (using 5 categories). All positions with less than 80% site coverage were eliminated.

588 Tree branch confidence values were calculated by running 1000 bootstrap replicates for NJ and 100

589 replicates for ML. The phylogenetic analyses were conducted in MEGA7 (Kumar et al., 2016).

590

591 CRISPR/Cas9 gene editing of the ALB3b insertase

592 All steps for performing CRISPR/Cas9 editing of the *ALB3b* insertase gene (Phatr2_46411;

593 XM_002180751) including selection of target site, ligation of adapter for target of interest into the

594 pKS diaCas9-sgRNA plasmid (Nymark et al., 2016), transformation of diatom cells, and screening

595 and identification of cells with biallelic mutations, were performed as described in the published

596 protocol for CRISPR/Cas9 gene editing in *P. tricornutum* (Nymark et al., 2017). *ALB3b* specific

597 oligos for creation of the adapter inserted into the sgRNA cassette of the CRISPR/Cas9 vector, and

598 primers used for screening of cells with CRISPR/Cas9-mediated mutations, are presented in

599 Supplementary Table S5. Three *alb3b* KO lines named *alb3b-14*, *alb3b-16* and *alb3b-19* were

600 selected for functional characterization. These three selected lines were checked for off-target

601 mutations by PCR amplification and sequencing of the regions containing the five most likely off-

602 target sites. To identify potential off-target sites, a custom-made Perl based script was used to

603 search the genome for sites with high homology to seed (PAM-proximal) region of the target site.

604 The script uses a string-based approach, which allows for up to 3 mismatches in the seed region.

605 Off-targets are ranked after their similarity to the target site as well as the position of the

606 mismatches. No off-target mutations were found at any of the investigated sites. The Phatr2 ID for

607 the genes containing the potential off-target sites and primers used for the screening process are

608 listed in Supplementary Table S5.

Allele-specific PCR

Allele-specific PCR was performed as an additional control as previously described (Serif et al., 2017). In short, primers for PCR were derived which include an allele-specific difference on the 3′ terminal base (see primers in Supplementary Table S5), thereby preventing polymerases without proofreading function from amplifying the respective other allele. Both alleles were amplified separately using HiDi polymerase (myPols, Konstanz, Germany) according to the manufacturer's instructions.

Isolation of thylakoid membranes

Cells were harvested by centrifugation at 1000 g for 8 min at 4 °C. The pelleted cells were resuspended in 50 mM Tricine – NaOH (pH 7.8) in ice-cold isolation buffer containing 300 mM sucrose, 5 mM MgCl, 10 mM NaCl, 2% PVP (w/v), 0.1% BSA (w/v) and 5 mM ascorbic acid. The pellet was washed twice with the described buffer to remove residual salts from the growth media. Cells were broken using a Branson 250 sonicator (pulse mode, 50% duty cycle, output power of 5) with a precooled tip for 45 s followed by 1 min of cooling in dim light. This process was repeated 4 times to ensure rupture of the majority of the cells. Unbroken cells were removed by centrifugation at 6500 rpm for 10 min at 4 °C. The thylakoid suspension was centrifuged at 75,000 g for 45 min at 4°C using a Beckman Coulter ultra-centrifuge. The thylakoid pellet was resuspended in 5 ml of ice-cold Tricine-NaOH (pH 7.8) buffer containing 10 mM NaCl and 5 mM Mg_2Cl. Samples were measured immediately upon preparation.

Spectrophotometric and kinetics analysis

Photosystem kinetics and PSI quantitation analysis were performed using a laboratory-constructed absorbance difference spectrophotometer (Melis and Brown, 1980; Melis, 1989). The premise for this method is that, under light limiting conditions, the rate of primary photochemistry is directly proportional to the light-harvesting antenna size (Melis, 1989). PSI (P_{700}) content was measured from the light-induced ΔA_{700} using a differential extinction coefficient of 64 mM^{-1} cm^{-1} (Hiyama and Ke, 1972). Actinic excitation was provided in the red region of the spectrum using a transmittance interference 670 nm filter combined with a yellow cut-off filter (CS 3-69). The

639   reaction mixture contained 50-100 μM Chl *a*, 0.02% SDS (w/v), 250 μM methyl viologen (MV)

640   and 2.5 mM Na-ascorbate. The sample was illumined once prior to measuring to ensure oxidation

641   of Cytochrome c_6 and possibly of Cytochrome *f*. 2-3 experimental replicates were measured, with

642   at least three technical replicates taken. Chl *a* concentration in the samples was calculated after

643   extraction in 90% acetone (v/v) for half an hour in the dark using the Jeffrey-Humphrey equation

644   for diatoms (Jeffrey and Humphrey, 1975). Photocatalytic kinetics of the two photosystems were

645   measured based on Chl *a* fluorescence induction for PSII and P_{700} oxidation for PSI (Melis, 1989).

646   Actinic illumination was provided in the red and green regions of the spectrum using narrow

647   interference filters with transmittance peaks at 670 nm and a 533 nm. These filters were chosen

648   after examination of the thylakoid absorbance spectra so that the 670 nm filter would excite

649   predominantly Chl *a,* whereas the 533 nm filter would excite Fx and other carotenoids. Incident

650   light intensity provided was 12 μmol photons m^{-2} s^{-1} in the green and 2.1 μmol m^{-2} s^{-1} in the red

651   region. The reaction mixture for the fluorescence kinetic measurements contained approximately 5-

652   10 μM Chl *a* and 20 μM 3-(3,4-dichlorophenyl)-1,1-dimethylurea (DCMU), and that for the P700

653   oxidation kinetics contained 100-200 μM Chl *a*, 250 μM MV and 20 μM DCMU.

654

Absorbance spectra

656   To avoid light scattering, absorption spectra were measured from thylakoid membrane extracts.

657   Prior to measurement, the samples were placed in darkness in an ice bath to avoid thermal

658   breakdown of thylakoid structure. Absorbance spectra of all extracts were scanned

659   spectrophotometrically from 400 to 750 nm with a Shimadzu UV-1800 UV−visible

660   spectrophotometer. The resuspension buffer was used as a blank and for baseline calibration.

661

In vivo fluorescence excitation

663   *In vivo* fluorescence excitation spectra (400-700 nm) were measured as described previously using a

664   Hitachi F-3000 spectrofluorometer (Nymark et al., 2013). Spectra were obtained by recording the

665   Chl *a* fluorescence intensity (Chl *a* fluorescence from PSII) at 1 nm spectral resolution (5 nm

666   bandwidth) at a fixed wavelength of emission (730 nm, 5 nm bandwidth). The emission of light

667   was measured as a function of absorbed light at different wavelengths for ML acclimated cultures.

668   All spectra were normalized to the red emission maximum of Chl *a* of the WT cultures, so as to

669   study the differences in excitation energy transfer efficiency (ETE) by the main photosynthetic

670 pigments Chl *a*, Chl *c* and Fx in the blue-green part of the PAR spectrum, where they exhibit their

671 maximum absorption.

672

77 K chlorophyll fluorescence emission measurements

674 Low-temperature fluorescence emission spectra were recorded for three biological replicates of ML

675 acclimated cell cultures using a custom-made 77 K fluorometer (Lamb et al., 2015).

676 Monochromatic LEDs with an emission centered around either 435 nm (LED435-12-30, Roithner

677 LaserTechnikor) or 470 nm (LED470 Roithner LaserTechnikor) were used as excitation

678 wavelengths. Fluorescence emission spectra were recorded between 600 and 800 nm. Samples were

679 adjusted to a Chl concentration of 1 µg/mL, transferred to glass tubes and frozen in liquid nitrogen

680 before measuring the 77 K fluorescence emission. All spectra were normalized to the WT emission

681 spectrum at 710 nm.

682

Protein isolation, SDS-PAGE and Western blot analysis

684 WT and *alb3b* mutant cultures acclimated to either LL or ML (three biological replicates for each

685 line and light condition) were harvested by filtration (Durapore Membrane Filters, pore size 0.65

686 µm; Merck Millipore). Filters were transferred to 2 ml tubes (Sarstedt) and 1 ml F/2 medium was

687 added. The tubes were vortexed for 10 s for resuspension of the cells, before removal of filters and

688 centrifugation of re-suspended cells at 16000 *g* for 1 min at 15 °C. The supernatant was removed

689 and the remaining pellet was flash frozen in liquid nitrogen and stored at -80 °C. A five mm pre-

690 cooled stainless-steel bead (QIAGEN) was added to each of the tubes with frozen cell pellets, and

691 the cells were mechanically broken and homogenized in two steps using the TissueLyser system

692 (QIAGEN). The samples were first placed in a precooled (-80 °C) adapter set followed by cell

693 disruption for 2 min at 25 Hz. Before the second shaking step (8 min at 25 Hz), the samples were

694 transferred to a room temperature (RT) adapter set and 700 µl lysis buffer (50 mM Tris, pH 6.8,

695 2% SDS) were added according to Juhas et al. (Juhas et al., 2014). Insoluble material was removed

696 by centrifugation (100 *g* for 30 min at 4 °C). The supernatant was transferred to new tubes and the

697 protein concentration was determined using the DC Protein Assay kit (BioRad) following the

698 manufacturer's instructions. In addition to the whole cell extracts, lysates were also obtained from

699 thylakoids isolated from cell cultures acclimated to either LL or ML conditions. Thylakoids were

resuspended in lysis buffer (50 mM Tris, pH 6.8, 2% (w/v) SDS) and protein extracts were obtained as above (the first step for cell breakage was omitted). Proteins were resolved on 12% or 15% SDS–PAGE gels, depending of the size of the protein of interest. 10 µg of the protein extracts were loaded onto the gel lanes. Western blot analyses were performed on either total protein extracts (detection of LHCF and LHCX proteins) or thylakoid extracts (detection of D1, D2 and PsaC proteins). The PsaC antibodies produced a signal only when using thylakoid extracts, whereas the antibody recognizing LHCX proteins produced optimal results when using whole cell extracts. LHC proteins and photosystem subunits were therefore analyzed in different extracts. The signal generated by AtpB polyclonal antibodies was used as loading controls on each blot, in addition to Coomassie stained gels that were run in parallel. 10 µg of the protein extracts were loaded onto the gels. Proteins were detected with the following antibodies: anti-D1 (AS05 084 Agrisera; 1:20000), anti-D2 (AS06 146 Agrisera; 1:5000), anti-PsaC (AS10 939 Agrisera; 1:1000), anti-AtpB (AS05 085, Agrisera; 1:4000), anti-LHCF1-11 (1:1000) and anti-FCP6 (LHCX; 1:1000) (kind gifts from C. Büchel, University of Frankfurt, Germany (Juhas et al., 2014)). Primary antibody incubation was performed overnight at 4°C for all antibodies. Polyclonal Goat Anti-Rabbit Immunoglobulins/Biotinylated (Dako) was used as secondary antibody with an incubation time of 2 h in RT, followed by incubation with Horseradish Peroxidase Streptavidin (Vector Laboratories) for 1 h also in RT. Protein-antibody cross-reactions were visualized with SuperSignal West Pico PLUS Chemiluminescent Substrate (Thermo Scientific) and documented with a G:BOX ChemiXRQ gel doc system (Syngene).

Transmission electron microscopy

Electron microscopy was used to examine the status of the thylakoid architecture in the *alb3b* mutant lines. WT and *alb3b-14* cell cultures acclimated to LL were harvested by a light centrifugation step (4000 g for 10 min) and fixated over night at room temperature in a F/2 medium buffer containing 2.5% Glutaraldehyde and 2% paraformaldehyde. Pellets were washed three times in F/2 medium buffer solution and embedded in a 5% gelatin solution. After post-fixation in 2% osmiumtetraoxide and 1.5% kaliumferrocyanid the samples were dehydrated in a gradient of ethanol. Samples were thereafter embedded with epoxy resins based on Bozzola and Russell's protocol (Bozzola and Russell, 1999) and sectioned with an ultramicrotome. Images were taken using a Tecnai 12 transmission electron microscope operating at 80KV. Images were captured using a MORADA CCD camera.

732

Measurements of malondialdehyde content

The malondialdehyde (MDA) content was determined using the Lipid Peroxidation (MDA) assay kit (Sigma-Aldrich). The MDA concentration was measured based on its reaction with thiobarbituric acid (TBA) and used as an index of lipid peroxidation. WT, *alb3b (alb3b-14, alb3b-19)* mutant cultures (three biological replicates for each line) acclimated to HL were harvested by filtration as described above. The cell pellet was resuspended in the MDA lysis buffer. To ensure complete lysis the cells were briefly sonicated. Thereafter, the MDA content was determined based on the manufacturer's instructions. In parallel samples were collected and manually counted to determine cell concentration.

Isolation of total RNA and quantitative real-time PCR

Three biological replicates of LL-acclimated WT and *alb3b* mutant cultures were harvested for isolation of total RNA in parallel to the samples harvested for protein analyses as described above. Total RNA isolation, quantification and verification of RNA integrity were performed as described in Nymark et al. (Nymark et al., 2009). Reverse transcription of RNA was performed with the QuantiTect Reverse Transcrition kit (Qiagen) following the recommended protocol. 1 μg of total RNA was used in each reaction. Quantitative real-time PCR (qRT-PCR) analysis were performed as described in Nymark et al. (Nymark et al., 2009) for calculation of relative expression ratios of four *LHCF* genes (*LHCF1*, *LHCF2*, *LHCF5* and *LHCF8*). The geNorm module in the qBasePLUS software (Biogazelle) was used for determining the expression stability of candidate reference gene. Based on the stability analysis *RPS5* (Phatr2_42848) and *DLST* (Phatr2_45557) were picked as reference genes (Nymark et al., 2013; Valle et al., 2014). LinRegPCR software (Ramakers et al., 2003; Ruijter et al., 2009) was used to calculate mean PCR efficiency per amplicon and cycle threshold (Ct) values per sample. These data were imported into the qBasePLUS software (Biogazelle), which calculated relative expression ratios (given as Calibrated Normalized Relative Quantities (CNRQ)) and performed statistical analyses on the results. The one-way ANOVA test integrated in the qBasePLUS software was used to evaluate the significance of the estimated relative expression ratios. Forward and reverse primers are listed in Supplementary Table S5.

Light shift time-series experiments

26

763   LL acclimated WT and *alb3b* KO lines were transferred to ML conditions and sampled after 0.5, 6,

764   24, 48 and 168 h following the shift in growth light intensity. LL samples (0 h) were harvested as

765   controls. Three biological replicates were set up for each line and time point to reach a cell

766   concentration of maximum 1 x 10^6 cells/mL at the day of harvesting. Samples were harvested for

767   pigment analyses, monitoring of cell concentrations, variable *in vivo* Chl *a* fluorescence (PAM), and

768   protein analyses.

769

770   **Pigment analyses**

771   The HPLC pigment analysis was performed according to Rodriguez et al. (Rodriguez et al., 2006)

772   using a Hewlett-Packard HPLC 1100 Series system. Pigment values from the HPLC analysis were

773   calculated as fmol pigment per cell. Cell numbers were calculated from flow cytometer counts as

774   described above.

775

776   **Measurements of photosynthetic parameters**

777   A PhytoPAM (System I, Walz, Germany) was used to measure variable Chl *a* fluorescence of the

778   harvested samples. The photosynthesis *vs.* irradiance relationship was obtained as described

779   previously (Nymark et al., 2009). An additional step at 1216 µmol photons $m^{-2} s^{-1}$ was added for the

780   samples that had been treated with ML for 1 week to ensure that light saturation levels were

781   reached. The maximum quantum yield of PSII (F_v/F_m), the maximum relative electron transport

782   rate ($rETR_{max}$), the maximum light utilization coefficient (α) and the light saturation index (E_k) were

783   calculated as described before (Nymark et al., 2009). The $rETR_{max}$ is an estimate of the maximum

784   photosynthetic capacity of the cells (~P_{max}), whereas the light saturation index E_k ($rETR_{max}/α$) is a

785   proxy for the threshold irradiance that separates light-limited and light-saturated photosynthesis

786   (Genty et al., 1989; Sakshaug et al., 1997). F_m at low light intensities is commonly observed to be

787   lower than the F_m' level under low actinic light in diatoms (Serôdio et al., 2006; Cruz and Serôdio,

788   2008; Cruz et al., 2011). NPQ was therefore calculated from the light-response curve from LL

789   acclimated samples, using the maximum F_m' level (F_m'$_{max}$;) instead of F_m as follows: NPQ =

790   (F_m'$_{max}$/ F_m') − 1 (Serôdio et al., 2006; Kalaji et al., 2017). NPQ development over time was

791   additionally calculated from LL acclimated cells exposed to 5 min of actinic light at an intensity

792   setting of 832 µmol photons $m^{-2} s^{-1}$. For the HL experiment, F_v/F_m was measured with an AquaPen-

793   C (Photon System Instruments) at the end of a 30 min dark acclimation period to relax the fast-

794 reversible component (qE) of NPQ so that only the photoinihibitory slowly reversible quenching
795 (qI), caused by damaged PSII reaction centers, would influence the F_v/F_m value.

796 Oxygen evolution was measured at 15 °C using a S1 Clark Type polarographic oxygen electrode
797 (Hansatech) increasingly illuminated with a 35 W cool white spot LED. The measurements were
798 done on cultures acclimated both to LL and ML. 2 ml cell suspension from mid-exponential phase
799 culture was added to a stirred chamber with temperature control and supplemented with sodium
800 bicarbonate (30 µl of a 0.5 M solution) so that the oxygen production would not be limited by
801 carbon availability. Prior to measuring, the Chl *a* concentration in the sample was adjusted to a
802 concentration lower than 1.2 µM to avoid cell shading in the chamber. Simultaneously, cell
803 concentration of the samples was determined by Flow cytometry counting. Oxygen consumption in
804 darkness was measured as a starting baseline, thereafter the sample was exposed to gradually
805 increasing light intensities and the oxygen evolution was measured continuously for at least 10 min.
806 Each light intensity was adjusted by measuring the light intensity in the middle of the electrode
807 chamber with a spherical US-SQS sensor (Waltz).

808

809 Complementation of *alb3b* KO lines

810 A modified version of the *ALB3b* gene was synthesized together with its native promoter by
811 GeneArt® Services Thermo Fisher Scientific Inc (Supplementary Figure S16). Modifications
812 consisted of changes of the codon usage in the PAM and target region of the *ALB3b* gene to avoid
813 gene editing by the functional CRISPR/Cas9 system incorporated into the genome of the *alb3b* KO
814 lines. MssI sites were included at the 5' and 3' ends of the module to facilitate blunt-end cloning
815 into the pM9_4Compln vector from Madhuri et al. (Madhuri et al., 2019) containing the *bsr* gene
816 conferring resistance to blasticidin-S. Transformation of all three *alb3b* KO lines with the
817 pM9_4Compln vector containing the synthesized *ALB3b* module was performed as described
818 previously (Nymark et al., 2017). The algae were transferred to low-salt selection plates (25% (v/v)
819 natural seawater supplemented with f/2-Si, 1% (w/v) agar, 4 µg/mL blasticidin-S (Thermo Fisher
820 Scientific)) ~ 24 h after transformation. Transformed colonies appeared 3-4 weeks after transfer to
821 selection plates. Colonies that had regained the normal brown color were randomly picked from the
822 selection plates. PCR amplification of the *ALB3b* gene and subsequent sequencing were used to test
823 for the presence of the modified version of the *ALB3b* gene (and the absence of WT sequence).

28

824 Primers used for both PCR amplification and sequencing were PtAlb3b-G1F and PtAlb3b-G1R
825 (Supplementary Table S5). One complemented *alb3b* colony, resulting from each of the
826 transformations performed with the *alb3b* KO lines, was cultivated for pigment and protein
827 analyses, as described above.

828

829

830 **ACCESSION NUMBERS:**

831 Accession numbers for ALBINO protein sequences extracted from GenBank NCBI, the iMicrobe
832 database (Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP)) and from
833 the JGI genome portal are listed in Supplementary Table S4.

834

835 **SUPPLEMENTAL DATA:**

836 **Supplemental Figure 1:** Phylogenetic relationship between members of the ALBINO3 family.

837 **Supplemental Figure 2:** C-terminal domain of diatom ALB3a and ALB3b proteins.

838 **Supplemental Figure 3:** DNA sequences for the *ALB3b* WT gene and the inserts in the *alb3b* KO
839 lines.

840 **Supplemental Figure 4:** Allele-specific amplification of the Cas9 target site within the *ALB3b*
841 gene in WT and *alb3b* mutant strains.

842 **Supplemental Figure 5.** Relative expression levels of LHCF genes in *alb3b* lines compared to WT.

843 **Supplemental Figure 6.** Transmission electron micrographs of WT and *alb3b-14* mutant line cells.

844 **Supplementary Figure 7.** NPQ development over time in WT and *alb3b* lines.

845 **Supplemental Figure 8.** Western blot analysis of LHCF and LHCX proteins from WT and *alb3b*
846 mutant lines.

847 **Supplemental Figure 9.** Re-evaluation of pigment concentration per cell for WT and *alb3b*
848 mutants at LL.

869

879

880 **TABLES:**

881

882 **Table 1. Photosystem absorption cross-section and Chl *a* content per P700 in *alb3b* mutants**
883 **compared to WT cells.** Photosystem absorption cross-section was measured as rate of 533 nm
884 (Fx) or 670 nm (Chl *a*) photons absorbed by the functional thylakoid membranes. The actinic light
885 intensity was adjusted to I_{670} = 2.1 µmol photons m^{-2} s^{-1} and I_{533} = 12 µmol photons m^{-2} s^{-1}. Rates
886 of light absorption and utilization are given in photons per second with ±SD. P700 quantification
887 was measured from the light induced ΔA_{700} with 670 nm (Chl *a*) actinic illumination.

888

		WT	*alb3b-14*	*alb3b-16*	*alb3b-19*	Average *alb3b*	*alb3b*/WT %
PSI	(Fx) 533 nm	2.61 s^{-1} ± 0.40	1.10 ± 0.08	1.17 ± 0.10	1.09 ± 0.00	1.10 ± 0.06 s^{-1}	42 %
	(Chl) 670 nm	1.93 s^{-1} ± 0.11	1.43 ± 0.05	1.39 ± 0.15	1.39 ± 0.14	1.40 ± 0.01 s^{-1}	72.5%
PSII	(Fx) 533 nm	32.30 s^{-1} ± 0.7	13.92 ± 1.96	8.10 ± 1.16	8.58 ± 0.93	10.17 ± 3.24 s^{-1}	35 %
	(Chl) 670 nm	12.62 s^{-1} ± 2.69	7.21 ± 0.00	7.08 ± 0.90	6.71 ± 1.36	7.00 ± 0.26 s^{-1}	55 %
	Chl *a* /P700	663±9 % : 1	466 ± 11 %	414 ± 9 %	394 ± 11 %	425:1	64%

889

890

31

891 **Table 2. Photosynthesis and respiration properties of the WT and the *alb3b* KO lines.**
892 Parameters are calculated from the light-saturation curves of photosynthesis based on oxygen
893 evolution of WT and *alb3b* KO lines (Figure 7; LL: Figure 7E, ML: Figure7F). Data for *alb3b* are
894 presented as an average of the three independent *alb3b* KO (*alb3b-14*, *alb3b-16, alb3b-19*) lines
895 ±SD. A minimum of three biological replicates were measured for each independent line.

896

	LL		ML	
	WT	*alb3b*	**WT**	*alb3b*
Respiration (µmol O_2/ mol Chl/s)	30.0 ± 13.6	23.8 ± 1.7	23.5 ± 5.9	24.9 ± 3.2
P_{max} (µmol O_2/ mol Chl/s)	57.7 ± 11.5	63.2 ± 3.1	55.7 ± 4.9	71.8 ± 7.6
E_s(Saturation intensity, µmol photons $m^{-2} s^{-1}$)	96.5	250	170	> 400
Maximum light utilization coefficient (α)	0.35	0.32	0.29	0.25

897

898

899 **Table 3. Growth rates of WT and *alb3b* mutant lines acclimated to different light intensities.**
900 Maximum cell division per day were calculated from three biological replicates of WT and *alb3b*
901 KO lines acclimated to LL (35 µmol photons $m^{-2} s^{-1}$), ML (200 µmol photons $m^{-2} s^{-1}$) or HL (480
902 µmol photons $m^{-2} s^{-1}$). Values are presented with ±SD. Growth rate for the *alb3b-16* mutant in HL
903 was not calculated because of cell aggregation.

	WT	*alb3b-14*	*alb3b-16*	*alb3b-19*	*alb3b* average
LL	1.6 ± 0.23	0.4 ± 0.02	0.6 ± 0.02	0.6 ± 0.03	0.5 ± 0.09
ML	2.2 ± 0.03	1.1 ± 0.01	1.2 ± 0.03	1.4 ± 0.05	1.2 ± 0.13
HL	2.0 ± 0.05	0.8 ± 0.17	n/a	0.9 ± 0.25	0.8 ± 0.19

904

905

906

907

908   **FIGURE LEGENDS:**

909   **Figure 1: Presentation of intact and truncated ALB3b protein.** A) The area of the ALB3b

910   protein corresponding to the 20 bp target region for CRISPR/Cas9-based gene editing is located

911   toward the N-terminal part of the protein (blue highlighting) with the PAM site located at the

912   reverse DNA strand (green highlighting). CTP: Chloroplast targeting peptide; 60 kD IMP: 60 kD

913   Inner Membrane Protein domain; CTD: conserved C-terminal domain. B) Overview of amino acid

914   sequences resulting from CRISPR/Cas9 induced inserts in the three *alb3b* KO lines causing

915   premature stop codons and truncated ALB3b proteins. Color coding: Blue: WT target sequence;

916   Green: amino acid corresponding to PAM site; Red letters: Insert; *: Premature stop. C) Protein

917   alignment based on the C-terminal domain (CTD) of ALB3b proteins in diatoms.

918

919   **Figure 2. Color differences and spectral characteristics of WT and *alb3b* mutants.** A) Visual

920   representation of the *alb3b* phenotype compared to WT at low light (LL; 35 µmol photons $m^{-2} s^{-1}$;

921   left side) and ML (200 µmol photons $m^{-2} s^{-1}$; right side). For comparison and visualization of the

922   color differences, all cultures were adjusted to equal cell densities (3×10^7 cells/ml). B) Absorbance

923   spectra and C) *in vivo* fluorescence excitation spectra of cultures acclimated to ML. Isolated intact

924   thylakoid membranes were used for recording of the absorption spectra to avoid scattering.

925   Fluorescence emission was measured at 730 nm to ensure origin from the reaction center II Chl *a*.

926   Insets: Difference spectra between: the absorbance of WT and *alb3b* KO lines B) and excitation

927   energy transfer in the blue-green region of the *in vivo* fluorescence excitation spectra C). WT:

928   Presented as an average of three biological replicates; *alb3b*: Presented as an average of the three

929   *alb3b* KO lines 14, 16 and 19 with ±SD for all data points indicated by the grey area around the

930   graphs. Three biological replicates were measured for each line.

931

932   **Figure 3. 77 K fluorescence emission spectra of WT and *alb3b* KO samples acclimated to ML.**

933   Samples were excited at either A) 435 nm or B) 470 nm. The emission spectra were normalized at

934   their 710 nm maximum. Data for *alb3b* is an average of the three *alb3b* KO lines 14, 16 and 19

935   with ±SD for all data points indicated by the grey area around the graphs. Three biological

33

936   replicates were measured for each line including the WT.

937

938   **Figure 4. Western blot analysis of thylakoid membrane proteins from WT and *alb3b* mutant**
939   **lines acclimated to low light (LL; 35 μmol photons m^{-2} s^{-1}) or medium light (ML; 200 μmol**
940   **photons m^{-2} s^{-1}) conditions.** A) Abundance of LHC proteins belonging to the LHCF group were
941   evaluated using an antibody recognizing LHCF1-11, whereas the LHCX proteins were recognized
942   by anti-FCP6 (a LHCX family member of *C. meneghiniana*). A dilution series of the WT samples
943   were used to assess the level of LHC proteins in *alb3b* mutants compared to WT. B) Protein
944   expression of PSII and PSI core proteins were evaluated with antibodies against the D1 (PSII), D2
945   (PSII) and PsaC (PSI) core subunits. A dilution series of the *alb3b* samples were used to assess the
946   level of photosystem subunits in *alb3b* mutants compared to WT. An antibody recognizing the β-
947   subunit of ATP synthase (AtpB) were used as loading control on each of the individual blots. Lanes
948   marked with 100% contain 10 μg (20 μg for analysis of LHCX levels) of protein extracts.
949   Images have been cropped.

950

951   **Figure 5. Pigment concentrations per cell for WT and *alb3b* mutant lines as a function of ML**
952   **exposure time.** Cellular pigment concentrations of A) Chl *a*, B) Fx, C) Ddx and D) Dtx in WT and
953   *alb3b* mutant cells as a function of time following a shift from LL conditions (0 h; 35 μmol photons
954   m^{-2} s^{-1}) to ML conditions (200 μmol photons m^{-2} s^{-1}) for 0.5, 6, 24, 48 and 168 h. Results are
955   presented as a mean of three biological replicates with ±SD.

956

957   **Figure 6. De-epoxidation state index and NPQ capacity of WT and *alb3b* mutants.** A) De-
958   epoxidation state index (DES = Dtx/(Dtx + Ddx)) calculated from the HPLC pigment data from LL
959   acclimated (0 h) WT and *alb3b* cultures exposed to ML for 0.5, 6, 24, 48 and 168 h. B) Capacity for
960   NPQ calculated from rapid light curves derived from LL acclimated cells approx. two months after
961   isolation of mutated single cells and C) after being maintained in culture for one more year. NPQ =
962   $(F_{m'max}/F_{m'}) - 1$. $F_{m'max}$ replaces the commonly used F_m since $F_{m'}$ values frequently occur that are
963   higher than the F_m from dark-treated diatom samples (Serôdio et al., 2006). Results are presented as
964   a mean of three biological replicates with ±SD.

965

966   **Figure 7. Photo-physiological responses of WT and *alb3b* mutant lines.** *In vivo* Chl *a*

967 fluorescence kinetics (PAM) were used to estimate A) the maximum quantum yield of PSII (F_v/F_m),

968 B) the maximum light utilization coefficient (α), C) the maximum relative light-saturated electron

969 transport rate ($rETR_{max}$) and D) the light saturation index (E_k) in LL (0h) acclimated WT and *alb3b*

970 KO lines as a function of ML exposure time (0.5-168 h). Values are presented with ±SD bars.

971 Light-saturation curves of photosynthesis based on oxygen evolution were produced for E) LL

972 acclimated and F) ML acclimated WT and *alb3b* KO lines. The oxygen concentration was

973 normalized on a per-Chl basis. The results were fit with curves based on a polynomial regression

974 using R. All values are presented as an average of three biological replicates for each line and ±SD

975 for each value can be found in Supplementary Table S3.

976

977 **Figure 8. Culture color, LHCF protein level and pigment concentration in complemented**

978 ***alb3b* lines compared to WT.** A) WT and complemented *alb3b* KO lines (*alb3b-14C*, *alb3b-16C*,

979 *alb3b-19C*) were acclimated to LL and ML conditions. All cultures were concentrated and adjusted

980 to equal cell densities (3 x 10^7 cells/ml) for comparison. B) Western blot analysis of LHCF proteins

981 in WT and complemented *alb3b* mutant lines acclimated to LL and ML conditions. LHCF protein

982 levels were evaluated using LHCF1-11 antibody. An antibody recognizing the β-subunit of ATP

983 synthase was used as loading control. 10 µg of total protein from cell lysates was loaded onto

984 the gel. C) Cellular pigment concentrations of Chl *a* and Fx in LL conditions. Results are presented

985 as a mean of three biological replicates with ±SD bars.

986

987 **Figure 9. Proposed model of the role of diatom ALB3 insertases in insertion/assembly of**

988 **thylakoid membrane proteins.** LHC proteins are synthesized on ribosomes on the cERM,

989 transported through the four membranes surrounding the secondary plastid of diatoms, and guided

990 to ALB3b by an unknown protein complex before incorporation into the thylakoid membrane (left

991 side). Chloroplast-encoded proteins are suggested to be integrated by the co-translational cpSRP

992 pathway including cpSRP54, FTSY and ALB3ba (right side). cERM: chloroplast ER membrane;

993 PPM: periplastidal membrane; OEM: plastid outer envelope membrane; IEM: plastid inner

994 envelope membrane. CpSRP54: chloroplast signal recognition particle protein 54; CpFTSY:

995 chloroplast SRP receptor; ALB3: chloroplast SRP insertase Albino3.

996

997

998

REFERENCES

Ago H, Adachi H, Umena Y, Tashiro T, Kawakami K, Kamiya N, Tian L, Han G, Kuang T,
Liu Z, Wang F, Zou H, Enami I, Miyano M, Shen JR (2016) Novel features of
eukaryotic photosystem II revealed by its crystal structure analysis from a red alga. J Biol
Chem **291:** 5676-5687

Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, Putnam NH, Zhou S, Allen AE,
Apt KE, Bechner M, Brzezinski MA, Chaal BK, Chiovitti A, Davis AK, Demarest MS,
Detter JC, Glavina T, Goodstein D, Hadi MZ, Hellsten U, Hildebrand M, Jenkins BD,
Jurka J, Kapitonov VV, Kroger N, Lau WW, Lane TW, Larimer FW, Lippmeier JC,
Lucas S, Medina M, Montsant A, Obornik M, Parker MS, Palenik B, Pazour GJ,
Richardson PM, Rynearson TA, Saito MA, Schwartz DC, Thamatrakoln K, Valentin
K, Vardi A, Wilkerson FP, Rokhsar DS (2004) The genome of the diatom *Thalassiosira*
pseudonana: ecology, evolution, and metabolism. Science **306:** 79-86

Austin JR, 2nd, Staehelin LA (2011) Three-dimensional architecture of grana and stroma
thylakoids of higher plants as determined by electron tomography. Plant Physiol **155:** 1601-
1611

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. Proc Natl Acad Sci U S A **107:** 18214-18219

Baroli I, Melis A (1996) Photoinhibition and repair in *Dunaliella salina* acclimated to different
growth irradiances. Planta **198:** 640-646

Bellafiore S, Ferris P, Naver H, Gohre V, Rochaix JD (2002) Loss of Albino3 leads to the
specific depletion of the light-harvesting system. Plant Cell **14:** 2303-2314

Ben-Shem A, Frolow F, Nelson N (2003) Crystal structure of plant photosystem I. Nature **426:**
630-635

Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, Kuo A, Maheswari U, Martens C,
Maumus F, Otillar RP, Rayko E, Salamov A, Vandepoele K, Beszteri B, Gruber A,
Heijde M, Katinka M, Mock T, Valentin K, Verret F, Berges JA, Brownlee C, Cadoret
JP, Chiovitti A, Choi CJ, Coesel S, De Martino A, Detter JC, Durkin C, Falciatore A,
Fournet J, Haruta M, Huysman MJ, Jenkins BD, Jiroutova K, Jorgensen RE, Joubert
Y, Kaplan A, Kroger N, Kroth PG, La Roche J, Lindquist E, Lommer M, Martin-
Jezequel V, Lopez PJ, Lucas S, Mangogna M, McGinnis K, Medlin LK, Montsant A,
Oudot-Le Secq MP, Napoli C, Obornik M, Parker MS, Petit JL, Porcel BM, Poulsen
N, Robison M, Rychlewski L, Rynearson TA, Schmutz J, Shapiro H, Siaut M, Stanley
M, Sussman MR, Taylor AR, Vardi A, von Dassow P, Vyverman W, Willis A, Wyrwicz
LS, Rokhsar DS, Weissenbach J, Armbrust EV, Green BR, Van de Peer Y, Grigoriev
IV (2008) The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes.
Nature **456:** 239-244

Bozzola JJ, Russell LD (1999) Electron Microscopy: Principles and Techniques for Biologists, Ed
2. Jones and Bartlett, Boston

Bricaud A, Claustre H, Ras J, Oubelkheir K (2004) Natural variability of phytoplanktonic
absorption in oceanic waters: Influence of the size structure of algal populations. J Geophys
Res-Oceans **109**

Brown JW, Sorhannus U (2010) A molecular genetic timescale for the diversification of

1043   autotrophic stramenopiles (Ochrophyta): substantive underestimation of putative fossil ages.
1044   PLoS One **5**

1045 **Büchel C** (2014) Fucoxanthin-Chlorophyll-proteins and non-photochemical fluorescence
1046   quenching of diatoms. *In* B Demmig-Adams, G Garab, W Adams III, Govindjee, eds, Non-
1047   photochemical quenching and energy dissipation in plants, algae and cyanobacteria.
1048   Advances in photosynthesis and respiration (Including bioenergy and related processes), Vol
1049   40. Springer, Dordrecht

1050 **Büchel C** (2015) Evolution and function of light harvesting proteins. J Plant Physiol **172:** 62-75

1051 **Chandrasekar S, Shan SO** (2017) Anionic phospholipids and the Albino3 translocase activate
1052   signal recognition particle-receptor interaction during light-harvesting chlorophyll *a/b*-
1053   binding protein targeting. J Biol Chem **292:** 397-406

1054 **Cruz S, Goss R, Wilhelm C, Leegood R, Horton P, Jakob T** (2011) Impact of chlororespiration
1055   on non-photochemical quenching of chlorophyll fluorescence and on the regulation of the
1056   diadinoxanthin cycle in the diatom *Thalassiosira pseudonana*. J Exp Bot **62:** 509-519

1057 **Cruz S, Serôdio J** (2008) Relationship of rapid light curves of variable fluorescence to
1058   photoacclimation and non-photochemical quenching in a benthic diatom. Aquat Bot **88:**
1059   256-264

1060 **De Martino A, Bartual A, Willis A, Meichenin A, Villazan B, Maheswari U, Bowler C** (2011)
1061   Physiological and molecular evidence that environmental changes elicit morphological
1062   interconversion in the model diatom *Phaeodactylum tricornutum*. Protist **162:** 462-481

1063 **De Martino A, Meichenin A, Shi J, Pan KH, Bowler C** (2007) Genetic and phenotypic
1064   characterization of *Phaeodactylum tricornutum* (Bacillariophyceae) accessions. J Phycol **43:**
1065   992-1009

1066 **Durnford DG, Aebersold R, Green BR** (1996) The fucoxanthin-chlorophyll proteins from a
1067   chromophyte alga are part of a large multigene family: structural and evolutionary
1068   relationships to other light harvesting antennae. Mol Gen Genet **253:** 377-386

1069 **Dünschede B, Bals T, Funke S, Schunemann D** (2011) Interaction studies between the chloroplast
1070   signal recognition particle subunit cpSRP43 and the full-length translocase Alb3 reveal a
1071   membrane-embedded binding region in Alb3 protein. J Biol Chem **286:** 35187-35195

1072 **Falk S, Ravaud S, Koch J, Sinning I** (2010) The C terminus of the Alb3 membrane insertase
1073   recruits cpSRP43 to the thylakoid membrane. J Biol Chem **285:** 5954-5962

1074 **Falk S, Sinning I** (2010) The C terminus of Alb3 interacts with the chromodomains 2 and 3 of
1075   cpSRP43. J Biol Chem **285:** le25-26; author reply le26-28

1076 **Falkowski PG, Barber RT, Smetacek V** (1998) Biogeochemical controls and feedbacks on ocean
1077   primary production. Science **281:** 200-206

1078 **Falkowski PG, Raven JA** (2007) Aquatic photosynthesis, Ed Second edition. Princeton University
1079   Press, Princeton

1080 **Fawley MW** (1984) Effects of llight intensity and temperature interactions on growth
1081   characteristics of *Phaeodactylum tricornutum* (Bacillariophyceae). J Phycol **20:** 67-72

1082 **Formighieri C, Melis A** (2017) Heterologous synthesis of geranyllinalool, a diterpenol plant
1083   product, in the cyanobacterium Synechocystis. Appl Microbiol Biotechnol **101:** 2791-2800

1084 **Genty B, Briantais JM, Baker NR** (1989) The relationship between the quantum yield of
1085   photosynthetic electron-transport and quenching of chlorophyll fluorescence. Biochim
1086   Biophys Acta **990:** 87-92

1087 **Gerdes L, Bals T, Klostermann E, Karl M, Philippar K, Hunken M, Soll J, Schunemann D**
1088   (2006) A second thylakoid membrane-localized Alb3/OxaI/YidC homologue is involved in
1089   proper chloroplast biogenesis in *Arabidopsis thaliana*. J Biol Chem **281:** 16632-16642

Giovagnetti V, Ruban AV (2017) Detachment of the fucoxanthin chlorophyll *a/c* binding protein (FCP) antenna is not involved in the acclimative regulation of photoprotection in the pennate diatom *Phaeodactylum tricornutum*. Biochim Biophys Acta **1858:** 218-230

Goss R, Lepetit B (2015) Biodiversity of NPQ. J Plant Physiol **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. J Proteome Res **10:** 5338-5353

Gu J, Zhou Z, Li Z, Chen Y, Wang Z, Zhang H, Yang J (2017) Photosynthetic properties and potentials for improvement of photosynthesis in pale green leaf rice under high light conditions. Front Plant Sci **8:** 1082

Guenther JE, Melis A (1990) The physiological significance of photosystem II heterogeneity in chloroplasts. Photosynth Res **23:** 105-109

Guillard RR, Ryther JH (1962) Studies of marine planktonic diatoms. I. Cyclotella nana Hustedt, and Detonula confervacea (cleve) Gran. Can J Microbiol **8:** 229-239

Gundermann K, Büchel C (2014) Structure and functional heterogeneity of fucoxanthin-chlorophyll proteins in diatoms. *In* M Hohmann-Marriott, ed, The Structural Basis of Biological Energy Generation. Advances in Photosynthesis and Respiration (Including Bioenergy and Related Processes), Vol 39. Springer Dordrecht

Göhre V, Ossenbuhl F, Crevecoeur M, Eichacker LA, Rochaix JD (2006) One of two alb3 proteins is essential for the assembly of the photosystems and for cell survival in *Chlamydomonas*. Plant Cell **18:** 1454-1466

Havaux M, Dall'Osto L, Bassi R (2007) Zeaxanthin has enhanced antioxidant capacity with respect to all other xanthophylls in *Arabidopsis* leaves and functions independent of binding to PSII antennae. Plant Physiol **145:** 1506-1520

Hennon SW, Soman R, Zhu L, Dalbey RE (2015) YidC/Alb3/Oxa1 family of insertases. J Biol Chem **290:** 14866-14874

Herbstova M, Bina D, Kana R, Vacha F, Litvin R (2017) Red-light phenotype in a marine diatom involves a specialized oligomeric red-shifted antenna and altered cell morphology. Sci Rep **7:** 11976

Hiyama T, Ke B (1972) Difference spectra and extinction coefficients of P700. Biochim Biophys Acta **267:** 160-171

Ikeda Y, Komura M, Watanabe M, Minami C, Koike H, Itoh S, Kashino Y, Satoh K (2008) Photosystem I complexes associated with fucoxanthin-chlorophyll-binding proteins from a marine centric diatom, *Chaetoceros gracilis*. Biochim Biophys Acta **1777:** 351-361

Jeffrey SW, Humphrey GF (1975) New spectrophotometric equations for determining chlorophylls *a*, *b*, *c1* and *c2* in higher plants, algae and natural phytoplankton. Biochem Physiol Pfl **167:** 191-194

Juhas M, Buchel C (2012) Properties of photosystem I antenna protein complexes of the diatom *Cyclotella meneghiniana*. J Exp Bot **63:** 3673-3681

Juhas M, von Zadow A, Spexard M, Schmidt M, Kottke T, Buchel C (2014) A novel cryptochrome in the diatom *Phaeodactylum tricornutum* influences the regulation of light-harvesting protein levels. Febs J **281:** 2299-2311

Kalaji HM, Schansker G, Brestic M, Bussotti F, Calatayud A, Ferroni L, Goltsev V, Guidi L, Jajoo A, Li P, Losciale P, Mishra VK, Misra AN, Nebauer SG, Pancaldi S, Penella C, Pollastrini M, Suresh K, Tambussi E, Yanniccari M, Zivcak M, Cetner MD, Samborska IA, Stirbet A, Olsovska K, Kunderlikova K, Shelonzek H, Rusinowski S, Baba W (2017) Frequently asked questions about chlorophyll fluorescence, the sequel.

Photosynth Res **132:** 13-66

Kirst H, Formighieri C, Melis A (2014) Maximizing photosynthetic efficiency and culture productivity in cyanobacteria upon minimizing the phycobilisome light-harvesting antenna size. Bba-Bioenergetics **1837:** 1653-1664

Kirst H, Gabilly ST, Niyogi KK, Lemaux PG, Melis A (2017) Photosynthetic antenna engineering to improve crop yields. Planta **245:** 1009-1020

Kirst H, Garcia-Cerdan JG, Zurbriggen A, Melis A (2012) Assembly of the light-harvesting chlorophyll antenna in the green alga *Chlamydomonas reinhardtii* requires expression of the TLA2-CpFTSY gene. Plant Physiol **158:** 930-945

Kirst H, Garcia-Cerdan JG, Zurbriggen A, Ruehle T, Melis A (2012) Truncated photosystem chlorophyll antenna size in the green microalga *Chlamydomonas reinhardtii* upon deletion of the TLA3-CpSRP43 gene. Plant Physiol **160:** 2251-2260

Kirst H, Melis A (2014) The chloroplast signal recognition particle (CpSRP) pathway as a tool to minimize chlorophyll antenna size and maximize photosynthetic productivity. Biotechnol Adv **32:** 66-72

Kirst H, Shen Y, Vamvaka E, Betterle N, Xu D, Warek U, Strickland JA, Melis A (2018) Downregulation of the CpSRP43 gene expression confers a truncated light-harvesting antenna (TLA) and enhances biomass and leaf-to-stem ratio in *Nicotiana tabacum* canopies. Planta **248:** 139-154

Kugelmann M, Fausser A, Ossenbuhl F, Brennicke A (2013) Phenotypes of Alb3p and carotenoid synthesis mutants show similarities regarding light sensitivity, thylakoid structure and protein stability. Photosynthetica **51:** 45-54

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol **33:** 1870-1874

Lamb J, Forfang K, Hohmann-Marriott M (2015) A practical solution for 77 K fluorescence measurements based on LED excitation and CCD array detector. PLoS One **10:** e0132258

Lang M, Kroth PG (2001) Diatom fucoxanthin chlorophyll a/c-binding protein (FCP) and land plant light-harvesting proteins use a similar pathway for thylakoid membrane Insertion. J Biol Chem **276:** 7985-7991

Lavaud J, Goss R (2014) The peculiar features of non-photochemical fluorescence quenching in diatoms and brown algae. *In* B Demmig-Adams, G Garab, IW Adams, Govindjee, eds, Non-photochemical quenching and energy dissipation in plants, algae and cyanobacteria. Advances in photosynthesis and respiration (Including bioenergy and related processes), Vol 40. Springer, Dordrecht

Lavaud J, Lepetit B (2013) An explanation for the inter-species variability of the photoprotective non-photochemical chlorophyll fluorescence quenching in diatoms. Biochim Biophys Acta **1827:** 294-302

Lavaud J, Rousseau B, van Gorkom HJ, Etienne AL (2002) Influence of the diadinoxanthin pool size on photoprotection in the marine planktonic diatom *Phaeodactylum tricornutum*. Plant Physiol **129:** 1398-1406

Le SQ, Gascuel O (2008) An improved general amino acid replacement matrix. Mol Biol Evol **25:** 1307-1320

Lepetit B, Gelin G, Lepetit M, Sturm S, Vugrinec S, Rogato A, Kroth PG, Falciatore A, Lavaud J (2017) The diatom *Phaeodactylum tricornutum* adjusts nonphotochemical fluorescence quenching capacity in response to dynamic light via fine-tuned Lhcx and xanthophyll cycle pigment synthesis. New Phytol **214:** 205-218

Lepetit B, Goss R, Jakob T, Wilhelm C (2012) Molecular dynamics of the diatom thylakoid

1184   membrane under different light conditions. Photosynth Res **111:** 245-257

1185 **Lepetit B, Sturm S, Rogato A, Gruber A, Sachse M, Falciatore A, Kroth PG, Lavaud J** (2013)
1186   High light acclimation in the secondary plastids containing diatom *Phaeodactylum*
1187   *tricornutum* is triggered by the redox state of the plastoquinone pool. Plant Physiol **161:**
1188   853-865

1189 **Lepetit B, Volke D, Gilbert M, Wilhelm C, Goss R** (2010) Evidence for the existence of one
1190   antenna-associated, lipid-dissolved and two protein-bound pools of diadinoxanthin cycle
1191   pigments in diatoms. Plant Physiol **154:** 1905-1920

1192 **Lewis NE, Marty NJ, Kathir KM, Rajalingam D, Kight AD, Daily A, Kumar TK, Henry RL,**
1193   **Goforth RL** (2010) A dynamic cpSRP43-Albino3 interaction mediates translocase
1194   regulation of chloroplast signal recognition particle (cpSRP)-targeting components. J Biol
1195   Chem **285:** 34220-34230

1196 **Madhuri S, Serif M, Rio Bartulos C, Lepetit B, Kroth PG** (2019) A strategy to complement
1197   PtAUREO1a in TALEN knockout strains of *Phaeodactylum tricornutum*. Algal Res **39:**
1198   101469

1199 **Melis A** (1989) Spectroscopic methods in photosynthesis - Photosystem stoichiometry and
1200   chlorophyll antenna size. Philos T Roy Soc B **323:** 397-409

1201 **Melis A, Brown JS** (1980) Stoichiometry of system I and system II reaction centers and of
1202   plastoquinone in different photosynthetic membranes. P Natl Acad Sci-Biol **77:** 4712-4716

1203 **Miloslavina Y, Grouneva I, Lambrev PH, Lepetit B, Goss R, Wilhelm C, Holzwarth AR**
1204   (2009) Ultrafast fluorescence study on the location and mechanism of non-photochemical
1205   quenching in diatoms. Biochim Biophys Acta **1787:** 1189-1197

1206 **Mock T, Otillar RP, Strauss J, McMullan M, Paajanen P, Schmutz J, Salamov A, Sanges R,**
1207   **Toseland A, Ward BJ, Allen AE, Dupont CL, Frickenhaus S, Maumus F, Veluchamy**
1208   **A, Wu T, Barry KW, Falciatore A, Ferrante MI, Fortunato AE, Glöckner G, Gruber**
1209   **A, Hipkin R, Janech MG, Kroth PG, Leese F, Lindquist EA, Lyon BR, Martin J,**
1210   **Mayer C, Parker M, Quesneville H, Raymond JA, Uhlig C, Valas RE, Valentin KU,**
1211   **Worden AZ, Armbrust EV, Clark MD, Bowler C, Green BR, Moulton V, van**
1212   **Oosterhout C, Grigoriev IV** (2017) Evolutionary genomics of the cold-adapted diatom
1213   *Fragilariopsis cylindrus*. Nature **541:** 536

1214 **Moore M, Harrison MS, Peterson EC, Henry R** (2000) Chloroplast Oxa1p homolog albino3 is
1215   required for post-translational integration of the light harvesting chlorophyll-binding protein
1216   into thylakoid membranes. J Biol Chem **275:** 1529-1532

1217 **Nelson N, Yocum CF** (2006) Structure and function of photosystems I and II. Annu Rev Plant Biol
1218   **57:** 521-565

1219 **Nicholas KB, Nicholas HBJ, Deerfield DWI** (1997) GeneDoc: analysis and visualization of
1220   genetic variation. . EMBNEW.NEWS **4**

1221 **Nymark M, Sharma AK, Hafskjold MC, Sparstad T, Bones AM, Winge P** (2017)
1222   CRISPR/Cas9 Gene editing in the marine diatom *Phaeodactylum tricornutum*. Bio-protocol
1223   **7:** e2442

1224 **Nymark M, Sharma AK, Sparstad T, Bones AM, Winge P** (2016) A CRISPR/Cas9 system
1225   adapted for gene editing in marine algae. Sci Rep-Uk **6**

1226 **Nymark M, Valle KC, Brembu T, Hancke K, Winge P, Andresen K, Johnsen G, Bones AM**
1227   (2009) An integrated analysis of molecular acclimation to high light in the marine diatom
1228   *Phaeodactylum tricornutum*. PLoS ONE **4:** e7743

1229 **Nymark M, Valle KC, Hancke K, Winge P, Andresen K, Johnsen G, Bones AM, Brembu T**
1230   (2013) Molecular and photosynthetic responses to prolonged darkness and subsequent

acclimation to re-illumination in the diatom *Phaeodactylum tricornutum*. PLoS ONE **8:** e58722

Oey M, Ross IL, Stephens E, Steinbeck J, Wolf J, Radzun KA, Kugler J, Ringsmuth AK, Kruse O, Hankamer B (2013) RNAi knock-down of LHCBM1, 2 and 3 Increases photosynthetic H$_2$ production efficiency of the green alga *Chlamydomonas reinhardtii*. Plos One **8**

Ossenbühl F, Gohre V, Meurer J, Krieger-Liszkay A, Rochaix JD, Eichacker LA (2004) Efficient assembly of photosystem II in *Chlamydomonas reinhardtii* requires Alb3.1p, a homolog of *Arabidopsis* ALBINO3. Plant Cell **16:** 1790-1800

Oudot-Le Secq MP, Grimwood J, Shapiro H, Armbrust EV, Bowler C, Green BR (2007) Chloroplast genomes of the diatoms *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*: comparison with other plastid genomes of the red lineage. Mol Genet Genomics **277:** 427-439

Polle JE, Kanakagiri SD, Melis A (2003) Tla1, a DNA insertional transformant of the green alga *Chlamydomonas reinhardtii* with a truncated light-harvesting chlorophyll antenna size. Planta **217:** 49-59

Powles SB, Critchley C (1980) Effect of light intensity during growth on photoinhibition of intact attached bean leaflets. Plant Physiol **65:** 1181-1187

Premvardhan L, Bordes L, Beer A, Büchel C, Robert B (2009) Carotenoid structures and environments in trimeric and oligomeric fucoxanthin chlorophyll *a*/*c*$_2$ proteins from resonance Raman spectroscopy. J Phys Chem B **113:** 12565-12574

Premvardhan L, Robert B, Beer A, Büchel C (2010) Pigment organization in fucoxanthin chlorophyll *a*/*c*$_2$ proteins (FCP) based on resonance Raman spectroscopy and sequence analysis. Biochim Biophys Acta **1797:** 1647-1656

Ramakers C, Ruijter JM, Deprez RH, Moorman AF (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. Neurosci Lett **339:** 62-66

Rodriguez F, Chauton M, Johnsen G, Andresen K, Olsen LM, Zapata M (2006) Photoacclimation in phytoplankton: implications for biomass estimates, pigment functionality and chemotaxonomy. Mar Biol **148:** 963-971

Ruijter JM, Ramakers C, Hoogaars WM, Karlen Y, Bakker O, van den Hoff MJ, Moorman AF (2009) Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucleic Acids Res **37:** e45

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol **4:** 406-425

Sakshaug E, Bricaud A, Dandonneau Y, Falkowski PG, Kiefer DA, Legendre L, Morel A, Parslow J, Takahashi M (1997) Parameters of photosynthesis: definitions, theory and interpretation of results. J Plankton Res **19:** 1637-1670

Schuenemann D, Gupta S, Persello-Cartieaux F, Klimyuk VI, Jones JD, Nussaume L, Hoffman NE (1998) A novel signal recognition particle targets light-harvesting proteins to the thylakoid membranes. Proc Natl Acad Sci U S A **95:** 10312-10316

Serif M, Lepetit B, Weißert K, Kroth PG, Rio Bartulos C (2017) A fast and reliable strategy to generate TALEN-mediated gene knockouts in the diatom Phaeodactylum tricornutum. Algal Res 23: 186-195

Serôdio J, Vieira S, Cruz S, Coelho H (2006) Rapid light-response curves of chlorophyll fluorescence in microalgae: relationship to steady-state light curves and non-photochemical quenching in benthic diatom-dominated assemblages. Photosynth Res **90:** 29-43

Sundberg E, Slagter JG, Fridborg I, Cleary SP, Robinson C, Coupland G (1997) ALBINO3,

1278   an Arabidopsis nuclear gene essential for chloroplast differentiation, encodes a chloroplast
1279   protein that shows homology to proteins present in bacterial membranes and yeast
1280   mitochondria. Plant Cell **9:** 717-730
1281 **Taddei L, Chukhutsina VU, Lepetit B, Stella GR, Bassi R, van Amerongen H, Bouly JP,**
1282   **Jaubert M, Finazzi G, Falciatore A** (2018) Dynamic changes between two LHCX-related
1283   energy quenching sites control diatom photoacclimation. Plant Physiol **177:** 953-965
1284 **Taddei L, Stella GR, Rogato A, Bailleul B, Fortunato AE, Annunziata R, Sanges R, Thaler M,**
1285   **Lepetit B, Lavaud J, Jaubert M, Finazzi G, Bouly JP, Falciatore A** (2016) Multisignal
1286   control of expression of the LHCX protein family in the marine diatom *Phaeodactylum*
1287   *tricornutum*. J Exp Bot **67:** 3939-3951
1288 **Theis J, Schroda M** (2016) Revisiting the photosystem II repair cycle. Plant Signal Behav **11:**
1289   e1218587
1290 **Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG** (1997) The CLUSTAL_X
1291   windows interface: flexible strategies for multiple sequence alignment aided by quality
1292   analysis tools. Nucleic Acids Res **25:** 4876-4882
1293 **Träger C, Rosenblad MA, Ziehe D, Garcia-Petit C, Schrader L, Kock K, Richter CV,**
1294   **Klinkert B, Narberhaus F, Herrmann C, Hofmann E, Aronsson H, Schunemann D**
1295   (2012) Evolution from the prokaryotic to the higher plant chloroplast signal recognition
1296   particle: the signal recognition particle RNA is conserved in plastids of a wide range of
1297   photosynthetic organisms. Plant Cell **24:** 4819-4836
1298 **Valle KC, Nymark M, Aamot I, Hancke K, Winge P, Andresen K, Johnsen G, Brembu T,**
1299   **Bones AM** (2014) System responses to equal doses of photosynthetically usable radiation of
1300   blue, green, and red light in the marine diatom Phaeodactylum tricornutum. PLoS One **9:**
1301   e114211
1302 **Yamagishi A, Ikeda Y, Komura M, Koike H, Satoh K, Itoh S, Shibata Y** (2010) Shallow sink in
1303   an antenna pigment system of photosystem I of a marine centric diatom, *Chaetoceros*
1304   *gracilis*, revealed by ultrafast fluorescence spectroscopy at 17 K. J Phys Chem B **114:** 9031-
1305   9038
1306 **Wang W, Yu L-J, Xu C, Tomizaki T, Zhao S, Umena Y, Chen X, Qin X, Xin Y, Suga M, Han**
1307   **G, Kuang T, Shen J-R** (2019) Structural basis for blue-green light harvesting and energy
1308   dissipation in diatoms. Science **363:** eaav0365
1309
1310
1311
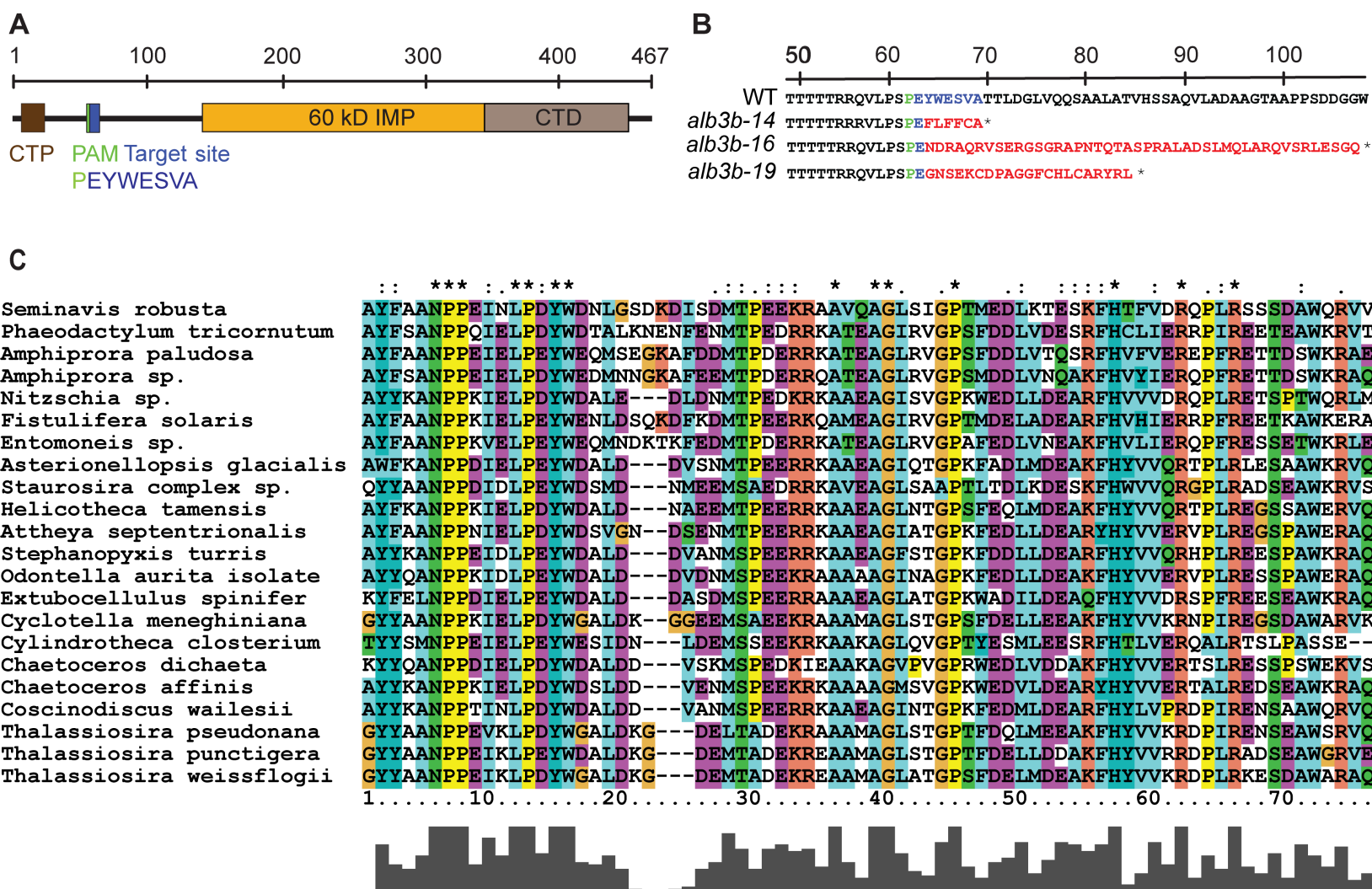
Figure 1: Presentation of intact and truncated ALB3b protein. A) The area of the ALB3b protein corresponding to the 20 bp target region for CRISPR/Cas9-based gene editing is located toward the N-terminal part of the protein (blue highlighting) with the PAM site located at the reverse DNA strand (green highlighting). CTP: Chloroplast targeting peptide; 60 kD IMP: 60 kD Inner Membrane Protein domain; CTD: conserved C-terminal domain. B) Overview of amino acid sequences resulting from CRISPR/Cas9 induced inserts in the three *alb3b* KO lines causing premature stop codons and truncated ALB3b proteins. Color coding: Blue: WT target sequence; Green: amino acid corresponding to PAM site; Red letters: Insert; *: Premature stop. C) Protein alignment based on the C-terminal domain (CTD) of ALB3b proteins in diatoms.
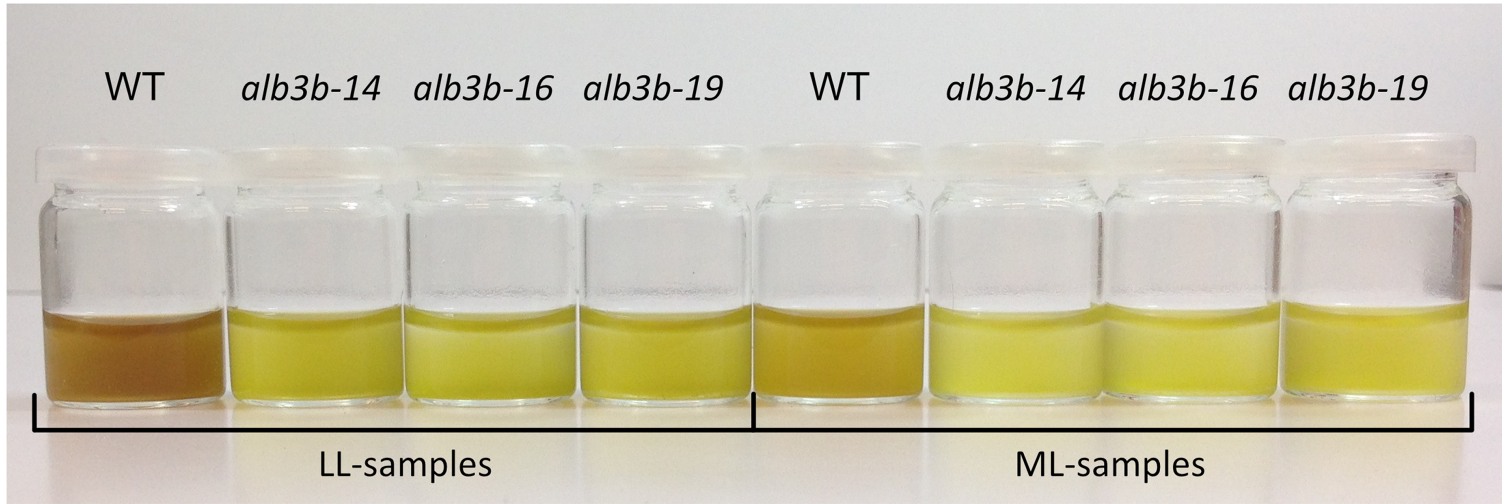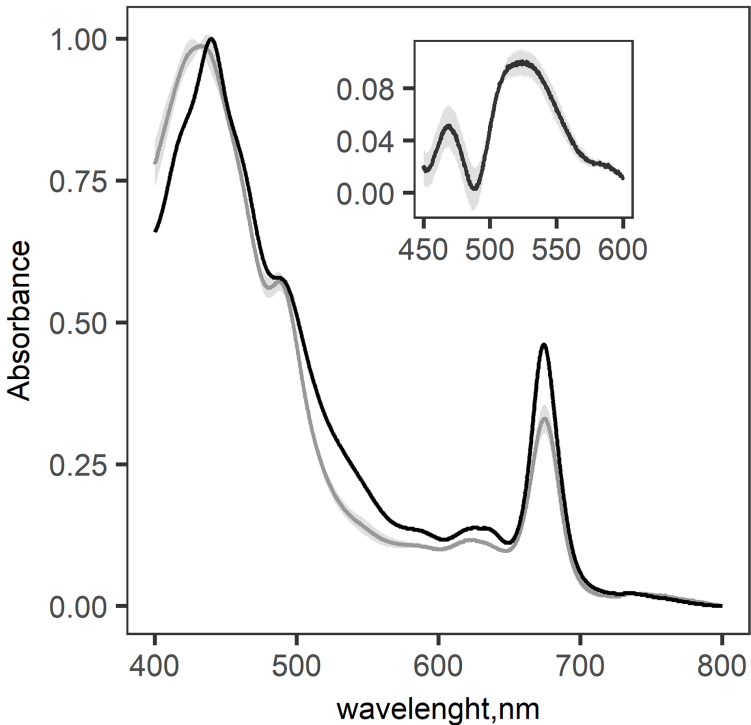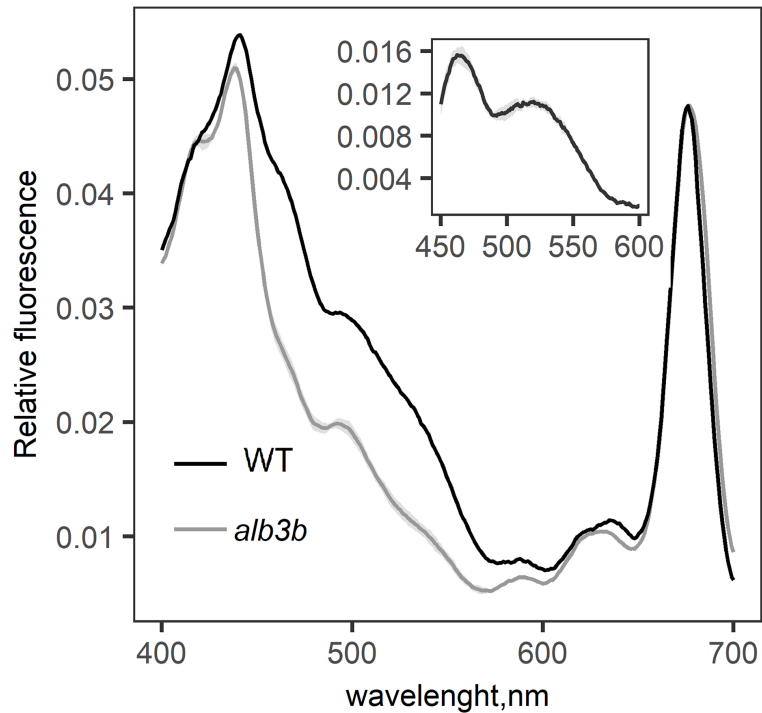
Figure 2. Color differences and spectral characteristics of WT and alb3b mutants. A) Visual representation of the *alb3b* phenotype compared to WT at low light (LL; 35 µmol photons m^{-2} s^{-1}; left side) and ML (200 µmol photons m^{-2} s^{-1}; right side). For comparison and visualization of the color differences, all cultures were adjusted to equal cell densities (3 x 10^7 cells/ml) B) Absorbance spectra and C) in vivo fluorescence excitation spectra of cultures acclimated to ML. Isolated intact thylakoid membranes were used for recording of the absorption spectra to avoid scattering. Fluorescence emission was measured at 730 nm to ensure origin from the reaction center II Chl *a*. Insets: Difference spectra between: the absorbance of WT and *alb3b* KO lines B), and excitation energy transfer in the blue-green region of the in vivo fluorescence excitation spectra C). WT: Presented as an average of three biological replicates; alb3b: Presented as an average of the three *alb3b* KO lines 14, 16 and 19 with ±SD for all data points indicated by the grey area around the graphs. Three biological replicates were measured for each line.

A
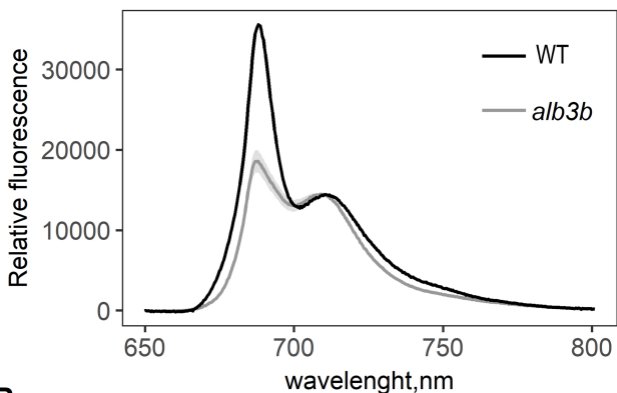
B
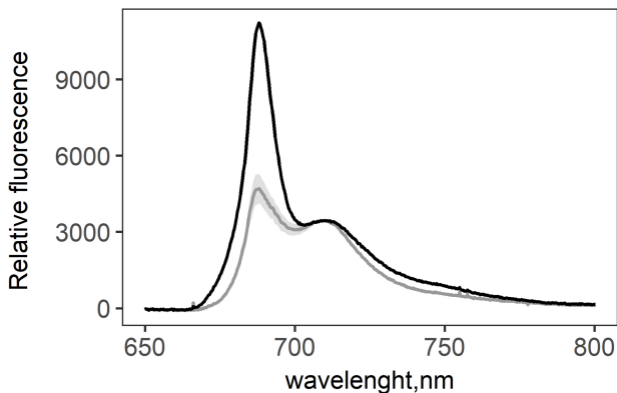
Figure 3. 77 K fluorescence emission spectra of WT and _alb3b_ KO samples acclimated to ML. Samples were excited at either 435 nm (A) or 470 nm (B). The emission spectra were normalized at their 710 nm maximum. Data for alb3b is an average of the three _alb3b_ KO lines 14, 16 and 19 with ±SD for all data points indicated by the grey area around the graphs. Three biological replicates were measured for each line including the WT.
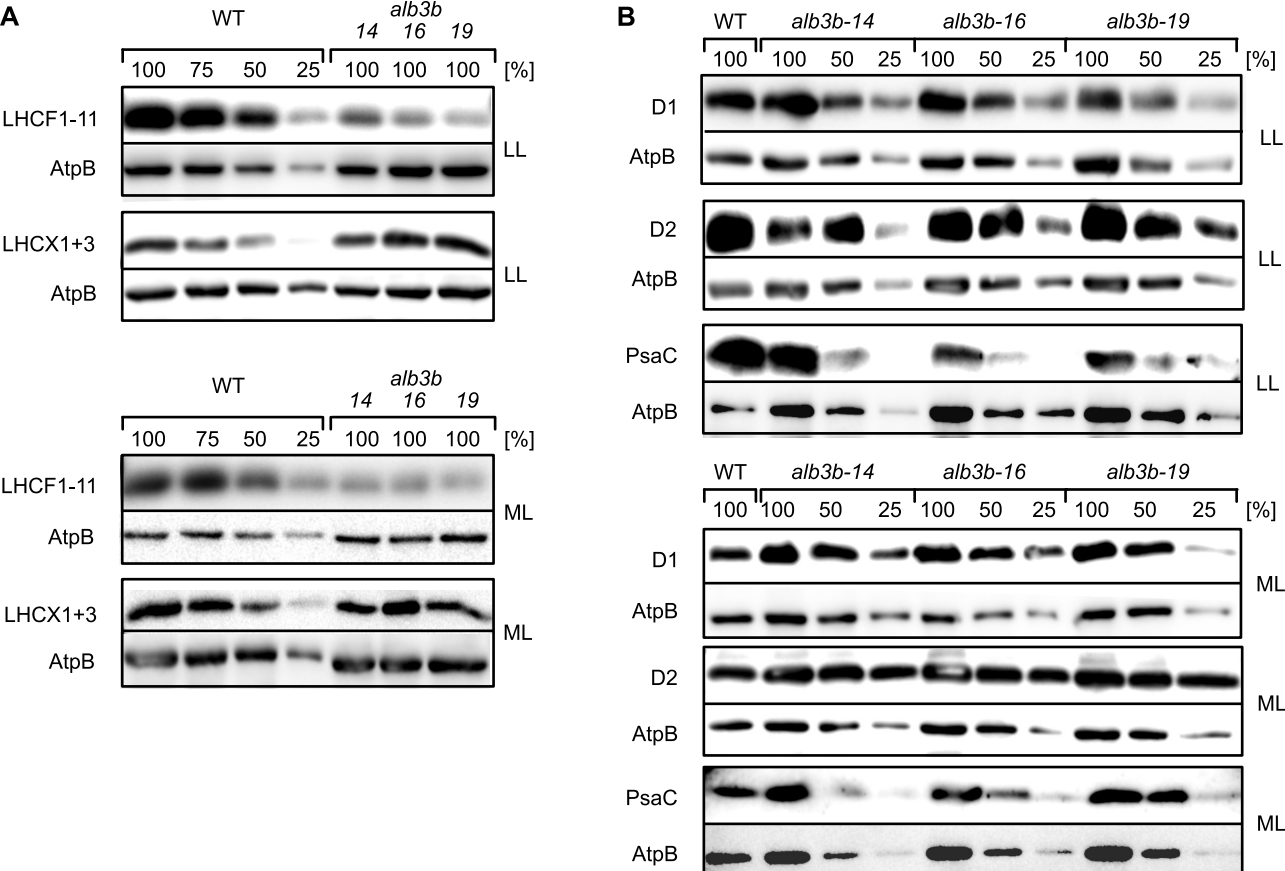
Figure 4. Western blot analysis of thylakoid membrane proteins from WT and *alb3b* mutant lines acclimated to low light (LL; 35 µmol photons m^{-2} s^{-1}) or medium light (ML; 200 µmol photons m^{-2} s^{-1}) conditions. A) Abundance of LHC proteins belonging to the LHCF group were evaluated using an antibody recognizing LHCF1-11, whereas the LHCX proteins were recognized by anti-FCP6 (a LHCX family member of *C. meneghiniana*). A dilution series of the WT samples were used to assess the level of LHC proteins in *alb3b* mutants compared to WT. B) Protein expression of PSII and PSI core proteins were evaluated with antibodies against the D1 (PSII), D2 (PSII) and PsaC (PSI) core subunits. A dilution series of the *alb3b* samples were used to assess the level of photosystem subunits in *alb3b* mutants compared to WT. An antibody recognizing the β-subunit of ATP synthase (AtpB) were used as loading control on each of the individual blots. Lanes marked with 100% contain 10 µg (20 µg for analysis of LHCX levels) of protein extracts. Images have been cropped.
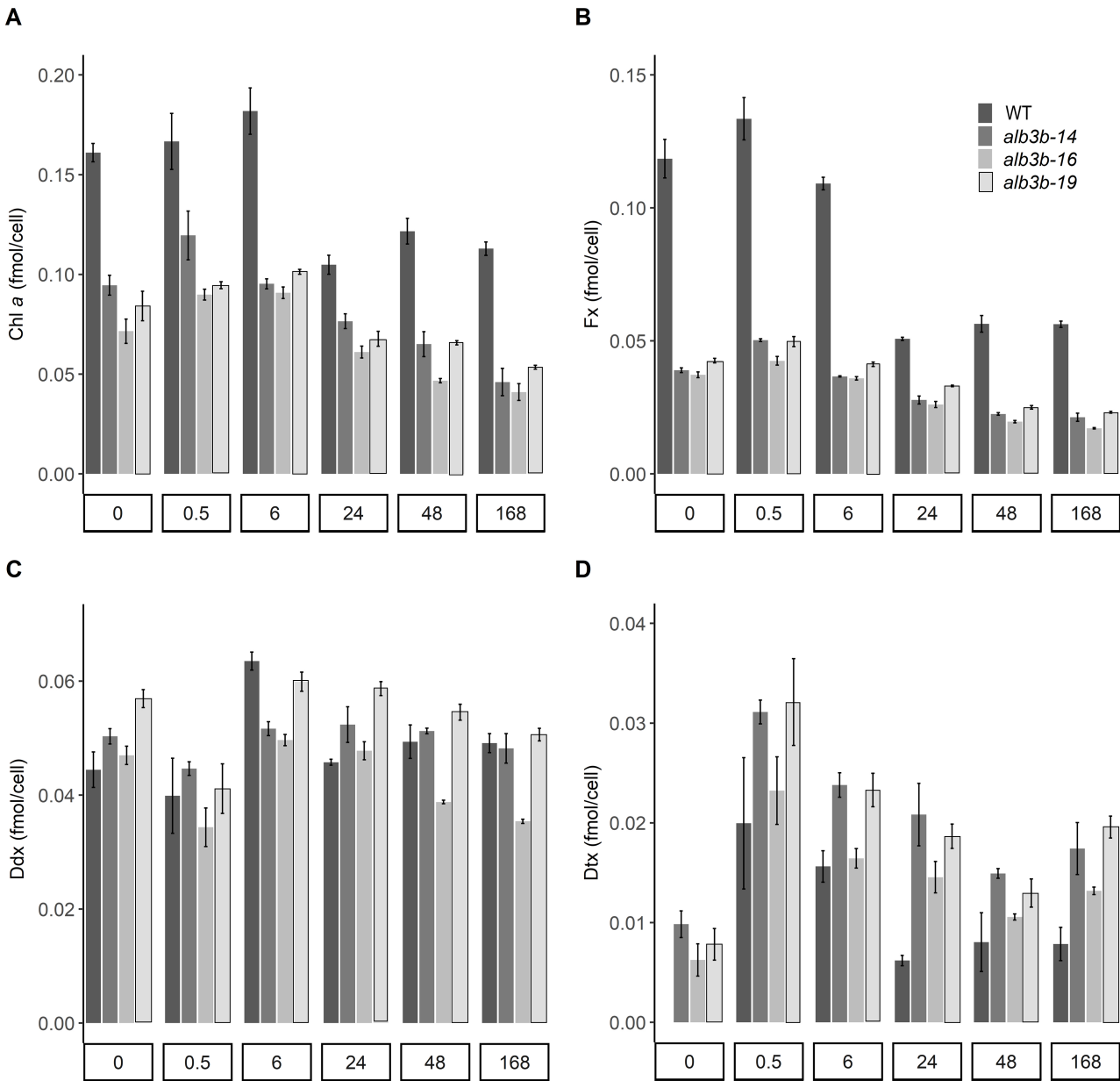
Figure 5. Pigment concentrations per cell for WT and *alb3b* mutant lines as a function of ML exposure time. Cellular pigment concentrations of A) Chl a, B) Fx, C) Ddx and D) Dtx in WT and *alb3b* mutant cells as a function of time following a shift from LL conditions (0 h; 35 µmol photons m^{-2} s^{-1}) to ML conditions (200 µmol photons m^{-2} s^{-1}) for 0.5, 6, 24, 48 and 168 h. Results are presented as a mean of three biological replicates with ±SD.
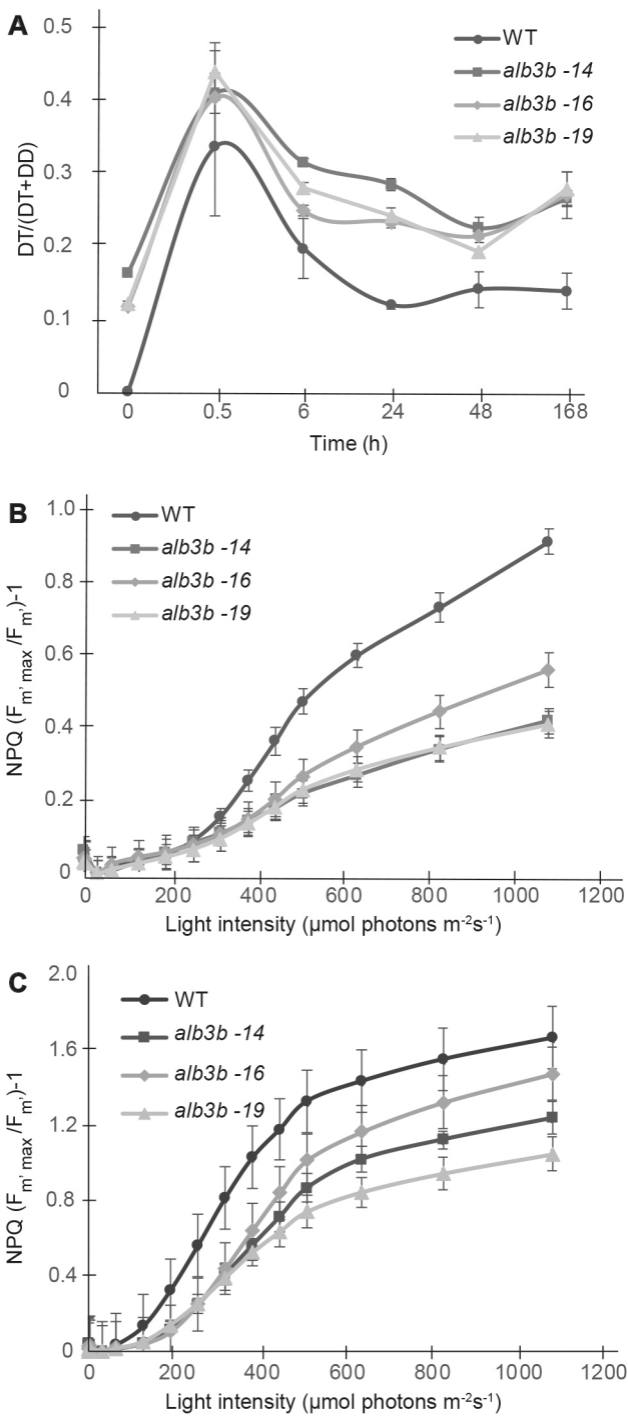
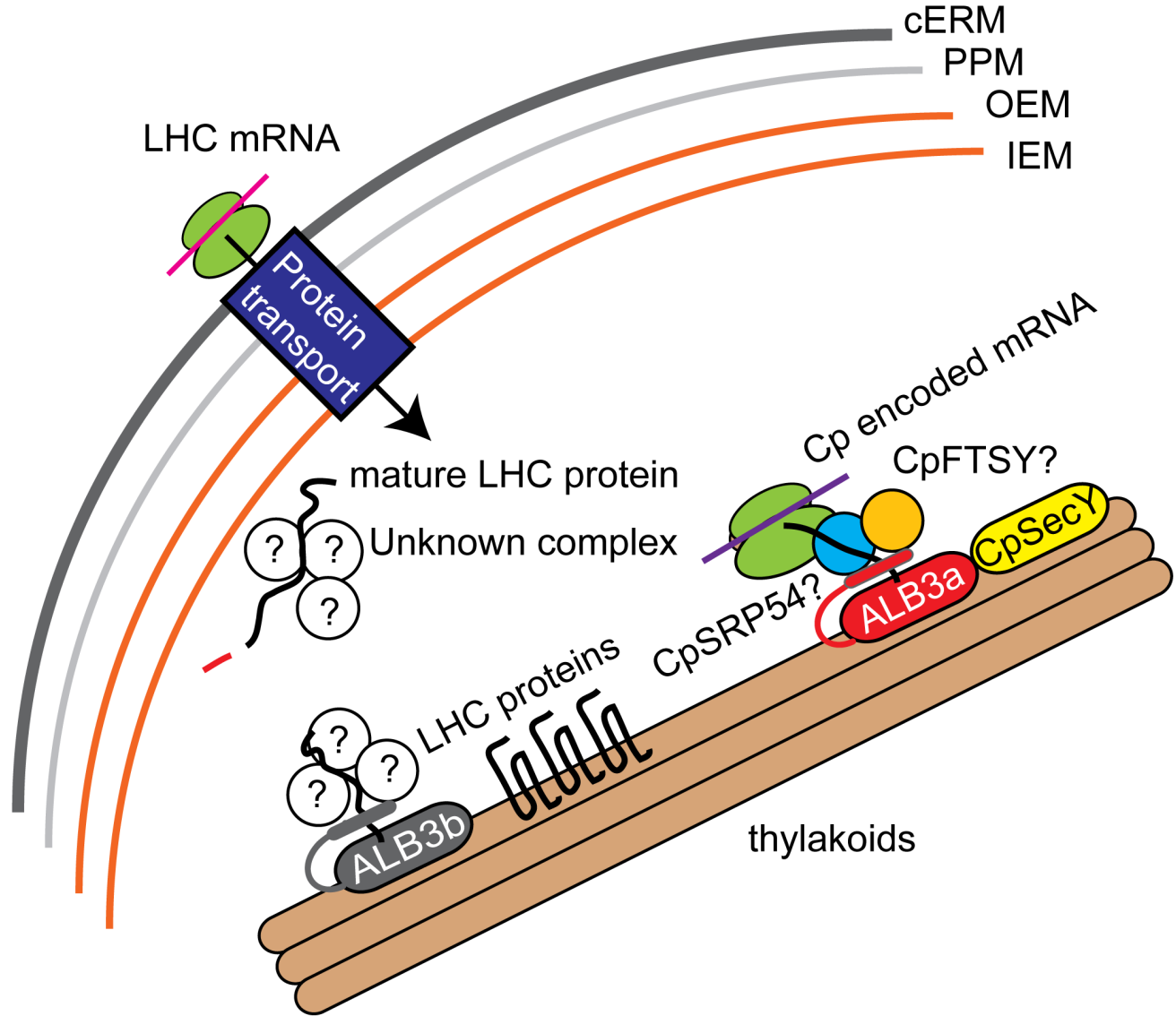**Figure 6. De-epoxidation state index and NPQ capacity of WT and *alb3b* mutants.** A) De-epoxidation state index (DES = Dtx/(Dtx + Ddx)) calculated from the HPLC pigment data from LL acclimated (0 h) WT and *alb3b* cultures exposed to ML for 0.5, 6, 24, 48 and 168 h. B) Capacity for NPQ calculated from rapid light curves derived from LL acclimated cells approx. two months after isolation of mutated single cells and C) after being maintained in culture for one more year. NPQ = $(F_{m'max}/F_{m'}) - 1$. $F_{m'max}$ replaces the commonly used F_m since $F_{m'}$ values frequently occur that are higher than the F_m from dark-treated diatom samples (Serodio et al., 2006). Results are presented as a mean of three biological replicates with ±SD.
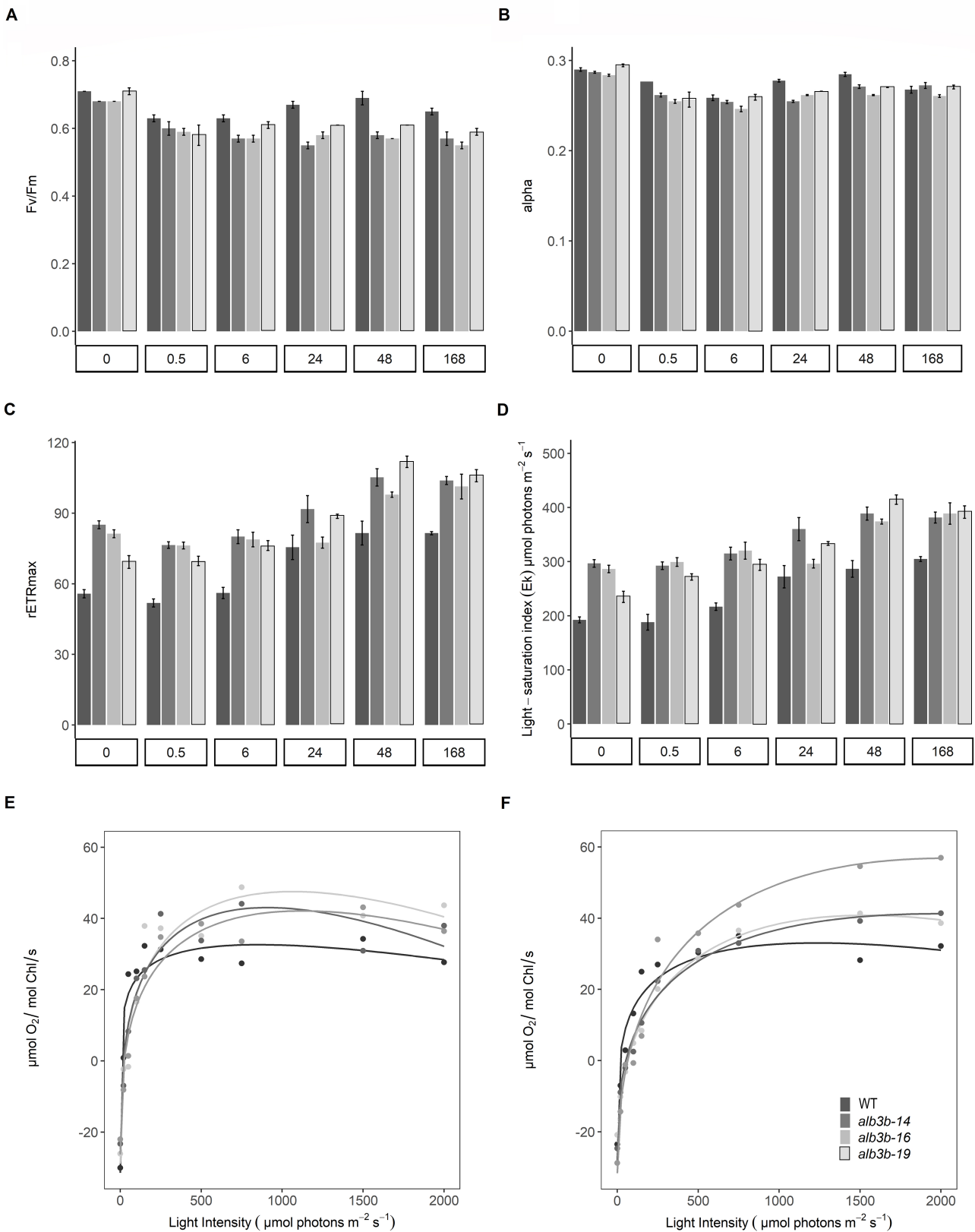
Figure 7. Photo-physiological responses of WT and alb3b mutant lines. In vivo Chl *a* fluorescence kinetics (PAM) were used to estimate A) the maximum quantum yield of PSII (Fv/Fm), B) the maximum light utilization coefficient (α), C) the maximum relative light-saturated electron transport rate (rETRmax) and D) the light saturation index (Ek) in LL (0h) acclimated WT and *alb3b* KO lines as a function of ML exposure time (0.5-168 h). Values are presented with ±SD bars. Light-saturation curves of photosynthesis based on oxygen evolution were produced for E) LL acclimated and F) ML acclimated WT and *alb3b* KO lines. The oxygen concentration was normalized on a per-Chl basis. The results were fit with curves based on a polynomial regression using R. All values are presented as an average of three biological replicates.

Figure 8. Culture color, LHCF protein level and pigment concentration in complemented *alb3b* lines compared to WT. A) WT and complemented *alb3b* KO lines (*alb3b*-14C, *alb3b*-16C, *alb3b*-19C) were acclimated to LL and ML conditions. All cultures were concentrated and adjusted to equal cell densities (3 x 10^7 cells/ml) for comparison. B) Western blot analysis of LHCF proteins in WT and complemented *alb3b* mutant lines acclimated to LL and ML conditions. LHCF protein levels were evaluated using LHCF1-11 antibody. An antibody recognizing the β-subunit of ATP synthase was used as loading control. 10 µg of total protein from cell lysates was loaded onto the gel. C) Cellular pigment concentrations of Chl *a* and Fx in LL conditions. Results are presented as a mean of three biological replicates with ±SD bars.

Figure 9. Proposed model of the role of diatom ALB3 insertases in insertion/assembly of thylakoid membrane proteins. LHC proteins are synthesized on ribosomes on the cERM, transported through the four membranes surrounding the secondary plastid of diatoms, and guided to ALB3b by an unknown protein complex before incorporation into the thylakoid membrane (left side). Chloroplast-encoded proteins are suggested to be integrated by the co-translational cpSRP pathway including cpSRP54, FTSY and ALB3ba (right side). cERM: chloroplast ER membrane; PPM: periplastidal membrane; OEM: plastid outer envelope membrane; IEM: plastid inner envelope membrane. CpSRP54: chloroplast signal recognition particle protein 54; CpFTSY: chloroplast SRP receptor; ALB3: chloroplast SRP insertase Albino3.