



HAL
open science

Allele-dependent expression and functionality of lipid enzyme phospholipid:diacylglycerol acyltransferase affect diatom carbon storage and growth

Yufang Pan, Wanting Zhang, Xiaofei Wang, Juliette Jouhet, Eric Maréchal,
Jin Liu, Xiao-Qin Xia, Hanhua Hu

► To cite this version:

Yufang Pan, Wanting Zhang, Xiaofei Wang, Juliette Jouhet, Eric Maréchal, et al.. Allele-dependent expression and functionality of lipid enzyme phospholipid:diacylglycerol acyltransferase affect diatom carbon storage and growth. *Plant Physiology*, 2024, 194, pp.1024-1040. 10.1093/plphys/kiad581 . hal-04297748

HAL Id: hal-04297748

<https://hal.science/hal-04297748v1>

Submitted on 21 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

1 **Short title:** ASE of *PDAT* regulates TAG accumulation

2

3 *Corresponding author: Hanhua Hu, hanhuahu@ihb.ac.cn

4

5 **Allele-dependent expression and functionality of lipid enzyme PDAT**
6 **affects diatom carbon storage and growth**

7

8 Yufang Pan,¹ Wanting Zhang,¹ Xiaofei Wang,² Juliette Jouhet,³ Eric Maréchal,³ Jin
9 Liu,² Xiao-Qin Xia,^{1,4} and Hanhua Hu^{1,4,*}

10 ¹Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, China

11 ²Laboratory for Algae Biotechnology and Innovation, College of Engineering, Peking
12 University, Beijing 100871, China

13 ³Laboratoire de Physiologie Cellulaire Végétale, Université Grenoble Alpes, CEA,
14 CNRS, INRA, IRIG-LPCV, 38054 Grenoble Cedex 9, France

15 ⁴University of Chinese Academy of Sciences, Beijing 100049, China

16

17 **One-sentence summary:** Allele-encoded phospholipid:diacylglycerol acyltransferase
18 responds specifically to nitrate stress and promotes triacylglycerol accumulation in the
19 model diatom *Phaeodactylum tricornutum*.

20

21 The author responsible for distribution of materials integral to the findings presented
22 in this article in accordance with the policy described in the Instructions for Authors
23 (<https://academic.oup.com/plphys/pages/General-Instructions>) is Hanhua Hu.

24

25

26

27 **ABSTRACT**

28 In the acyl-CoA-independent pathway of triacylglycerol (TAG) synthesis unique to
29 plants, fungi and algae, TAG formation is catalyzed by the enzyme
30 phospholipid:diacylglycerol acyltransferase (PDAT). The unique *PDAT* gene of the
31 model diatom *Phaeodactylum tricornutum* strain CCMP2561 boasts 47 single
32 nucleotide variants within protein coding regions of the alleles. To deepen our
33 understanding of TAG synthesis, we observed the allele-specific expression of *PDAT*
34 by the analysis of 87 published RNA-seq data and experimental validation. The
35 transcription of one of the two *PDAT* alleles, allele 2, could be specifically induced by
36 decreasing nitrogen concentrations. Overexpression of allele 2 in *P. tricornutum*
37 substantially enhanced the accumulation of TAG by 44~74% under nutrient stress;
38 however, overexpression of allele 1 resulted in little increase of TAG accumulation.
39 Interestingly, a more serious growth inhibition was observed in the *PDAT* allele 1
40 overexpression strains compared with allele 2 counterparts. Heterologous expression
41 in yeast (*Saccharomyces cerevisiae*) showed that enzymes encoded by *PDAT* allele 2
42 but not allele 1 had TAG biosynthetic activity, and 7 N-terminal and 3 C-terminal
43 amino acid variants between the two allele-encoded proteins substantially affected
44 enzymatic activity. *P. tricornutum* PDAT, localized in the innermost chloroplast
45 membrane, used monogalactosyldiacylglycerol and phosphatidylcholine as
46 acyl-donors as demonstrated by the increase of the two lipids in *PDAT* knockout lines,
47 which indicated a common origin in evolution with green algal PDATs. Our study
48 reveals unequal roles among allele-encoded PDATs in mediating carbon storage and
49 growth in response to nitrogen stress and suggests an unsuspected strategy towards
50 lipid and biomass improvement for biotechnological purposes.

51 **Keywords:** allele-specific expression, diatom, nitrogen, phospholipid:diacylglycerol
52 acyltransferase, triacylglycerol.

53 **Introduction**

54 Triacylglycerol (TAG) is present in all eukaryotes and some prokaryotes, and its
55 biosynthetic pathways, which appear to be conserved from bacteria to humans
56 (Athenstaedt and Daum, 2006), have received substantial attention in the last decade.
57 Two TAG biosynthetic pathways, the acyl CoA-dependent and acyl-CoA-independent
58 pathway, have been clarified. In the former pathway, commonly known as the
59 Kennedy pathway, diacylglycerol (DAG) and fatty acyl-CoAs are substrates of DAG
60 acyltransferase (DGAT) that catalyzes the terminal step. In the acyl-CoA-independent
61 pathway, which is unique in plants, algae and fungi, the formation of TAG is
62 catalyzed by phospholipid:diacylglycerol acyltransferase (PDAT) using membrane
63 glycerolipids as acyl donors (Dahlqvist et al., 2000).

64 PDAT belongs to the lecithin:cholesterol acyltransferase (LCAT) family and shares
65 structural similarity with human LCAT (Pan et al., 2015; Falarz et al., 2020). The
66 *PDAT* gene was identified from yeast (*Saccharomyces cerevisiae*) (Dahlqvist et al.,
67 2000), and subsequently *PDAT* was also identified in Arabidopsis (*Arabidopsis*
68 *thaliana*) (*AtPDAT1* and *AtPDAT2*) (Ståhl et al., 2004) together with several land
69 plants (Kim et al., 2011; van Erp et al., 2011; Pan et al., 2013; Aznar-Moreno and
70 Durrett, 2017; Marmon et al., 2017; Yuan et al., 2017; Xu et al., 2021),
71 *Chlamydomonas reinhardtii* (Yoon et al., 2012), *Myrmecia incise* (Liu et al., 2016c)
72 and *Nannochloropsis* (Fattore et al., 2022; Yang et al., 2022). Usually, land plants
73 contain multiple *PDAT* copies in contrast to the single copy in yeast and microalgae
74 (Pan et al., 2015). PDAT is indispensable to TAG biosynthesis at the exponential
75 growth stage in yeast (Oelkers et al., 2000) and *C. reinhardtii* (Yoon et al., 2012),
76 while no change of fatty acid (FA) content or composition was observed in the seeds
77 of *AtPDAT1* overexpression and knockout lines of Arabidopsis (Ståhl et al., 2004;
78 Mhaske et al., 2005). Nevertheless, silencing of *AtPDAT1* under a *dgat1* knockout
79 background resulted in a severe decrease in seed TAG level, and the double mutation
80 led to the lack of visible oil bodies (Zhang et al., 2009), suggesting that PDAT1 and

81 DGAT1 have overlapping functions in TAG biosynthesis in developing seeds and
82 pollen. Furthermore, AtPDAT1 is essential for TAG biosynthesis and maintaining
83 lipid homeostasis in growing tissues in Arabidopsis (Fan et al., 2013, 2014). In
84 contrast, *Nannochloropsis oceanica* PDAT (NoPDAT) acts under nutrient stress, and an
85 impairment of NoPDAT does not trigger any compensatory mechanism via the
86 acyl-CoA-dependent pathway (Yang et al., 2022).

87 Diatoms are abundant unicellular algae found in most aquatic habitats, and
88 responsible for about 20% of the primary production on Earth (Field et al., 1998;
89 Falkowski et al., 2004). *Phaeodactylum tricornutum* is one of the model diatoms with
90 the capacity to accumulate eicosapentaenoic acid, fucoxanthin, and neutral lipids
91 (mostly TAGs), and thus is perceived as a microalgal cell factory for high-value
92 products and oil production (Butler et al., 2020). Six DGAT encoding genes were
93 identified in *P. tricornutum* (Guihéneuf et al., 2011; Cui et al., 2013; Gong et
94 al., 2013), and several of them have been overexpressed to improve TAG
95 accumulation (Niu et al., 2013; Cui et al., 2018; Dinamarca et al., 2017; Haslamet
96 al., 2020; Zhang et al., 2021). In particular, overexpression of *PtDGAT1* produced
97 more than double the amount of TAG and total lipids (Zhang et al., 2021). However, *P.*
98 *tricornutum* PDAT (PtPDAT), a single copy gene annotated in the genome (Bowler et
99 al., 2008), remains yet to be studied.

100 Gene expression is a complex trait that is regulated by genetic and epigenetic
101 variation and also by developmental and environmental cues (Pastinen, 2010).
102 Although most genes are expressed equally from both alleles, some genes are
103 differentially expressed. The preferential expression of one of the alleles is termed as
104 allele-specific expression (ASE). ASE has been found across human tissues (Aguet et
105 al., 2017) and is regulated by genetic and environmental effects (Buil et al., 2014).
106 ASE has also been demonstrated in land plants such as Arabidopsis (Zhang and
107 Borevitz, 2009; Todesco et al., 2010) and rice (*Oryza sativa*) (He et al., 2010;
108 Chodavarapu et al., 2012; Shao et al., 2019). In diatoms, ASE was reported in the polar

109 *Fragilariopsis cylindrus* and the allelic imbalance in gene expression might play an
110 important role in the adaptation to new environments (Mock et al., 2017). Nomaguchi
111 et al. (2018) revealed that there were subgenomic preferences in about 61% of
112 homoeologous gene pairs in the allodiploid diatom *Fistulifera solaris*. ASE in *P.*
113 *tricornutum* has also been analyzed using whole genome RNA-seq data and
114 polymorphism calling, which indicated that 22% of the genes showed moderate bias
115 in allelic expression and 1% showed nearly complete monoallelic expression (Hoguin
116 et al., 2021). *PDAT* was considered to be allele-specifically expressed in both *F.*
117 *solaris* and *P. tricornutum* (Nomaguchi et al., 2018; Hoguin et al., 2021).

118 In this study, the ASE of *PtPDAT* in response to various environmental stimuli was
119 analyzed using previously published RNA-seq data and was verified by experiments.
120 Sequences of the two *PtPDAT* alleles were compared, and *in vivo* and heterologous *in*
121 *vivo* functions of their encoded proteins were studied. We reveal the critical and
122 unequal role of PDATs encoded by alleles in mediating TAG synthesis. Furthermore,
123 ASE might be a key strategy in response to nutrient stress in the diatom.

124

125 **Results**

126 **Sequence and expression analysis of the *PDAT* alleles in *P. tricornutum***

127 PDAT belongs to the lecithin:cholesterol acyltransferase (LCAT) family, and *P.*
128 *tricornutum* possesses two LCATs (Phatr3_J49702 and Phatr3_J8860) (Huang et al.
129 2019). Only Phatr3_J8860 harbors domains conserved in PDAT polypeptides (Falarz
130 et al., 2020), and expressed sequence tag (EST) libraries (Maheswari et al., 2010)
131 show some single nucleotide variants (SNVs) within the coding region of *PtPDAT*
132 gene. A large number of single nucleotide polymorphisms (SNPs) were found in ten
133 sequenced accessions of *P. tricornutum* (Pt1-Pt10 strains). The 10 accessions were
134 subdivided into four genetic clades (A–D) in our previous study (Rastogi et al, 2020),
135 and the coding regions of *PtPDAT* were assembled using whole genome resequencing
136 data of Pt1- Pt10. In Genotype A (Pt1, Pt2, Pt3 and Pt9), the coding region of *PDAT*

137 alleles (allele 1 and allele 2) boasted 47 SNVs (Supplemental Figure S1), resulting in
138 17 amino acid differences between the encoded proteins (Supplemental Figure S2).
139 SNVs were also found in *PDAT* allelic sequences of Genotype D (Pt6, Pt7 and Pt8).
140 However, there is no difference between the *PDAT* allelic sequences of Genotype B
141 (Pt4) and Genotype C (Pt5 and Pt10). *PDAT* was allele-specifically expressed in strain
142 CCMP2561 (Pt1) (Hoguin et al., 2021), the most commonly used strain in
143 laboratories, so this strain was used to analyze the expression of *PDAT* alleles and
144 published RNA sequencing data from five bioprojects from which 87 samples were
145 selected (Supplemental Dataset S1).

146 *PtPDAT* transcript level was up-regulated in response to nitrogen-free (PT1_-N)
147 treatment for 48 h (Levitan et al., 2015). *PDAT* allele 2-specific transcript level was
148 higher under nitrogen-free condition compared with nitrogen-replete (PT1_+N)
149 condition, however, *PDAT* allele 1-specific transcript decreased under nitrogen-free
150 condition (PT1_-N) (Figure 1A). These results indicated that the up-regulation of
151 *PtPDAT* transcript level can be attributed to the increased expression of allele 2.
152 Consistently, allele 2-specific transcript was also higher under no nitrate treatment for
153 4, 8 and 20 h (Matthijs et al., 2016) compared with that of the respective control, and
154 incubation in N-free media for 2 h (Mccarthy et al., 2017) or 18 h (Smith et al., 2019)
155 resulted in higher allele 2-specific transcript relative to the incubation in NH_4^+ media
156 (Figure 1A; Supplemental Figure S3). Furthermore, allele 2-specific transcript
157 increased gradually from 4 to 20 h during the cultivation (Figure 1A), and tended to
158 increase from 18 to 66 h but then remained relatively stable from 90 to 162 h ($P >$
159 0.05). In addition, neither *PtPDAT* nor allele 2-specific transcript was up-regulated by
160 early ('36 h' and '4 d') phosphorus depletion (Matthijs et al., 2016; Cruz de Carvalho
161 et al., 2016), while late ('8 d') phosphorus depletion (Cruz de Carvalho et al., 2016)
162 gave rise to increased *PtPDAT* transcript owing to the increased transcripts of both
163 alleles (Figure 1B). Growth with different nitrogen sources (NH_4^+ , NO_2^- or NO_3^-)
164 (Smith et al., 2019) resulted in comparable transcript levels of *PtPDAT*, and a slight

165 increase of allele 2-specific transcript was observed under nitrate and nitrite
166 conditions compared with ammonium condition (Supplemental Figure S3). Dark
167 treatment induced the simultaneous increase of transcript level in both *PDAT* alleles,
168 leading to the dramatic increase of *PtPDAT* transcript. No obvious difference of
169 *PtPDAT* transcript was found in samples treated with nocodazole (Matthijs et al.,
170 2016), and transcripts of alleles 1 and 2 showed no difference under different CO₂
171 concentrations (Supplemental Figure S3). Although the iron limitation decreased
172 allele 1 transcription, Fe- or vitamin B12-limitation had no significant ($P > 0.05$) effect
173 on the mRNA levels of allele 2 (Allen et al., 2008) (Supplemental Figure S3). These
174 results indicated that allele 2 transcript could be specifically induced by nitrogen
175 stress. According to our immunoblotting assay, PDAT protein level in the batch
176 culture soared from 12 to 36 h (nitrate in the media was nearly exhausted at 36 h, Ge
177 et al., 2004), and then a modest increase was observed from 36 to 84 h followed by a
178 decrease at 108 and 156 h (Figure 1C).

179

180 **Overexpression of *PDAT* allele 2 significantly promotes TAG accumulation**

181 To understand the function of PDAT, two *PDAT* alleles were separately overexpressed
182 in *P. tricornutum* cells. A substantial increase (1.6 to 12.4-fold) of *PtPDAT*
183 transcription occurred in all the four overexpression (OE) lines on day 5, detected by
184 RT-qPCR using universal primers, though some OE strains showed no or very little
185 up-regulation compared with wild-type (WT) on day 3 and 4 (Figure 2A).
186 Subsequently, allele-specific primers were used to detect *PDAT* expression, and *PDAT*
187 showed up-regulation (4.4~8.3-fold on day 3 and over 38-fold on day 5) only in allele
188 1 OE lines with allele 1-specific primers, indicating that the primer pair was highly
189 specific (Figure 2A). When detected with allele 2-specific primers instead of universal
190 primers, the relative level of *PDAT* was substantially higher in allele 2 OE lines not
191 only on day 5 but also on day 3, but the increase fold of the relative level lowered
192 from 8.1~12.4-fold to 2.3~3.3-fold on day 5 in allele 1 OE lines. Thus, the specificity

193 of allele 2-specific primer pair was acceptable but that of allele 1-specific primer pair
194 was more admirable.

195 The relative expression results using universal primers and highly specific allele 1
196 primers respectively were used to evaluate the background expression of *PDAT* alleles
197 (transgenic-derived *PDAT* transcripts were excluded). After calculation using the
198 formula provided in the Methods, the expression of allele 2 was higher than that of
199 allele 1 especially on day 5 (about 3 times), which is consistent with the RNA
200 sequencing data that allele 2-specific transcript was higher under nitrogen limitation.
201 In addition, RT-qPCR and western blotting results both indicated that the relative
202 expression levels of *PDAT* in allele 1 OE lines were higher than those in allele 2 OE
203 lines, and that *PDAT* relative expression was up-regulated under nitrogen limitation
204 (day 4~5) in all lines (Figure 2, A and B).

205 Overexpression of allele 1 resulted in remarkable inhibition of growth on day 10
206 (decreased to 77-79% of WT) while no substantial difference of growth was observed
207 during exponential phase with 500 μ M nitrate (LN+NP) (Figure 3A), and it had
208 almost no effect on both nitrate utilization and neutral lipid content (Figure 3, B and
209 C). Compared with WT, allele 2 OE lines exhibited lower cell densities throughout the
210 whole growth phase (decreased by 11–16% on day 10) (Figure 3D). Although there
211 was no difference in nitrate utilization between allele 2 OE lines and WT (Figure 3E),
212 neutral lipid content detected by Nile red was higher (increased by 43-59% and 44–54%
213 on day 8 and day 10, respectively) in allele 2 OE lines during the whole cultivation
214 period (Figure 3F). Increased TAG contents were also demonstrated by the thin-layer
215 chromatography (TLC) analysis in allele 2 OE lines on day 8 and 10 (Supplemental
216 Figure S4a).

217 To observe how allele 2 functioned in response to different nutrient concentrations,
218 allele 2 OE lines were cultivated in the media with high nitrate and high phosphate
219 (HN+HP), high nitrate and normal phosphate (HN+NP), and low nitrate and high
220 phosphate (LN+HP) respectively. Under HN+HP condition, no differences were

221 observed in growth (Figure 4A), and nitrate and phosphate utilization (Figure 4B)
222 between allele 2 OE lines and WT, meanwhile nearly no neutral lipid was detected in
223 the OE lines and WT (Figure 4C). Overexpression of *PDAT* allele 2 inhibited the
224 growth under HN+NP (Figure 4D) condition after day 8 (decreased by 11-14% on day
225 14, $P < 0.05$) but did not affect the nitrate and phosphate utilization (Figure 4E).
226 However, allele 2 OE lines possessed higher neutral lipid content (increased by 53-74%
227 on day 14) (Figure 4F). Growth of allele 2 OE lines tended to decrease slightly
228 (decreased by 10-19% on day 10) compared with WT under LN+HP (Figure 4G)
229 condition, and nitrate and phosphate utilization was not different between them
230 (Figure 4H) while neutral lipid content was higher in the OE lines, especially on day
231 10 (increased by 58-63%) (Figure 4I). These results showed that the limitation of
232 either nitrate or phosphate in the media could induce neutral lipid accumulation, while
233 nitrate limitation in the media enhanced neutral lipid accumulation to a higher extent
234 than phosphate limitation (Figure 4, F and I). It is suggested that the combination of
235 allele 2 overexpression and nitrate starvation gave rise to much higher TAG
236 accumulation. In contrast, overexpression of *PDAT* allele 1 gave rise to a slightly
237 higher cell density on day 14 (HN+HP, HN+NP) or 10 (LN+HP), and had almost no
238 effect on the neutral lipid content (Supplemental Figure S5)

239

240 **The enzyme encoded by *PDAT* allele 2 rather than allele 1 shows TAG**
241 **biosynthetic activity in yeast**

242 As the overexpression of *PDAT* allele 2 instead of allele 1 enhanced TAG
243 accumulation *in vivo* (Figure 3; Supplemental Figure S4a), the coding sequences
244 (CDS) of the two alleles and four reconstructed *PDAT*s (*PDAT3-6*) were expressed in
245 the TAG-deficient yeast H1246 to verify the activity of proteins encoded by the 6
246 *PDAT*s. Reconstructed *PDAT3-6* were generated by replacing the fragment in *PDAT*
247 allele 1 with the corresponding segment from allele 2 (Figure 5A; Supplemental
248 Figure S2). Obviously, yeast cells carrying the protein encoded by allele 2

249 accumulated the TAG-filled lipid droplets (LDs, indicated by green fluorescence). In
250 contrast, similar to EV (negative control), yeast cells carrying the protein encoded by
251 allele 1 failed to accumulate detectable LDs (Figure 5B). LDs in yeast cells
252 expressing the four reconstructed *PDATs* were less than those in H1246 cells
253 expressing *PDAT* allele 2, and as for the 4 reconstructed *PDATs*, yeast cells carrying
254 *PDAT6* and *PDAT4* exhibited the strongest and weakest green fluorescence signal
255 respectively (Figure 5B). TLC (Figure 5C) and GC–MS analysis (Figure 5D) further
256 confirmed that considerably high amount of TAG was accumulated in yeast cells
257 carrying the protein encoded by allele 2, *PDAT6* or *PDAT3*, while yeasts expressing
258 *PDAT4* or *PDAT5* synthesized only a trace amount of TAG. TAG was almost not
259 detectable in yeast cells carrying the protein encoded by allele 1. The relative
260 abundance of fatty acids in TAG was similar in yeast cells expressing the three
261 functional *PDATs* (Figure 5E).

262 Although *PDAT* allele 1 encoding protein showed little TAG biosynthetic activity,
263 replacing the N-terminal (*PDAT3*, including 7 amino acid variants) or C-terminal
264 peptides (*PDAT6*, including 3 amino acid variants) of allele 1 encoding protein with
265 the corresponding allele 2 encoding peptides recovered the enzyme activity. It is
266 suggested that these 10 amino acid variants had a positive impact on the protein
267 activity. Structure predictions revealed subtle differences in the distance between the
268 catalytic triad residues (S-D-H) in all the 6 *PDATs* (Supplemental Figure S6a).

269 *PDAT1* and *PDAT2* had similar 3D structures and they were well superposed with
270 each other except the flexible coil region and transmembrane helix α_1 at the
271 N-terminal part and the last helix α at the C-terminal (Figure 5F; Supplemental Figure
272 S6). The amino acid differences in the N-terminus of *PDAT1* and *PDAT2* were mainly
273 located at the flexible coil region before the transmembrane helix (Supplemental
274 Figure 7a), which might lead to different substrate recognition and binding activities
275 for *PDAT1* and *PDAT2*. Electrostatic surface potential of *PDATs* indicated that
276 negatively charged amino acids in the C-terminus of *PDAT1* (T708R and F717R),

277 compared to PDAT2, drew the last helix closer to the core structure (Supplemental
278 Figure 7b). These differences might lead to different TAG biosynthetic activities.
279 Therefore, PDAT3 and PDAT6, which possess similar structures to PDAT2 at the
280 N-terminus and C-terminus respectively (Supplemental Figure S6), recovered the
281 partial enzyme activity.

282

283 **PtPDAT uses phosphatidylcholine (PC) and monogalactosyldiacylglycerol**
284 **(MGDG) as substrates**

285 To generate *PtPDAT* knockout (KO) lines, two sgRNAs (Psg1 and Psg2) were
286 designed targeting the N-terminus, and one mutant at each site was selected from the
287 obtained transformants for further analysis. PDAT-KO34 line (the Psg1 mutant) and
288 PDAT-KO52 line (the Psg2 mutant) had a 2 bp deletion and 1 bp insertion,
289 respectively (Supplemental Figure S8). Both mutants introduced a premature stop
290 codon and neither of the two KO lines harbored wild-type alleles. The two *PDAT*-KO
291 lines exhibited increased growth compared with WT cells from day 6 to 14 (Figure
292 6A), and nitrate utilization of the two *PDAT*-KO lines was slightly slower than that of
293 WT on day 2 (Figure 6B). Neutral lipid content detected by Nile red in *PDAT*-KO
294 cells decreased throughout the cultivation period compared with what was observed in
295 WT cells (Figure 6C). TLC analysis also showed that TAG content was decreased by
296 30-32% on days 12 and 14 in the two *PDAT*-KO lines compared with WT
297 (Supplemental Figure S4b).

298 To uncover the *PDAT*-KO-associated global changes of lipid composition, lipid
299 class profiling by mass spectrometry analysis was conducted for WT and *PDAT*-KO
300 cultures harvested on day 3, 4 and 5 under air-bubbling condition. Most abundant
301 polar lipids were three plastid glycolipids, i.e. MGDG,
302 digalactosyldiacylglycerol (DGDG) and sulfoquinovosyldiacylglycerol (SQDG). In
303 the course of cultivation, the content of SQDG, MGDG, phosphatidylglycerol (PG),
304 phosphatidylinositol (PI), phosphatidylethanolamine (PE) and PC decreased, while

305 that of TAG increased with time (Figure 7A). WT cells contained TAG of about
306 30 mol%, a content comparable to the previous works (Lupette et al., 2019), on day 4
307 when nitrogen was exhausted. Consistent with Nile red and TLC assay results, TAG
308 content on day 4 and 5 decreased by 19-57% and 12-38% respectively in the two
309 *PDAT*-KO lines compared with WT cells. Accordingly, the KO lines accumulated a
310 much higher level of MGDG (increased by 17-54%) and PC (increased by 70-80%)
311 on day 5 (Figure 7A), indicating that MGDG and PC could be the substrates of
312 PtPDAT.

313 The composition of acyl molecular species, especially in MGDG, PC and TAG, was
314 different between *PDAT*-KO lines and WT strain (Figure 7, B, C and D; Supplemental
315 Figure S9). The major MGDG species detected in our study were 20:5/16-C and
316 16-C/16-C (Figure 7B), which indicated that they were chloroplast-derived or
317 “prokaryotic” lipids. Similar fatty acid composition could also be observed in DGDG
318 (Supplemental Figure S9c). The increased MGDG species in *PDAT*-KO lines,
319 compared with WT cells, were mainly unsaturated 16-C/16:3 and 16:1/16:2 (Figure
320 7B), while the decreased TAG species in *PDAT*-KO lines were mainly 48:1
321 (16:0/16:0/16:1) and some more unsaturated species such as 52:6 (20:5/16:0/16:1) and
322 52:7 (20:5/16:1/16:1) (Figure 7C). The PC species of 16:1/16-C, 20:5/16:1 and
323 20:5/16:2 in *PDAT*-KO lines were higher than those in WT (Figure 7D). These
324 acyl-donors were in line with those fatty acids in the decreased species of TAG.
325 Saturated 16-C acyl-donors in MGDG, which might not be transferred to TAG due to
326 the loss of PDAT activity in *PDAT*-KO lines, were rapidly desaturated resulting in a
327 lower mol% of MGDG 16:1/16:0 (Figure 7B) and increased unsaturated MGDG. In
328 addition, mol% of unsaturated DGDGs, PEs and DGTA
329 (diacylglycerylhydroxymethyltrimethyl- β -alanine, 20:5/16:1 and 20:5/16:2) also
330 increased in *PDAT*-KO lines (Supplemental Figure S9, c, d and f). The composition of
331 acyl molecular species in PG, DAG and SQDG showed little difference between KO
332 lines and WT (Supplemental Figure S9, a, b and e). The relative abundance of C16:0

333 and C18:3 in the total lipids was lower in *PDAT*-KO lines, while that of C16:1, C16:2,
334 C16:3 and C18:1 tended to increase (Supplemental Figure S10). Considering the
335 species of fatty acids that are strongly decreased in TAG and those were
336 correspondingly increased in membrane lipids, the dynamic recycling, mainly of 16:1,
337 was impaired in *PDAT*-KO strains. In addition, *PDAT* knockout had contribution to
338 the feeding of membrane lipids with eicosapentaenoic acid (C20:5).

339

340 **PtPDAT shares a closer phylogenetic relationship with green algal PDATs**

341 Our previous study suggested that stramenopile PDATs were polyphyletic and diatom
342 PDATs did not clustered together with *Nannochloropsis* PDATs (Yang et al., 2022).
343 Maximum-likelihood tree showed PtPDAT and *Thalassiosira pseudonana* PDAT
344 exhibited a closer phylogenetic relationship with green algal PDATs and were
345 clustered with six clades of land plant PDATs (clades II–VII) (Pan et al., 2015)
346 (Figure 8A). Like *N. oceanica* PDAT, PtPDAT also harbors domains conserved in
347 PDAT polypeptides based on protein sequence alignment (Yang et al., 2022). We used
348 HECTAR, a method to predict subcellular targeting in Heterokonts/Stramenopiles
349 (Gschloessl et al., 2008), to predict the subcellular localization of PtPDAT. It
350 contained a signal anchor while had no targeting sequence to chloroplast or
351 mitochondrion. Prediction by TargetP (<http://www.cbs.dtu.dk/services/TargetP/>) did
352 not support the localization in chloroplast or mitochondrion, either. Besides, PtPDAT
353 was not predicted to be a plastid-targeted protein by ASAFind
354 (<https://rocaplab.ocean.washington.edu/tools/asafind/>). Laser confocal scanning
355 microscopy analyses showed that the green eGFP signal formed regular layers
356 surrounding the chlorophyll autofluorescence in the PtPDAT-eGFP transformants
357 (Figure 8B). PtPDAT was predicted to have a transmembrane domain (Supplemental
358 Figure S2), and the green fluorescence signal was not identical to the reported patterns
359 of marker proteins located in the three outermost membranes of the plastid (Liu et al.,
360 2016b), but was consistent with the reported fluorescence patterns of two

361 endosymbiont-derived inner membrane (IEM)-localized triose phosphate translocators
362 (TPTs, TPT4a and TPT4b) (Moog et al., 2015). Thus, PtPDAT should be localized in
363 IEM. A three-dimensional visualization of PtPDAT-eGFP and chlorophyll
364 autofluorescence signals further supported the IEM localization of PtPDAT (Figure
365 8B; Supplemental Figure S11).

366

367 **Discussion**

368 **Allelic variation and allele-specific expression of the *PtPDAT* gene**

369 Allelic variation is considered to be an important driving force in adaptive evolution
370 to distinct environmental conditions. In plants, the naturally occurring intraspecific
371 allelic diversity has been reported to contribute to phenotypic variations by altering
372 the structure, function, or expression level of the encoded protein (Tao et al., 2009;
373 Kloosterman et al., 2013; Zhao et al., 2016). Natural allelic variation has also been
374 widely found in diatoms from predominantly coastal regions and the gene diversity
375 (heterozygosity) estimates range from 39 to 88% (Godhe and Rynearson, 2017). In *F.*
376 *cylindrus* genome, about 29% of the predicted protein-coding genes are represented
377 by diverged alleles, among which about 66% showed allelic imbalance in at least one
378 experiment (Mock et al., 2017). Sequence divergence between alleles in *F. cylindrus*
379 is up to 6%, and there was a 2.6% nucleotide base difference of DNA sequence
380 between *PtPDAT* allele 1 and allele 2 in our study (Supplemental Figure S1). More
381 than 80% of the predicted protein-coding genes have homoeologous counterparts and
382 61% of homoeologous genes show differential expression in the allodiploid diatom *F.*
383 *solaris* (Tanaka et al., 2015). In *P. tricornutum*, about 60% of the genes have at least
384 one SNV in transcripts of two alleles and 38% of the genes with SNVs in their alleles
385 are not equally expressed (Hoguin et al., 2021).

386 *PtPDAT* was reported as an ASE gene in previous analysis of RNA-Seq data of WT
387 *P. tricornutum* at late exponential phase (Hoguin et al., 2021). In our study, 87 sample
388 data related to different nitrogen sources, sufficient or limited nitrogen or phosphate,

389 and dark or nocodazole treatment were analyzed. The results showed that ASE of the
390 *PtPDAT* gene was mainly regulated by the nitrate concentration in the medium
391 (Figure 1; Supplemental Figure S3) and the *PDAT* allele 2-specific transcript level
392 dramatically increased in response to nitrogen limitation, indicating that ASE plays a
393 key role in TAG accumulation in *P. tricornutum*. Promoter region analysis indicated
394 the variants of the regulatory elements in the two *PDAT* alleles (Supplemental Figure
395 S12; Supplemental Dataset S2), and these variants might be highly responsive to the
396 fluctuating nitrate concentrations but further experiments should be performed to
397 interpret the regulation modes. Abida et al. (2015) found the TAG accumulation with
398 different time scales and magnitudes in strains Pt1-Pt10 under nitrogen deprivation,
399 which might result from the gene diversity of *PtPDAT* in the 10 strains (Supplemental
400 Figure S1). Except for the different expression levels of *PDAT* in the 10 strains
401 (Supplemental Figure S13a), the amino acid variation of *PDATs* in the four Genotypes
402 (Supplemental Figure S2) that changed the protein folding (Supplemental Figure S13b)
403 might also have an impact on TAG biosynthetic activities.

404

405 **Distinct activity of proteins encoded by *PDAT* allele 1 and allele 2**

406 Proteins encoded by *PDAT* allele 1 and allele 2, two differentially expressed alleles in
407 *P. tricornutum*, showed distinct enzymatic activities both in *P. tricornutum* and in
408 yeast. Overexpression of allele 2 rather than allele 1 gave rise to a substantial TAG
409 accumulation; likewise no TAG biosynthetic activity was detected in TAG-deficient
410 yeast expressing *PDAT* allele 1 encoding protein (Figure 5). In particular, even before
411 day 4 when the nitrate was not exhausted (Figure 3), TAG content was much higher in
412 allele 2 OE lines than that of WT, indicating TAG accumulation occurred at the
413 exponential growth stage (the nonstress condition). In contrast, a negligible increase
414 was present at the stationary stage (day 8 and 10) in the allele 1 OE strains (Figure 3).
415 Distinct enzymatic activities should be attributed to the amino acid variants between
416 the two allele encoding proteins, though these variants did not exist in the seven most

417 characteristically conserved domains of PDAT polypeptides except the substitution of
418 alanine by valine and of aspartate by glutamate in conserved domains I and II
419 respectively (Supplemental Figure S2). Both of the substitutions took place in
420 segment 2 of PtPDAT in our study, and replacing this segment in *PDAT* allele 1 with
421 the corresponding segment from allele 2 did not change the enzymatic activity,
422 indicating the two amino acid variants had no effect on the activity. Seemingly trivial
423 amino acid changes in N-terminal and the C-terminal should be taken into
424 consideration to interpret the causes of the loss of enzymatic activity in the allele 1
425 encoding protein. Amino acid substitutions/insertions in animal or plant DGATs have
426 been reported to substantially change the enzymatic activities *in vivo* or *in vitro*. A
427 nonconservative K232A substitution in bovine *DGAT1* gene has a major effect on
428 milk fat content (Grisart et al., 2002; Winter et al., 2002), and substitutions of two
429 amino acids (N6D or A26P) near the extreme N-terminus in peanut DGAT2 increased
430 the enzymatic activity level in yeast (Zheng et al., 2017). A phenylalanine insertion in
431 maize DGAT1-2 in the predicted transmembrane domain alters the overall protein
432 structure and influences the enzymatic activity (Zheng et al., 2008). The +1 position
433 of the catalytic serine has been proved to be particularly important for the function of
434 PDATs and mutagenesis at this site (L206M or L206Y) in human LCAT is likely to
435 change the conformation of the active site pocket (Falarz et al., 2020). Although
436 predicted structures of the two PDAT allele encoding proteins are very similar, protein
437 structure resolution is important to identify the effect of amino acid changes on the
438 function of the two allele encoding proteins in the future.

439 Surprisingly, more severe inhibition of growth was observed in allele 1 OE strains
440 compared with allele 2 OE strains during the period of TAG accumulation (from day
441 6 to 10, Figure 3), suggesting the potential but unknown activities of the
442 acyltransferase encoded by allele 1, and thus further studies are required. Indeed, *C.*
443 *reinhardtii* PDAT has been reported to be a multifunctional protein possessing
444 acyltransferase and lipase activities with broad substrate specificity *in vitro* (Yoon et

445 al., 2012).

446

447 **Subcellular location and substrates of PtPDAT**

448 Phylogenetic analyses indicated that diatom PDATs grouped in a sister clade to green
449 algae clade with a strong support, and they were far away from PDATs of the
450 secondary endosymbiotic algae *N. oceanica* and *N. gaditana* (Figure 8A). *N. oceanica*
451 PDAT resided at the chloroplast endoplasmic reticulum (Yang et al., 2022), while
452 PtPDAT was proved to be located in IEM here. These results revealed different
453 origins of PDATs in the two secondary endosymbiotic algae. *C. reinhardtii* PDAT
454 harbors a chloroplast transit peptide and was identified in the chloroplast based on the
455 proteomic analysis (Terashima et al., 2011). Phylogenetic relationship and subcellular
456 location suggested that diatom PDAT shared the same origin in evolution with green
457 algal PDATs.

458 *Arabidopsis* PDAT1 and *S. cerevisiae* PDAT possess similar substrate specificity,
459 and prefer the zwitterionic phospholipids PE and PC, particularly PE, as acyl donor
460 for TAG biosynthesis (Dahlqvist et al., 2000; Ghosal et al., 2007; Fan et al., 2013,
461 2014; Feng et al., 2019); and *N. oceanica* PDAT uses PE and PG as acyl-donors *in*
462 *vivo* and shows an apparent preference for PG *in vitro* (Yang et al., 2022). *C.*
463 *reinhardtii* PDAT, as an acyltransferase and a lipase, can use a broad range of
464 substrates, including TAG, phospholipids, galactolipids, and cholesteryl esters (Yoon
465 et al., 2012) *in vitro*. However, *in vivo*, it uses chloroplast membrane lipids,
466 particularly MGDG, SQDG and PG, as substrates to synthesize TAG, and only
467 MGDG could be efficiently utilized as an acyl donor (Yoon et al., 2012). The green
468 alga *M. incise* preferentially used PC as the substrate as revealed by *in vitro* enzyme
469 activity assay (Liu et al., 2016c). IEM-localized PtPDAT also used chloroplast lipids
470 MGDG and PC as substrates *in vivo* in our study, and MGDG was preferred, which
471 further supports a common origin in evolution of PtPDAT and green algal PDATs.
472 Thus, like plant PDATs, PtPDAT contributes to membrane remodeling and

473 degradation in this diatom.

474

475 **Materials and methods**

476 **Strain and Growth Conditions**

477 Axenic culture of the diatom *Phaeodactylum tricornutum* Bohlin (CCMP2561) was
478 obtained from the Provasoli-Guillard National Center for Marine Algae and
479 Microbiota at Bigelow Laboratory for Ocean Sciences, USA. For the analysis of lipids,
480 reverse transcription quantitative PCR (RT-qPCR) and western blot, cultures were
481 inoculated with 2×10^5 cells mL⁻¹ in artificial seawater enriched with f/2 (Guillard,
482 1975) and cultivated at 22°C bubbling with filtrated air under continuous illumination
483 of 100 $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$. For the growth experiment, nitrate and phosphate in the
484 media were set at different concentrations (500 μM nitrate for nitrogen-limited; 7056
485 μM nitrate/289 μM phosphate for nitrogen/phosphorus-replete) and cells were
486 cultivated on a shaking table with continuous shaking at 60 rpm.

487

488 ***PtPDAT* Sequence Analysis *in Silico* and Experimental Verification**

489 Whole genome resequencing data of ten accessions of *P. tricornutum* (Rastogi et al.,
490 2020) were used for the detection of single nucleotide variants (SNVs) of *PtPDAT*
491 gene. Sequences of two allelic *PtPDAT* promoters were assembled using
492 corresponding fragments in the resequencing data of Pt1 (Rastogi et al., 2020).
493 Reference sequence was from the latest version of the *P. tricornutum* genome
494 (http://protists.ensembl.org/Phaeodactylum_tricornutum; *Phatr3_J8860*). The
495 full-length cDNA of *PtPDAT* was further verified based on EST libraries (Maheswari
496 et al., 2010) and 5'-RACE data. 5'-RACE was performed with specific antisense
497 primers for *5race1*, *5race2*, and *5race3* (Supplemental Dataset S3) in combination
498 with a universal sense primer according to the instructions of the manufacturer
499 (5'-Full RACE Kit, TaKaRa, Japan).

500 RNA-seq data with NCBI accession IDs PRJNA242990, PRJEB11970,

501 PRJNA382762, PRJNA311568, PRJNA278661, PRJNA101889 and PRJNA322663
502 were used for *in silico* analysis of *PtPDAT* ASE. CDS of the two *PtPDAT* alleles
503 (designated as *PDAT1* and *PDAT2*) were used as the reference sequence. Reads with
504 remaining adapters or with a low quality value ($Q \leq 20$ more than 30%) were discarded
505 by NGS QC Toolkit (v2.3.3), and then the clean data were aligned to the two *PDAT*
506 alleles using Bowtie2 (v2.1.0) (Langmead and Salzberg, 2012) and alignments
507 without mismatch or gap were selected for subsequent statistical analysis. We counted
508 the number of reads matched to *PDAT* allele 1 and *PDAT* allele 2 and got the
509 coverage of the reads on each base of two alleles using bedtools (v2.25.0). An
510 in-house bash script was used to calculate the number of reads at SNP of the two
511 alleles. Then RPKM values were calculated from the number of reads.

512

513 **Western blot analysis**

514 For the analysis of PtPDAT protein expression, wild-type cells grown for 12, 24, 36,
515 48, 60, 72, 84, 108 and 156 h were harvested by centrifugation (3000 g, 15 min), and
516 so were WT and PDAT overexpression (OE) lines grown for 2, 3 and 4 d. By
517 ultrasonication for 10 min at 10% intensity, total protein was extracted from cells
518 washed with PBS before liquid nitrogen freezing in western and IP lysis buffer
519 supplemented with 1 mM phenylmethanesulfonyl fluoride (Beyotime Biotechnology,
520 China). Extracts were centrifuged and protein concentration was determined using the
521 Beyotime BCA protein assay kit. After denaturation, total protein was separated by 10%
522 SDS-PAGE gel and then transferred to polyvinylidene fluoride membrane. The
523 membranes were incubated with anti-PtPDAT antibody followed by HRP-conjugated
524 secondary antibody, and blots were developed with Immobilon Western
525 Chemiluminescent HRP Substrate (Millipore, USA). Images of the blots were
526 obtained using an ImageQuant LAS 4000 mini apparatus (GE Healthcare Life
527 Science). Polyclonal antibody of PtPDAT was generated by Genscript Corporation
528 (Nanjing, China) against a synthetic peptide CDEKSSRTWHPTR.

529

530 **Reverse transcription quantitative PCR (RT-qPCR)**

531 For RNA extraction, cells of WT and PDAT OE lines grown for 3, 4 and 5 days were
532 harvested by centrifugation (3000 g, 15 min), and then washed with PBS. Total RNA
533 was extracted with TRIzol reagent (Invitrogen) from liquid nitrogen-frozen cells.
534 DNA was removed by gDNA Eraser and first-strand cDNA was synthesized using
535 PrimeScript RT reagent Kit (TaKaRa). RT-qPCR analysis was performed with
536 LightCycler 480 SYBR Green I Master Mix and a LightCycler 480 Machine (Roche).
537 Primers used for RT-qPCR are shown in Supplemental Dataset S3. *Histone H4*, whose
538 mRNA expression was stable under the cultivation condition (Siaut et al., 2007),
539 served as the internal control gene for normalizing the expression of *PtPDAT*.
540 Technical triplicates were performed for duplicate cultures. The mean values of
541 relative mRNA quantities were calculated according to the $2^{-\Delta\Delta C_t}$ -method.

542 Ratio of background expression level of *PDAT* allele 2 and 1 (*N*) was estimated
543 using the following formula: $N = (L2-L1)/(L1-1)$, where *L1* is the relative level of
544 *PDAT* in allele 1 OE lines detected by RT-qPCR using universal primers, and *L2* is
545 the relative level of *PDAT* in allele 1 OE lines detected by RT-qPCR using allele
546 1-specific primers. The specificity of *PDAT* allele-specific primers was validated by
547 PCR, RT-qPCR, and sequencing (Supplemental Figure S14).

548 For the comparison of *PDAT* mRNA levels in different accessions of *P. tricornutum*
549 (Pt1-Pt10 strains), exponential growth cells were used for the total RNA extraction.

550

551 **Overexpression, GFP Fusion and Knockout Construction**

552 To generate allele-specific overexpression constructs, full-length CDS of two
553 *PtPDAT* alleles were PCR amplified respectively with different allele-specific primer
554 pairs using cDNA as template and cloned into the pPha-T1 vector (between the *Xba*I
555 and *Hind*III sites). For the localization analysis of PtPDAT, the CDS of *PDAT2* was
556 inserted into the pPhaT1-eGFP vector between the *Eco*RI and *Kpn*I sites (Zhang and

557 Hu, 2014) to generate a C-terminal GFP fusion protein PtPDAT-eGFP. Inserts of all
558 constructs were sequenced to confirm that no unintended nucleotide changes were
559 present.

560 The CRISPR/Cas9 vector developed by Nymark et al. (2016) was employed to edit
561 *PDAT* in *P. tricornutum*. Candidate sgRNAs were generated using the Phyto
562 CRISPEX online package (Rastogi et al., 2016) and two 20 nt guide RNAs Psg1
563 (*PDAT_293*: GTTTCGGTGTCCTGGATCCT) and Psg2 (*PDAT_523*:
564 TATCCTGTGGTCATGATGCC) were chosen based on the position and allelic
565 consensus sequence within *PDAT* gene. Complementary oligos containing flanking
566 overhang sequences were then synthesized and ligated into pKS diaCas9_sgRNA
567 plasmid (Nymark et al., 2016). The constructs were verified via sequencing and were
568 co-transformed into WT cells with the pPha-T1 vector endowing zeocin-resistance.

569 Vectors were linearized before transformation and then introduced into *P.*
570 *tricornutum* by electroporation (Zhang and Hu, 2014). OE and eGFP fusion
571 transformants were screened by checking the integration of the target gene with the
572 primers *pPhaT1_fw* and *pPhaT602_rev*. Putative PtPDAT knockout lines were
573 screened via PCR (primers *oe-fw* and *523-rev* for *PDAT_293* KO lines, primers
574 *293-fw* and *979-rev* for *PDAT_523* KO lines) and the PCR product was sequenced
575 with primers *oe-fw* or *979-rev* to identify the mutation in the corresponding region.
576 All primers for the PCR are listed in Supplemental Dataset S3.

577

578 **Growth and TAG Content Analysis**

579 Growth was determined by cell enumeration every 2 d using a Malassez chamber.
580 Concentrations of nitrate in the culture were measured spectrophotometrically at 220
581 nm (Collos et al., 1999) and those of phosphate with the colorimetric
582 phosphomolybdate complex method at 885 nm (Harrison 1988). Relative abundance
583 of the neutral lipid was estimated by fluorescence spectroscopy (Perkin-Elmer, LS 55)
584 with the fluorescent dye Nile red (Sigma-Aldrich) according to procedures described

585 before (Ge et al., 2014). Total lipids were extracted from the cell pellet (Bligh and
586 Dyer, 1959) and were separated by one-dimensional thin-layer chromatography (TLC)
587 for the comparison of TAG contents (Reiser and Somerville, 1997) on silica gel plates
588 60 F254 (Merck KGaA, Darmstadt, Germany) with triolein as the standard
589 (Sigma-Aldrich). TAGs were identified by staining with iodine.

590

591 **Functional complementation of PtPDAT in the TAG-deficient yeast**

592 The full-length CDS of two *PtPDAT* alleles were PCR amplified (primers *HR-fw* and
593 *p1HR-rev* for *PDAT1* allele, primers *HR-fw* and *p2HR-rev* for *PDAT2* allele) with *P.*
594 *tricornutum* cDNA as template. To reconstruct the *PDAT* CDS, the coding region of
595 *PtPDAT* gene was divided into four segments as shown in Figure 5A and
596 Supplemental Figure S2. Each segment was amplified with the corresponding primer
597 pair and template as shown in Supplemental Dataset S3 to generate four reconstructed
598 *PDATs* (*PDAT3-6*). The six *PDAT* sequences were cloned into the yeast expression
599 vector pYES2-CT (Invitrogen), respectively, using a ClonExpress Ultra One Step
600 Cloning Kit (Vazyme, C115, China). The resulting vectors were confirmed by
601 sequencing and then were transformed into the yeast (*Saccharomyces cerevisiae*)
602 TAG-deficient strain H1246 using S.c. EasyComp Transformation Kit (Invitrogen).
603 The empty vector (EV) pYES2-CT was used as the negative control. Transformants
604 were selected on solid SC-uracil medium (Teknova, Hollister, CA, USA). Yeast cells
605 were grown in liquid SC-uracil medium with 2% (w/v) raffinose for 24 h and
606 harvested by centrifugation (1,500 g, 5 min). For induction, cells were resuspended in
607 SC-uracil medium containing 2% (w/v) galactose with an initial OD₆₀₀ of 0.4. After
608 the induction of 2 d, total lipids from yeast cells were extracted as described by Mao
609 et al. (2019) and were separated by one-dimensional TLC on silica gel plates 60 using
610 hexane/tert-butylmethyl ether/acetic acid (80/20/2, by volume) as the mobile phase.
611 Lipids were detected by spraying the TLC plate with 10% (w/v) CuSO₄ in 8% (v/v)
612 phosphoric acid, followed by charring at 180°C for 3 min.

613

614 **Fluorescent and Confocal Microscopy Analyses**

615 Neutral lipids in yeast cells were stained with the fluorescence BODIPY™ 505/515
616 (Invitrogen) and then visualized under an Olympus BX51 Fluorescence Microscope
617 (Olympus) with excitation at 488 nm and emission between 505 and 530 nm (Zhang
618 et al., 2021). To observe the subcellular localization of PtPDAT, transformants were
619 visualized under a Leica TCS SP8 laser scanning confocal microscope with Leica
620 HyD hybrid detector in photon-counting mode (gain 10%). Chlorophyll
621 autofluorescence and GFP fluorescence were excited at 488 nm (30% intensity), and
622 were detected with 630–690 nm and 500–550 nm, respectively.

623

624 **Lipid analysis**

625 Fatty acid methyl esters prepared from the transesterification of yeast lipids were
626 analyzed using an Agilent 7890 capillary gas chromatograph equipped with a 5975 C
627 mass spectrometry detector and a HP-88 capillary column (60 m × 0.25 mm; Agilent
628 Technologies) (Liu et al., 2016a).

629 Triplicate cultures of wild-type *P. tricornutum* and *PtPDAT* KO lines were
630 harvested by centrifugation (3000 g, 10 min) and pellets were quickly frozen in liquid
631 nitrogen and then lyophilized. Lipids from the 100 mg freeze-dried cells per sample
632 were extracted and quantified by mass spectrometric analysis according to Lupette et
633 al. (2019).

634

635 **Protein Structure Prediction and Phylogenetic Analysis**

636 The protein models were predicted by AlphaFold2. The structural models were viewed
637 using PYMOL software. The majority of PDAT protein families (Supplemental
638 Dataset S4) were incorporated into a maximum-likelihood tree by PHYML algorithm.

639

640 **Statistical treatments**

641 Data were statistically analyzed using analysis of variance (ANOVA) followed by the
642 LSD test ($P < 0.05$). Student's t test was performed to assess the significant
643 differences ($P < 0.05$) between two data.

644

645 **Accession Numbers**

646 Sequence data of *P. tricornutum* are from the latest version of the *P. tricornutum*
647 genome (http://protists.ensembl.org/Phaeodactylum_tricornutum). The accession
648 numbers used for phylogenetic analysis are in Supplemental Dataset S4.

649

650 **Supplemental Data**

651 **Supplemental Figure S1.** Single nucleotide variants (SNVs) within protein coding
652 regions of the *PDAT* gene in different accessions of *P. tricornutum*.

653 **Supplemental Figure S2.** Sequence alignment of the amino acids of PDAT in
654 different accessions of *P. tricornutum*.

655 **Supplemental Figure S3.** The analysis of allele-specific expression of *PDAT* gene in
656 response to various environmental stimuli based on the published RNA sequencing
657 data.

658 **Supplemental Figure S4.** Triacylglycerol (TAG) contents in *P. tricornutum* analyzed
659 by thin-layer chromatography (TLC).

660 **Supplemental Figure S5.** Effect of *PDAT* allele 1 overexpression on growth and
661 neutral lipid accumulation.

662 **Supplemental Figure S6.** Predicted structure of PDATs by AlphaFold.

663 **Supplemental Figure S7.** Close up view of PDAT predicted structures by AlphaFold.

664 **Supplemental Figure S8.** Targeted mutations of *PDAT* in *P. tricornutum* induced by
665 CRISPR-Cas9.

666 **Supplemental Figure S9.** Composition of PG, DAG, DGDG, PE, SQDG and DGTA
667 (A-F) (mol% within their class, respectively) in wild-type (WT) and *PDAT* knockout
668 lines (PDAT-KO34 and PDAT-KO52).

669 **Supplemental Figure S10.** Fatty acid composition of total lipids (mol%) in wild-type

670 (WT) and *PDAT* knockout lines (PDAT-KO34 and PDAT-KO52).
671 **Supplemental Figure S11.** Synthetic 3D fluorescent microscope images of cells
672 expressing PDAT-eGFP.
673 **Supplemental Figure S12.** Sequence alignment and functional analysis of *PDAT*
674 alleles in the promoter region.
675 **Supplemental Figure S13.** Transcript levels and structures of PDATs in different *P.*
676 *tricornutum* Genotypes.
677 **Supplemental Figure S14.** Validation on the specificity of allele-specific primers.
678 **Supplemental Dataset S1.** Reads and RPKM values of *PDAT* alleles and abundance
679 of each SNP by re-analysis of previously published RNA sequencing data.
680 **Supplemental Dataset S2.** Promoter elements of *PDAT* gene.
681 **Supplemental Dataset S3.** Primers used in this study for 5'-RACE, RT-qPCR,
682 plasmid construction, and transformant identification.
683 **Supplemental Dataset S4.** Accessions of PDAT proteins used for the phylogenetic
684 tree construction.

685

686 **Funding information**

687 This work was supported by the National Natural Science Foundation of China
688 (41976119, 42376131 and 31800302) and International Partnership Program of
689 Chinese Academy of Sciences (Grant No. 075GJHZ2022014MI). E.M. and J.J. were
690 supported by the French National Research Agency (GRAL Labex
691 ANR-10-LABEX-04, EUR CBS ANR-17-EURE-0003; ANR-15-IDEX-02;
692 ANR-20-CE02-0020; NR-21-CE02-0021) and Institut Carnot 3BCAR. E.M. and H.H.
693 were supported by a CEA-CAS bilateral program.

694 *Conflict of interest statement.* None declared.

695

696 **Author contributions:** H.H. conceived and designed the research. Y.P. and X.W.
697 performed the experiments. Y.P., W.Z., J.J., E.M., J.L., X.X. and H.H. analyzed the

698 data. Y.P. drafted the manuscript and H.H. made critical revisions. All authors read
699 and approved the final article.

700

701

702 **Figure legends**

703 **Figure 1** Transcription and protein levels of PDAT in *P. tricornutum* strain
704 CCMP2561 (Pt1). A. The analysis of allele-specific expression (ASE) of *PDAT* gene
705 in response to nitrogen stress based on the published RNA sequencing data. B. The
706 analysis of ASE of *PDAT* gene in response to phosphate stress based on the published
707 RNA sequencing data. C. Time course of expression of PDAT protein detected by
708 western blotting during batch culture. *PDAT-com* refers to the reads mapped to the
709 regions with no difference between *PDAT* alleles; *PDAT1-sp* and *PDAT2-sp* refer to
710 the reads specifically mapped to the two *PDAT* alleles 1 and 2, respectively. The
711 transcriptome data of *PTI+N* (nitrogen-replete, i.e. treated with 882 μM nitrate for
712 48 h) and *PTI-N* (nitrogen-free treatment for 48 h) were obtained from Levitan et al.
713 (2015); data of *control 4h, 8h, 20h, 36h* (88.2 μM nitrate and 36.3 μM phosphate),
714 data of *no nitrate 4h, 8h, 20h*, and data of *no phosphate 36h* were obtained from
715 Matthijs et al., 2016; data of *WT_NH4* (mid-exponential phase cells with 880 μM
716 NH_4^+), *WT_2h N-free* (2 h of incubation in N-free media) and *WT_1h, 18h, 42h, 66h,*
717 *90h, 114h, 138h, 162h, 228h* (1~228 h after the transfer of cultures into 300 μM
718 NO_3^- media) were obtained from McCarthy et al. (2017); and data of *control 4d, 8d*
719 (grown with 36.3 μM phosphate for 4 and 8 days), *Pi depletion 4d, 8d* (grown without
720 phosphate for 4 and 8 days), and *recovery 4d* (phosphate supplemented on day 4 to Pi
721 depleted cultures, then grown for 4 days) were obtained from Cruz de Carvalho et al.
722 (2016). * (by *t*-test and compared with *PTI+N*) and different letters (by one-way
723 ANOVA, lowercase and upper letters are refer allele 1 and 2, respectively) above the
724 bars indicate statistically significant differences ($P < 0.05$). Data represent mean \pm *SD*
725 ($n=3$).

726 **Figure 2** Expression of PDAT in allele-specific overexpression strains of *P.*
727 *tricornutum*. A. Relative mRNA levels of *PDAT* detected by real-time quantitative
728 PCR using universal primers (*PDAT*) or allele-specific primers (*PDAT1* or *PDAT2*).
729 Error bars represent standard deviation (SD) of triplicate technical replicates of
730 duplicate cultures. B. Expression of PDAT protein detected by western blotting during
731 batch culture. WT: wild-type; PDAT1-OE1 and PDAT1-OE4: two overexpression
732 lines of *PDAT* allele 1; PDAT2-OE2 and PDAT2-OE4: two overexpression lines of
733 *PDAT* allele 2.

734 **Figure 3** Effect of *PDAT* gene overexpression on growth and neutral lipid
735 accumulation in *P. tricornutum*. A and D: growth, B and E: nitrate utilization, C and F:
736 accumulation of neutral lipid detected by Nile Red assay (fluorescence intensity
737 normalized to cell number) of wild-type (WT) and four *PDAT* allele-specific
738 overexpression strains grown in f/2 (NaNO₃ concentration was reduced to 500 μM)
739 enriched artificial seawater medium. PDAT1-OE1 and PDAT1-OE4: two
740 overexpression lines of *PDAT* allele 1; PDAT2-OE2 and PDAT2-OE4: two
741 overexpression lines of *PDAT* allele 2. Data represent mean ± SD (*n*=3).

742 **Figure 4** Effect of *PDAT* allele 2 overexpression on growth and neutral lipid
743 accumulation. A, D and G: growth, B, E and H: nitrate (black lines) and phosphate
744 (blue lines) utilization, C, F and I: accumulation of neutral lipid detected by Nile Red
745 assay (fluorescence intensity normalized to cell number) of wild-type (WT) and two
746 overexpression lines of *PDAT* allele 2 (PDAT2-OE2 and PDAT2-OE4) grown in f/2
747 enriched artificial seawater medium with nitrate and phosphate of different
748 concentrations (HN+HP: high nitrate and high phosphate, HN+NP: high nitrate and
749 normal phosphate, LN+HP: low nitrate and high phosphate). Data represent mean ±
750 SD (*n*=3).

751 **Figure 5** Schematic diagram, functional analysis in yeast and predicted structure of
752 PDATs. A. Schematic diagram of the PDAT proteins encoded by alleles 1 (*PDAT1*)
753 and 2 (*PDAT2*), and that of the four reconstructed PDATs (PDAT3-6) obtained by

754 replacing the fragment in *PDAT1* with the corresponding segment from *PDAT2*. Pink
755 sections and light blue sections represent the fragments from *PDAT1* and *PDAT2*,
756 respectively. Seventeen different amino acids, the putative transmembrane domain
757 (TM), the catalytic triad (S-D-H) and the pfam of lecithin cholesterol acyltransferase
758 (LCAT) are marked in *PDAT1* and *PDAT2*. B. Microscopic observation of
759 *PDAT*-expressing H1246 cells (the TAG-deficient yeast strain) stained by Bodipy.
760 Green fluorescence indicates the Bodipy-bound TAG-filled lipid droplets. C. TAG
761 contents in the *PDAT*-expressing H1246 cells analyzed by TLC. D. The percentage of
762 TAG based on total fatty acids (TFA). E. Fatty acid composition of TAG from
763 *PDAT*-expressing H1246 cells. F. Predicted structure of *PDATs*. Empty vector (EV)
764 was used as the negative control (B to E). Data in D and E are expressed as
765 mean \pm *SD* (*n*=3). FFA, free FA. Bar, 5 μ m.

766 **Figure 6** Effect of *PDAT* knockout on growth and neutral lipid accumulation. A:
767 growth, B: nitrate utilization, C: accumulation of neutral lipid detected by Nile Red
768 assay (fluorescence intensity normalized to cell number) of wild-type (WT) and two
769 *PDAT* knockout lines (*PDAT-KO34* and *PDAT-KO52*) grown in f/2 (NaNO₃
770 concentration was reduced to 500 μ M) enriched artificial seawater medium. Data
771 represent mean \pm *SD* (*n*=3).

772 **Figure 7** Lipidomic analysis of wild-type (WT) and *PDAT* knockout lines
773 (*PDAT-KO34* and *PDAT-KO52*). A. Lipid content (mol% of total lipid). B. MGDG
774 composition (mol% within its class). C. TAG composition (mol% within its class). D.
775 PC composition (mol% within its class). SQDG, sulfoquinovosyldiacylglycerol;
776 MGDG, monogalactosyldiacylglycerol; DGDG, digalactosyldiacylglycerol; PG,
777 phosphatidylglycerol; PI, phosphatidylinositol; PE, phosphatidylethanolamine; PC,
778 phosphatidylcholine; DGTA, diacylglycerylhydroxymethyltrimethyl- β -alanine; DAG,
779 diacylglycerol; TAG, triacylglycerol. Data represent mean \pm *SD* (*n*=3).

780 **Figure 8** Phylogenetic relationship among the *PDAT* protein family and subcellular
781 localization of *P. tricornutum* *PDAT*. A. The maximum-likelihood tree of *PDAT*

782 protein family generated by PHYML algorithm. Numbers above branches represent
783 the support values. The scale bar represents an estimated evolutionary distance of 1. B.
784 Fluorescent microscope images of cells expressing PDAT-eGFP. The four panels
785 above from left to right show microscopical images of transmitted light, chlorophyll
786 autofluorescence (red), and GFP fluorescence (green) together with a merged image.
787 Below is the synthetic 3D image of chlorophyll autofluorescence and GFP
788 fluorescence.

789

790 **References**

- 791 **Abida H, Dolch LJ, Meř C, Villanova V, Conte M, Block MA, Finazzi G, Bastien**
792 **O, Tirichine L, Bowler C, et al.** (2015) Membrane glycerolipid remodeling
793 triggered by nitrogen and phosphorus starvation in *Phaeodactylum tricornutum*.
794 *Plant Physiol* **167**: 118–136
- 795 **Aguet F, Guigó Serra R, Montgomery SB** (2017) Genetic effects on gene
796 expression across human tissues. *Nature* **550**: 204–213
- 797 **Allen AE, LaRoche J, Maheswari U, Lommer M, Schauer N, Lopez PJ, Finazzi**
798 **G, Fernie AR, Bowler C** (2008) Whole-cell response of the pennate diatom
799 *Phaeodactylum tricornutum* to iron starvation. *Proc Natl Acad Sci USA* **105**:
800 10438–10443
- 801 **Athenstaedt K, Daum G** (2006) The life cycle of neutral lipids: synthesis, storage
802 and degradation. *Cell Mol Life Sci* **63**:1355–69
- 803 **Aznar-Moreno JA, Durrett TP** (2017) Simultaneous targeting of multiple gene
804 homeologs to alter seed oil production in *Camelina sativa*. *Plant Cell Physiol* **58**:
805 1260–1267
- 806 **Bligh EG, Dyer WJ** (1959) A rapid method of total lipid extraction and purification.
807 *Can J Biochem Physiol* **37**: 911–917
- 808 **Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, Kuo A, Maheswari U,**
809 **Martens C, Maumus F, Otiillar RP, et al.** (2008) The *Phaeodactylum* genome

810 reveals the evolutionary history of diatom genomes. *Nature* **456**: 239–244

811 **Buil A, Brown AA, Lappalainen T, Viñuela A, Davies MN, Zheng HF, Richards**

812 **JB, Glass D, Small KS, Durbin R, et al.** (2015) Gene-gene and gene-environment

813 interactions detected by transcriptome sequence analysis in twins. *Nat Genet* **47**:

814 88–91

815 **Butler T, Kapoore RV, Vaidyanathan S** (2020) *Phaeodactylum tricornutum*: A

816 diatom cell factory. *Trends Biotechnol* **38**: 606–622

817 **Chodavarapu RK, Feng S, Ding B, Simon SA, Lopez D, Jia Y, Wang GL, Meyers**

818 **BC, Jacobsen SE, Pellegrini M** (2012) Transcriptome and methylome interactions

819 in rice hybrids. *Proc Natl Acad Sci USA* **109**: 12040–12045

820 **Collos Y, Mornet F, Sciandra A, Waser N, Larson A, Harrison PJ** (1999) An

821 optical method for the rapid measurement of micromolar concentrations of nitrate

822 in marine phytoplankton cultures. *J Appl Phycol* **11**: 179–184

823 **Cruz de Carvalho MH, Sun HX, Bowler C, Chua NH** (2016) Noncoding and

824 coding transcriptome responses of a marine diatom to phosphate fluctuations. *New*

825 *Phytol* **210**: 497–510

826 **Cui Y, Zhao J, Wang Y, Qin S, Lu Y** (2018) Characterization and engineering of a

827 dual-function diacylglycerol acyltransferase in the oleaginous marine diatom

828 *Phaeodactylum tricornutum*. *Biotechnol Biofuels* **11**: 32

829 **Cui Y, Zheng G, Li X, Lin H, Jiang P, Qin S** (2013) Cloning and characterization of

830 a novel diacylglycerol acyltransferase from the diatom *Phaeodactylum tricornutum*.

831 *J Appl Phycol* **25**: 1509–1512

832 **Dahlqvist A, Ståhl U, Lenman M, Banaś A, Lee M, Sandager L, Ronne H,**

833 **Stymne S** (2000) Phospholipid:diacylglycerol acyltransferase: an enzyme that

834 catalyzes the acyl-CoA-independent formation of triacylglycerol in yeast and plants.

835 *Proc Natl Acad Sci USA* **97**: 6487–6492

836 **Dinamarca J, Levitan O, Kumaraswamy GK, Lun DS, Falkowski PG** (2017)

837 Overexpression of a diacylglycerol acyltransferase gene in *Phaeodactylum*

838 *tricornutum* directs carbon towards lipid biosynthesis. J Phycol **53**: 405–414

839 **Falarz LJ, Xu Y, Caldo KMP, Garroway CJ, Singer SD, Chen G** (2020)

840 Characterization of the diversification of phospholipid:diacylglycerol

841 acyltransferases in the green lineage. Plant J **103**: 2025–2038

842 **Falkowski PG, Katz ME, Knoll AH, Quigg A, Raven JA, Schofield O, Taylor FJR**

843 (2004) The evolution of modern eukaryotic phytoplankton. Science **305(5682)**:

844 354–360

845 **Fan J, Yan C, Roston R, Shanklin J, Xu C** (2014) *Arabidopsis* lipins, PDAT1

846 acyltransferase, and SDP1 triacylglycerol lipase synergistically direct fatty acids

847 toward β -oxidation, thereby maintaining membrane lipid homeostasis. Plant Cell **26**:

848 4119–4134

849 **Fan J, Yan C, Xu C** (2013) Phospholipid:diacylglycerol acyltransferase-mediated

850 triacylglycerol biosynthesis is crucial for protection against fatty acid-induced cell

851 death in growing tissues of *Arabidopsis*. Plant J **76**: 930–942

852 **Fattore N, Bucci F, Bellan A, Bossi S, Maffei ME, Morosinotto T** (2022) An

853 increase in the membrane lipids recycling by PDAT overexpression stimulates the

854 accumulation of triacylglycerol in *Nannochloropsis gaditana*. J Biotechnol **357**:

855 28–37

856 **Feng Y, Zhang Y, Ding W, Wu P, Cao X, Xue S** (2019) Expanding of phospholipid:

857 diacylglycerol acyltransferase (PDAT) from *Saccharomyces cerevisiae* as

858 multifunctional biocatalyst with broad acyl donor/acceptor selectivity. Appl

859 Biochem Biotechnol **88**: 824–835

860 **Field CB, Behrenfeld MJ, Randerson JT, Falkowski P** (1998) Primary production

861 of the biosphere: integrating terrestrial and oceanic components. Science **281**: 237–

862 240

863 **Ge F, Huang W, Chen Z, Zhang C, Xiong Q, Bowler C, Yang J, Xu J, Hu H** (2014)

864 Methylcrotonyl-CoA carboxylase regulates triacylglycerol accumulation in the

865 model diatom *Phaeodactylum tricornutum*. Plant Cell **26(4)**: 1681–1697

866 **Ghosal A, Banas A, Ståhl U, Dahlqvist A, Lindqvist Y, Stymne S** (2007)
867 *Saccharomyces cerevisiae* phospholipid:diacylglycerol acyl transferase (PDAT)
868 devoid of its membrane anchor region is a soluble and active enzyme retaining its
869 substrate specificities. *Biochim Biophys Acta* **1771**: 1457–1463

870 **Godhe A, Rynearson T** (2017) The role of intraspecific variation in the ecological
871 and evolutionary success of diatoms in changing environments. *Philos Trans R Soc*
872 *Lond B Biol Sci* **372**: 20160399

873 **Gong Y, Zhang J, Guo X, Wan X, Liang Z, Hu CJ, Jiang M** (2013) Identification
874 and characterization of PtDGAT2B, an acyltransferase of the DGAT2
875 acyl-Coenzyme A: Diacylglycerol acyltransferase family in the diatom
876 *Phaeodactylum tricornutum*. *FEBS Lett* **587**: 481–487

877 **Grisart B, Coppieters W, Farnir F, Karim L, Ford C, Berzi P, Cambisano N, Mni**
878 **M, Reid S, Simon P, et al.** (2002) Positional candidate cloning of a QTL in dairy
879 cattle: identification of a missense mutation in the bovine *DGATI* gene with major
880 effect on milk yield and composition. *Genome Res* **12**: 222–231

881 **Gschloessl B, Guermeur Y, Cock JM** (2008) HECTAR: a method to predict
882 subcellular targeting in heterokonts. *BMC Bioinform* **9**: 393

883 **Guihéneuf F, Leu S, Zarka A, Khozin-Goldberg I, Khalilov I, Boussiba S.** (2011)
884 Cloning and molecular characterization of a novel acyl-CoA:diacylglycerol
885 acyltransferase 1-like gene (*PtDGATI*) from the diatom *Phaeodactylum*
886 *tricornutum*. *FEBS J* **278**: 3651–3666

887 **Guillard RRL** (1975) Culture of phytoplankton for feeding marine invertebrates. *In*
888 *WL Smith and MH Canley, eds, Culture of Marine Invertebrate Animals. Plenum*
889 *Press, New York, NY, pp 29–60*

890 **Harrison PJ** (1988) Determining phosphate uptake rates of phytoplankton. *In* *CS*
891 *Lobban, DJ Chapman and BP Kremer, eds, Experimental phycology: a laboratory*
892 *manual. Cambridge University Press, New York, NY, pp 186–195*

893 **Haslam RP, Hamilton ML, Economou CK, Smith R, Hassall KL, Napier JA,**

894 **Sayanova O** (2020) Overexpression of an endogenous type 2 diacylglycerol
895 acyltransferase in the marine diatom *Phaeodactylum tricornutum* enhances lipid
896 production and omega-3 long-chain polyunsaturated fatty acid content. *Biotechnol*
897 *Biofuels* **13**: 87

898 **He G, Zhu X, Elling AA, Chen L, Wang X, Guo L, Liang M, He H, Zhang H,**
899 **Chen F, et al.** (2010) Global epigenetic and transcriptional trends among two rice
900 subspecies and their reciprocal hybrids. *Plant Cell* **22**: 17–33

901 **Hoguin A, Rastogi A, Bowler C, Tirichine L** (2021) Genome-wide analysis of
902 allele-specific expression of genes in the model diatom *Phaeodactylum*
903 *tricornutum*. *Sci Rep* **11**: 2954

904 **Huang B, Marchand J, Blanckaert V, Lukomska E, Ulmann L, Wielgosz-Collin**
905 **G, Rabesaotra V, Moreau B, Bougaran G, Mimouni V, et al.** (2019) Nitrogen
906 and phosphorus limitations induce carbon partitioning and membrane lipid
907 remodelling in the marine diatom *Phaeodactylum tricornutum*. *Eur J Phycol* **54**:
908 342–358

909 **Kim HU, Lee KR, Go YS, Jung JH, Suh MC, Kim JB** (2011) Endoplasmic
910 reticulum-located PDAT1-2 from castor bean enhances hydroxy fatty acid
911 accumulation in transgenic plants. *Plant Cell Physiol* **52**: 983–993

912 **Kloosterman B, Abelenda JA, Gomez MDMC, Oortwijn M, de Boer JM,**
913 **Kowitwanich K, Horvath BM, van Eck HJ, Smaczniak C, Prat S, et al.** (2013)
914 Naturally occurring allele diversity allows potato cultivation in northern
915 latitudes. *Nature* **495**: 246-250

916 **Langmead B, Salzberg SL** (2012) Fast gapped-read alignment with Bowtie 2. *Nat*
917 *Methods* **9**: 357–359

918 **Levitan O, Dinamarca J, Zelzion E, Lun DS, Guerra LT, Kim MK, Kim J, Van**
919 **Mooy BAS, Bhattacharya D, Falkowski PG** (2015) Remodeling of intermediate
920 metabolism in the diatom *Phaeodactylum tricornutum* under nitrogen stress. *Proc*
921 *Natl Acad Sci USA* **112**: 412–417

922 **Liu J, Han D, Yoon K, Hu Q, Li Y** (2016a) Characterization of type 2 diacylglycerol
923 acyltransferases in *Chlamydomonas reinhardtii* reveals their distinct substrate
924 specificities and functions in triacylglycerol biosynthesis. *Plant J* **86**: 3–19

925 **Liu X, Hempel F, Stork S, Bolte K, Moog D, Heimerl T, Maier UG, Zauner S**
926 (2016b) Addressing various compartments of the diatom model organism
927 *Phaeodactylum tricoratum* via sub-cellular marker proteins. *Algal Res* **20**: 249–
928 257

929 **Liu XY, Ouyang LL, Zhou ZG** (2016c) Phospholipid: diacylglycerol acyltransferase
930 contributes to the conversion of membrane lipids into triacylglycerol in *Myrmecia*
931 *incisa* during the nitrogen starvation stress. *Sci Rep* **6**: 26610

932 **Lupette J, Jaussaud A, Seddiki K, Morabito C, Brugière S, Schaller H, Kuntz M,**
933 **Putaux JL, Jouneau PH, Rébeillé F, et al.** (2019) The architecture of lipid
934 droplets in the diatom *Phaeodactylum tricoratum*. *Algal Res* **38**: 101415

935 **Maheswari U, Jabbari K, Petit JL, Porcel BM, Allen AE, Cadoret JP, De**
936 **Martino A, Heijde M, Kaas R, Roche JL, et al.** (2010) Digital expression
937 profiling of novel diatom transcripts provides insight into their biological functions.
938 *Genome Biol* **11**: R85

939 **Mao X, Wu T, Kou Y, Shi Y, Zhang Y, Liu J** (2019) Characterization of type I and
940 type II diacylglycerol acyltransferases from the emerging model alga *Chlorella*
941 *zofingiensis* reveals their functional complementarity and engineering potential.
942 *Biotechnol Biofuels* **12**: 28

943 **Matthijs M, Fabris M, Broos S, Vyverman W, Goossens A** (2016) Profiling of the
944 early nitrogen stress response in the diatom *Phaeodactylum tricoratum* reveals a
945 novel family of RING-domain transcription factors. *Plant Physiol* **170**: 489–498

946 **McCarthy JK, Smith SR, McCrow JP, Tan M, Zheng H, Beerl K, Roth R, Lichtle**
947 **C, Goodenough U, Bowler CP, et al.** (2017) Nitrate reductase knockout uncouples
948 nitrate transport from nitrate assimilation and drives repartitioning of carbon flux in
949 a model pennate diatom. *Plant Cell* **29**: 2047–2070

950 **Mhaske V, Beldjilali K, Ohlrogge J, Pollard M** (2005) Isolation and
951 characterization of an *Arabidopsis thaliana* knockout line for
952 phospholipid:diacylglycerol transacylase gene (At5g13640). *Plant Physiol Biochem*
953 **43**: 413–417

954 **Mock T, Otilar RP, Strauss J, McMullan M, Paaajanen P, Schmutz J, Salamov A,**
955 **Sanges R, Toseland A, Ward BJ, et al.** (2017) Evolutionary genomics of the
956 cold-adapted diatom *Fragilariopsis cylindrus*. *Nature* **541**: 536–540

957 **Moog D, Rensing SA, Archibald JM, Maier UG, Ullrich KK** (2015) Localization
958 and evolution of putative triose phosphate translocators in the diatom
959 *Phaeodactylum tricorutum*. *Genome Biol Evol* **7(11)**: 2955–2969

960 **Niu Y-F, Zhang M-H, Li D-W, Yang W-D, Liu J-S, Bai W-B, Li H-Y** (2013)
961 Improvement of neutral lipid and polyunsaturated fatty acid biosynthesis by
962 overexpressing a type 2 diacylglycerol acyltransferase in marine diatom
963 *Phaeodactylum tricorutum*. *Mar Drugs* **11**: 4558–4569

964 **Nomaguchi T, Maeda Y, Yoshino T, Asahi T, Tirichine L, Bowler C, Tanaka T**
965 (2018) Homoeolog expression bias in allopolyploid oleaginous marine diatom
966 *Fistulifera solaris*. *BMC Genom* **19**: 330

967 **Nymark M, Sharma AK, Sparstad T, Bones A, Winge P** (2016) A CRISPR/Cas9
968 system adapted for gene editing in marine algae. *Sci Rep* **6**: 24951

969 **Oelkers P, Tinkelenberg A, Erdeniz N, Cromley D, Billheimer JT, Sturley SL**
970 (2000) A lecithin cholesterol acyltransferase-like gene mediates diacylglycerol
971 esterification in yeast. *J Biol Chem* **275**: 15609–15612

972 **Pan X, Peng FY, Weselake RJ** (2015) Genome-wide analysis of
973 *Phospholipid-diacylglycerol acyltransferase (PDAT)* genes in plants reveals the
974 eudicot-wide PDAT gene expansion and altered selective pressures acting on the
975 core eudicot PDAT paralogs. *Plant Physiol* **167**: 887–904

976 **Pan X, Siloto RMP, Wickramarathna AD, Mietkiewska E, Weselake RJ** (2013)
977 Identification of a pair of phospholipid:diacylglycerol acyltransferases from

978 developing flax (*Linum usitatissimum* L.) seed catalyzing the selective production
979 of trilinolenin. J Biol Chem **288**: 24173–24188

980 **Pastinen T** (2010) Genome-wide allele-specific analysis: insights into regulatory
981 variation. Nat Rev Genet **11**: 533–538

982 **Rastogi A, Murik O, Bowler C, Tirichine L** (2016) PhytoCRISP-Ex: a web-based
983 and stand-alone application to find specific target sequences for CRISPR/CAS
984 editing. BMC Bioinform **17**: 261

985 **Rastogi A, Vieira FRJ, Deton-Cabanillas A-F, Veluchamy A, Cantrel C, Wang G,**
986 **Vanormelingen P, Bowler C, Piganeau G, Hu H, et al.** (2020) A genomics
987 approach reveals the global genetic polymorphism, structure, and functional
988 diversity of ten accessions of the marine model diatom *Phaeodactylum*
989 *tricornutum*. ISME J **14**: 347–363

990 **Reiser S, Somerville C** (1997) Isolation of mutants of *Acinetobacter calcoaceticus*
991 deficient in wax ester synthesis and complementation of one mutation with a gene
992 encoding a fatty acyl coenzyme A reductase. J Bacteriol **179**: 2969–2975

993 **Shao L, Xing F, Xu C, Zhang Q, Che J, Wang X, Song J, Li X, Xiao J, Chen L-L,**
994 **et al.** (2019) Patterns of genome-wide allele-specific expression in hybrid rice and
995 the implications on the genetic basis of heterosis. Proc Natl Acad Sci USA **116**:
996 5653–5658

997 **Siaut M, Heijde M, Mangogna M, Montsant A, Coesel S, Allen A, Manfredonia A,**
998 **Falciatore A, Bowler C** (2007) Molecular toolbox for studying diatom biology in
999 *Phaeodactylum tricornutum*. Gene **406**: 23–35

1000 **Smith SR, Dupont CL, McCarthy JK, Broddrick JT, Oborník M, Horák A,**
1001 **Füßy Z, Cihlář J, Kleessen S, Zheng H, et al.** (2019) Evolution and regulation of
1002 nitrogen flux through compartmentalized metabolic networks in a marine diatom.
1003 Nat Commun **10**: 4552

1004 **Ståhl U, Carlsson AS, Lenman M, Dahlqvist A, Huang B, Banaś W, Banaś A,**
1005 **Stymne S** (2004) Cloning and functional characterization of a

1006 phospholipid:diacylglycerol acyltransferase from *Arabidopsis*. *Plant Physiol* **135**:
1007 1324–1335

1008 **Tanaka T, Maeda Y, Veluchamy A, Tanaka M, Abida H, Maréchal E, Bowler C,**
1009 **Muto M, Sunaga Y, Tanaka M, et al.** (2015) Oil accumulation by the oleaginous
1010 diatom *Fistulifera solaris* as revealed by the genome and transcriptome. *Plant Cell*
1011 **27**:162–76

1012 **Tao Z, Liu H, Qiu D, Zhou Y, Li X, Xu C, Wang S** (2009) A pair of allelic WRKY
1013 genes play opposite roles in rice-bacteria interactions. *Plant Physiol* **151**: 936–948

1014 **Terashima M, Specht M, Hippler M** (2011) The chloroplast proteome: a survey
1015 from the *Chlamydomonas reinhardtii* perspective with a focus on distinctive
1016 features. *Curr Genet* **57**: 151–168

1017 **Todesco M, Balasubramanian S, Hu TT, Traw MB, Horton M, Epple P, Kuhns C,**
1018 **Sureshkumar S, Schwartz C, Lanz C, et al.** (2010) Natural allelic variation
1019 underlying a major fitness trade-off in *Arabidopsis thaliana*. *Nature* **465**: 632–636

1020 **van Erp H, Bates PD, Burgal J, Shockey J, Browse J** (2011) Castor
1021 phospholipid:diacylglycerol acyltransferase facilitates efficient metabolism of
1022 hydroxy fatty acids in transgenic *Arabidopsis*. *Plant Physiol* **155**: 683–693

1023 **Winter A, Kramer W, Werner FA, Kollers S, Kata S, Durstewitz G, Buitkamp J,**
1024 **Womack JE, Thaller G, Fries R** (2002) Association of a lysine-232/alanine
1025 polymorphism in a bovine gene encoding acyl-CoA: diacylglycerol acyltransferase
1026 (*DGATI*) with variation at a quantitative trait locus for milk fat content. *Proc Natl*
1027 *Acad Sci USA* **99**: 9300–9305

1028 **Xu Y, Caldo KMP, Singer SD, Mietkiewska E, Greer MS, Tian B, Dyer JM,**
1029 **Smith M, Zhou X-R, Qiu X, et al.** (2021) *Physaria fendleri* and *Ricinus communis*
1030 lecithin: cholesterol acyltransferase-like phospholipases selectively cleave hydroxy
1031 acyl chains from phosphatidylcholine. *Plant J* **105**: 182–196

1032 **Yang J, Liu J, Pan Y, Maréchal E, Amato A, Liu M, Gong Y, Li Y, Hu H** (2022)
1033 PDAT regulates PE as transient carbon sink alternative to triacylglycerol in

1034 *Nannochloropsis*. Plant Physiol **189**: 1345–1362

1035 **Yoon K, Han D, Li Y, Sommerfeld M, Hu Q** (2012) Phospholipid:diacylglycerol
1036 acyltransferase is a multifunctional enzyme involved in membrane lipid turnover
1037 and degradation while synthesizing triacylglycerol in the unicellular green
1038 microalga *Chlamydomonas reinhardtii*. Plant Cell **24**: 3708–3724

1039 **Zhang C, Hu H** (2014) High-efficiency nuclear transformation of the diatom
1040 *Phaeodactylum tricorutum* by electroporation. Mar Genomics **16**: 63–66

1041 **Zhang M, Fan J, Taylor DC, Ohlrogge JB** (2009) DGAT1 and PDAT1
1042 acyltransferases have overlapping functions in *Arabidopsis* triacylglycerol
1043 biosynthesis and are essential for normal pollen and seed development. Plant Cell
1044 **21**: 3885–3901

1045 **Zhang X, Borevitz JO** (2009) Global analysis of allele-specific expression in
1046 *Arabidopsis thaliana*. Genetics **182**: 943–954

1047 **Zhang Y, Pan Y, Ding W, Hu H, Liu J** (2021) Lipid production is more than doubled
1048 by manipulating a diacylglycerol acyltransferase in algae. GCB Bioenergy **13**:
1049 185–200

1050 **Zhao Y, Huang J, Wang Z, Jing S, Wang Y, Ouyang Y, Cai B, Xin X-F, Liu X,**
1051 **Zhang C, et al.** (2016) Allelic diversity in an NLR gene *BPH9* enables rice to
1052 combat planthopper variation. Proc Natl Acad Sci USA **113**:12850-12855

1053 **Zheng L, Shockey J, Bian F, Chen G, Shan L, Li X, Wan S, Peng Z** (2017) Variant
1054 amino acid residues alter the enzyme activity of peanut type 2 diacylglycerol
1055 acyltransferases. Front Plant Sci **8**: 1751

1056 **Zheng P, Allen WB, Roesler K, Williams ME, Zhang S, Li J, Glassman K, Ranch**
1057 **J, Nubel D, Solawetz W, et al.** (2008) A phenylalanine in DGAT is a key
1058 determinant of oil content and composition in maize. Nat Genet **40**: 367–372















