



HAL
open science

Monogalactosyldiacylglycerol synthase isoforms play diverse roles inside and outside the diatom plastid

Nolwenn Guéguen, Yannick Sérès, Félix Cicéron, Valérie Gros, Grégory Si Larbi, Denis Falconet, Etienne Deragon, Siraba D Gueye, Damien Le Moigne, Marion Schilling, et al.

► To cite this version:

Nolwenn Guéguen, Yannick Sérès, Félix Cicéron, Valérie Gros, Grégory Si Larbi, et al.. Monogalactosyldiacylglycerol synthase isoforms play diverse roles inside and outside the diatom plastid. *The Plant cell*, 2024, 2024, pp.koae275. 10.1093/plcell/koae275 . hal-04731720

HAL Id: hal-04731720

<https://hal.science/hal-04731720v1>

Submitted on 11 Oct 2024

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 4.0 International License

1 **RESEARCH ARTICLE**2 **Monogalactosyldiacylglycerol synthase isoforms play diverse roles inside**
3 **and outside the diatom plastid**4
5 **Authors:**6 Nolwenn Guéguen^{1,§}, Yannick Sérès^{1,§}, Félix Cicéron^{1,¶}, Valérie Gros^{1,#}, Grégory Si Larbi¹,
7 Denis Falconet¹, Etienne Deragon¹, Siraba D. Gueye¹, Damien Le Moigne¹, Marion
8 Schilling¹, Mathilde Cussac¹, Dimitris Petroutsos^{1,2}, Hanhua Hu³, Yangmin Gong^{4,5}, Morgane
9 Michaud¹, Juliette Jouhet¹, Juliette Salvaing¹, Alberto Amato^{1,*}, Eric Maréchal^{1,*}.10 **Affiliations:**11 ¹Laboratoire de Physiologie Cellulaire et Végétale, Institut National de Recherche pour
12 l'Agriculture, l'Alimentation et l'Environnement, Centre National de la Recherche
13 Scientifique, Commissariat à l'Energie Atomique et aux Energies Alternatives, Université
14 Grenoble Alpes; IRIG, CEA-Grenoble, 17 rue des Martyrs; 38000 Grenoble, France15 ²Department of Organismal Biology, Uppsala University, 75236, Uppsala, Sweden16 ³Key Laboratory of Algal Biology, Institute of Hydrobiology, Chinese Academy of Sciences,
17 Wuhan 430072, China18 ⁴Oil Crops Research Institute of the Chinese Academy of Agricultural Sciences, Wuhan
19 430062, China20 ⁵Key Laboratory of Biology and Genetic Improvement of Oil Crops, Ministry of Agriculture,
21 Oil Crops Research Institute of Chinese Academy of Agricultural Sciences, Wuhan 430062,
22 China23 [¶]Present address: Living Systems Institute - Exeter University Stocker Road; EX4 4QD,
24 Exeter, UK25 [#]Present address: Institut des Biomolécules Max Mousseron; Pôle Chimie Balard Recherche,
26 1919 route de Mende, 34293 Montpellier27 [§]Equal contribution28 ***Corresponding authors:**29 Alberto Amato (alberto.amato@cea.fr) and Eric Maréchal (eric.marechal@cea.fr)© The Author(s) 2024. Published by Oxford University Press on behalf of American Society of Plant Biologists. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

30

31 **Short title:** Multigenic MGD family in diatoms

35

36 The authors responsible for distribution of materials integral to the findings presented in this
37 article in accordance with the policy described in the Instructions for Authors
38 (<https://academic.oup.com/plcell/pages/General-Instructions>) are Alberto Amato
39 (alberto.amato@cea.fr) and Eric Maréchal (eric.marechal@cea.fr).

ACCEPTED MANUSCRIPT

40 **Abstract.**

41 Diatoms derive from a secondary endosymbiosis event, which occurred when a eukaryotic
42 host cell engulfed a red alga. This led to the formation of a complex plastid enclosed by four
43 membranes: two innermost membranes originating from the red alga chloroplast envelope,
44 and two additional peri- and epiplastidial membranes (PPM, EpM). The EpM is linked to the
45 endoplasmic reticulum (ER). The most abundant membrane lipid in diatoms is
46 monogalactosyldiacylglycerol (MGDG), synthesized by galactosyltransferases called MGDG
47 synthases (MGDs), conserved in photosynthetic eukaryotes and considered to be specific to
48 chloroplast membranes. Similar to angiosperms, a multigenic family of MGDs has evolved in
49 diatoms, but through an independent process. We characterized MGD α , MGD β and MGD γ in
50 *Phaeodactylum tricornutum*, combining molecular analyses, heterologous expression in
51 *Saccharomyces cerevisiae*, and studying overexpressing and CRISPR-Cas9-edited lines.
52 MGD α localizes mainly to thylakoids, MGD β to the PPM, and MGD γ to the ER and EpM.
53 MGDs have distinct specificities for diacylglycerol, consistent with their localization. Results
54 suggest that MGD α is required for thylakoid expansion under optimal conditions, while
55 MGD β and MGD γ play roles in plastid and non-plastid membranes and in response to
56 environmental stress. Functional compensation among MGDs likely contributes to diatom
57 resilience under adverse conditions and to their ecological success.

58

ACCEPTED MANUSCRIPT

59 **Introduction.**

60 Diatoms form a monophyletic phylum of photosynthetic eukaryotes, representing one
61 of the largest groups of stramenopiles (de Vargas et al., 2015; Serôdio and Lavaud, 2020).
62 They dominate phytoplankton biodiversity in oceans and freshwater ecosystems (Mann and
63 Droop, 1996; Mann, 1999; Kooistra et al., 2007; Guiry, 2012; de Vargas et al., 2015). Our
64 understanding of their cellular functional organization relies mainly on the model species
65 *Phaeodactylum tricorutum*, for which intensive efforts have been made to sequence its
66 genome (Bowler et al., 2008), develop molecular tools and resources (Siaut et al., 2007; De
67 Riso et al., 2009; Daboussi et al., 2014; Kroth et al., 2018), and characterize membrane
68 glycerolipids at whole-cell level (Abida et al., 2015) and in organelle-enriched fractions
69 (Lupette et al., 2019; Huang et al., 2024).

70 They have a cell architecture comprising membrane systems with no equivalent in
71 ‘simpler’ photosynthetic models, like *Chlamydomonas reinhardtii* or *Arabidopsis*
72 (*Arabidopsis thaliana*), containing primary plastids surrounded by only two membranes. This
73 complex membrane compartmentation derives from their evolutionary origin following two
74 endosymbiosis events. An initial event at the origin of most photosynthetic eukaryotes
75 involved the engulfment of a cyanobacterium by an unknown heterotrophic cell, 1-1.5 b.y.a.
76 (Jensen and Leister, 2014; Marechal, 2024). This primary endosymbiosis led to the formation
77 of the classical chloroplast, bounded by a two-membrane envelope (Sato and Awai, 2017;
78 Marechal, 2024). Lineages radiated into chlorophytes (green algae), rhodophytes (red algae)
79 and glaucophytes. Emergence of diatoms has occurred more recently, ~180-250 m.y.a.
80 (Kooistra et al., 2003; Sorhannus, 2007), stemming from the engulfment of a red algal
81 symbiont by a secondary eukaryotic host cell (Reyes-Prieto et al., 2007; Bowler et al., 2008;
82 Benoiston et al., 2017).

83 Diatom photosynthetic organelle derives vertically from the symbiont cell, reduced in
84 the form of a ‘complex’ or ‘secondary plastid’, bounded by four membranes (Figure 1). The
85 two innermost membranes, *i.e.* the ‘inner’ and ‘outer envelope membranes’ (iEM and oEM,
86 respectively; Figure 1) correspond to the chloroplast envelope of the engulfed alga. The
87 outermost one, named the ‘epiplastidial membrane’ (EpM, Figure 1), is thought to derive
88 from the host cell phagosome. It is continuous with the outer membrane of the nuclear
89 envelope (oNE) and is connected to the endoplasmic reticulum (ER) (Murakami and
90 Hashimoto, 2009; Tanaka et al., 2015; Flori et al., 2016). Underneath, the ‘periplastidial
91 membrane’ (PPM, Figure 1) is usually considered to derive from the symbiont plasma

92 membrane (Grosche et al., 2014). An alternative hypothesis is that the PPM could derive from
93 the host ER, like the EpM (Gould et al., 2015). A ‘blob-like’ structure, or in short a ‘blob’,
94 containing a vesicular network (VN) visible in confocal and transmission electron microscopy
95 (Figure 1) has been detected between the PPM and the oEM, but its organization and function
96 remain elusive (Kilian and Kroth, 2005; Flori et al., 2016). Thomas Cavalier-Smith postulated
97 that the blob corresponds to a relic of the symbiont ER (Cavalier-Smith, 2018). The EpM,
98 PPM and blob are therefore ontogenetically related to the endomembrane system. To picture
99 diatoms’ sophisticated plastid architecture completely, membrane contact sites occur between
100 the inner membrane of the nuclear envelope (iNE) and the PPM (Flori et al., 2016) (Figure 1).

101 It is challenging to identify the protein and lipid components of the iEM, oEM, PPM
102 and EpM, and to unravel the biogenesis, dynamics and function of each of these membranes,
103 since little knowledge can be transferred from simple eukaryotic models. Concerning proteins,
104 the import of some nuclear-encoded precursors was dissected molecularly in *P. tricornutum*
105 and *Thalassiosira pseudonana*. Plastid protein precursors contain a bipartite N-terminal
106 targeting peptide, with a signal peptide (Sp) fused to a chloroplast transit peptide-like
107 sequence (Ctp), and harbor a conserved amino acid motif called ‘ASAFAP’ (Apt et al., 2002;
108 Kilian and Kroth, 2005; Gruber et al., 2007; Gruber et al., 2015). Based on the detection of
109 Sp-Ctp sequences and ASAFAP motifs, using the two predictions tools ASAFind (Gruber et
110 al., 2015) and HECTAR (Gschloessl et al., 2008) only 1,471 and 561 proteins, respectively,
111 were considered chloroplast-targeted among the >12,000 nuclear-encoded proteins in *P.*
112 *tricornutum*. These predictions are likely underestimated. By contrast, our recent proteomic
113 study of *P. tricornutum* purified plastids, allowed the detection of 2,758 nuclear-encoded and
114 92 plastid-encoded proteins (Huang et al., 2024), possibly including contaminants, and
115 requesting experimental validation by expressing protein fusions with fluorescent reporters.

116 Concerning lipids, the composition of the different plastid membranes and the precise
117 sites of their biosynthesis are unknown. Membrane glycerolipids contain two fatty acids (FA)
118 esterified at positions *sn-1* and *sn-2* of a glycerol backbone, and harbor a polar head group at
119 position *sn-3*. The nature of the head group defines glycerolipid ‘classes’, while FA chain-
120 lengths and unsaturation levels define ‘molecular species’ (Li-Beisson et al., 2010; Petroustos
121 et al., 2014). It was considered plausible but not demonstrated, that the iEM and oEM may
122 contain the lipid classes conserved in cyanobacteria and primary chloroplast, *i.e.* two
123 galactolipids, monogalactosyl- and digalactosyldiacylglycerol (MGDG and DGDG), a
124 sulfolipid, sulfoquinovosyldiacylglycerol (SQDG) and a phospholipid, phosphatidylglycerol

125 (PG) (Boudiere et al., 2014; Petroutsos et al., 2014). The current consensus is that this lipid
126 quartet is a marker of plastids in eukaryotes (Boudiere et al., 2014; Petroutsos et al., 2014).
127 The EpM and PPM were hypothesized to have an ER/endomembrane-type lipid composition,
128 with phospholipids, *e.g.* phosphatidylcholine (PC) and phosphatidylethanolamine (PE), and
129 betaine lipids, *e.g.* diacylglyceryl hydroxymethyltrimethyl- β -alanine (DGTA) (Dolch and
130 Marechal, 2015). Extraplasmidial membranes were speculated to be either devoid of
131 galactolipids and sulfolipids or to contain very low levels of these lipid classes. Contradicting
132 this view, a quantitative analysis of glycerolipids between whole cells and plastid-enriched
133 sub-fractions showed similar compositions (Huang et al., 2024), suggesting that MGDG,
134 DGDG or SQDG presence could expand to the PPM, EpM and to non-plastid membranes in
135 diatom cells.

136 Here we addressed the committing step of galactolipid synthesis, *i.e.* the production of
137 MGDG by monogalactosyldiacylglycerol synthase (MGD) enzymes. MGDs catalyze the
138 transfer of a galactosyl residue from a uridine-diphosphate-galactose (UDP-Gal) to the *sn*-3
139 position of a diacylglycerol (DAG):



141 In *P. tricornutum*, MGDG is the most abundant membrane lipid in whole extracts,
142 representing one third of all lipid classes, supporting its presence at least in thylakoids, the
143 largest membrane system (Abida et al., 2015). It is the main platform for the production of
144 highly unsaturated 16-carbon FAs deriving from plastid palmitoyl-ACP (16:0-ACP), in
145 particular 16:2, 16:3 and 16:4 (Dolch and Marechal, 2015; Smith et al., 2021), mostly
146 esterified at its *sn*-2 position (Abida et al., 2015). *P. tricornutum* MGDG molecular species
147 are also enriched in eicosapentaenoic acid (20:5), a very-long-chain polyunsaturated FA.
148 Biosynthesis of 20:5 follows a sequence of acyl-chain-elongations and desaturations, likely
149 occurring at the ER (Moog et al., 2011; Dolch and Marechal, 2015; Dolch et al., 2017a; Billey
150 et al., 2021b; Smith et al., 2021) and/or at the EpM (Huang et al., 2024), on acyl chains
151 esterified on phospholipids and betaine lipids. The route of 20:5 from its biosynthesis site in
152 the ER/EpM to inner compartments of the plastid, tentatively called the “ Ω -pathway”
153 (Petroutsos et al., 2014; Dolch and Marechal, 2015; Dolch et al., 2017a) is unresolved.

154 Three MGDs have been predicted in *P. tricornutum* based on sequence similarity
155 (Petroutsos et al., 2014; Hori et al., 2016; Shang et al., 2022). Since *MGD* genes were
156 previously assigned confusing numbering schemes, the three gene loci *Phatr3_J14125* (on

157 chromosome 13), *Phatr3_J43938* (on chromosome 3), and *Phatr3_J9619* (on chromosome 1),
158 (Bowler et al., 2008) were renamed here *MGD α* , *MGD β* and *MGD γ* , respectively. MGD
159 proteins were detected in *P. tricornutum* proteome analyses (Lupette et al., 2022; Huang et al.,
160 2024), validating their expression. It would be tempting to assume that multiple isoforms of
161 MGDs may be important for lipid production in the different membrane compartments of the
162 secondary plastid. However, other lineages in stramenopiles, like eustigmatophytes, harbor
163 also a four-membrane plastid but contain only one *MGD* in their genome (Vieler et al., 2012;
164 Corteggiani Carpinelli et al., 2014). Thus, the presence of three isoforms in diatoms is
165 intriguing.

166 In this study, we characterized the three MGD isoforms in *P. tricornutum*, addressing
167 the question of their subcellular localization, nature of DAG substrates, involvement in the
168 different MGDG and DGDG species productions, as well as their putative role under
169 environmental stress, taking nitrogen starvation as model stress condition.

170

171

172

ACCEPTED MANUSCRIPT

173 **Results.**

174 **Gene structure and molecular phylogeny of *P. tricornutum* MGD α , MGD β and MGD γ**

175 MGD sequences were recovered from genomic data of *P. tricornutum* CCAP 1055/1
 176 via the EnsemblProtist web portal (Yates et al., 2022) and compared to entries in NCBI and
 177 Uniprot databases (Supplemental Table 1). The *Phatr3_J14125/Phatr3_EG02525* gene was
 178 previously numbered as either *MGD1* (Bullmann et al., 2010) or *MGD3* (Dolch et al., 2017b;
 179 Shang et al., 2022). To resolve this ambiguity we named genes with Greek alphabet letters.
 180 The presence and length of introns was confirmed by cDNA sequencing and curated
 181 manually. The structure of MGD α , MGD β and MGD γ coding regions, with 1,615, 1,805 and
 182 2,099 bp lengths respectively, is summarized in Supplemental Figure 1.

183 Molecular phylogeny of MGDs has previously shown that angiosperm and diatom
 184 sequences fall in distinct groups (Hori et al., 2016; Shang et al., 2022). We confirmed this
 185 finding with a midpoint rooted molecular phylogeny, updated with sequences of
 186 stramenopiles and other secondary endosymbionts available from public databases
 187 (Supplemental Table 2; Supplemental Figure 2). Angiosperm MGD encoded sequences form
 188 two clusters, corresponding to the two types identified in *Arabidopsis thaliana*, namely types
 189 A and B, reflecting a gene duplication and functional differentiation, at the base of flowering
 190 plants' radiation (Awai et al., 2001). The MGD sequences encoded in bryophyta and
 191 lycopodiophyta cluster with type A, highlighting that this type is the ancestral isoform in
 192 plants.

193 MGDs from stramenopiles are close to rhodophyta (Supplemental Figure 2), reflecting
 194 their origin from the red algal symbiont. In eustigmatophytes, like *Microchloropsis gaditana*
 195 (Fawley et al., 2015), a unique MGD sequence is detected, clustering with diatom MGD α ,
 196 along with pelagophyceae and phaeophyceae sequences, in a sister clade to the alveolates.
 197 This supports the conservation of an ancestral type corresponding to MGD α . The multigenic
 198 family of diatom MGDs highlights two additional clusters corresponding to β and γ , which
 199 have likely stemmed from gene duplication at the base of diatoms' radiation (Supplemental
 200 Figure 2).

201 **MGD α , MGD β and MGD γ predicted targeting sequences and protein structures**

202 *P. tricornutum* protein sequences were analyzed using *Arabidopsis thaliana* MGDs as
 203 a reference (Supplemental Figure 3 and 4). In *Arabidopsis*, AtMGD1, AtMGD2 and AtMGD3
 204 have distinct subcellular localizations and roles (Awai et al., 2001). AtMGD1 belongs to 'type

205 A', defined by the presence of an N-terminal Ctp that targets its precursor to the iEM of the
206 chloroplast (Awai et al., 2001). It is the major isoform producing MGDG for thylakoid
207 membranes and its knockout (KO) mutation is lethal in photoautotrophic conditions
208 (Kobayashi et al., 2007a, b). AtMGD2 and AtMGD3 belong to 'type B', lacking any N-
209 terminal Ctp. They are dispensable under optimal conditions (Kobayashi et al., 2009a). Type
210 B MGDs localize to the oEM of chloroplasts (Awai et al., 2001) and are crucial for the
211 adaptation of plant cells to nutrient stress (Kobayashi et al., 2009a). A recent work also shows
212 a localization of AtMGD2 in the cytosol of pollen tube (Billey et al., 2021a). This suggested
213 that angiosperm MGDs might have also some unknown function outside of plastids.

214 In *P. tricornutum*, nuclear-encoded sequences harboring a Sp-Ctp bipartite pre-
215 sequence and an ASAFAP motif were shown to go across the four limiting membranes of the
216 plastid (Apt et al., 2002). ASAFAP-dependent targeting can be tuned by changing some
217 residues, leading to the arrest of pre-proteins after crossing the EpM, at the level of the PPM
218 and in the so-called periplastidial compartment (PPC), locally forming a blob-like structure
219 (Apt et al., 2002; Kilian and Kroth, 2005).

220 In MGD α protein sequence, a Sp-Ctp could be detected using SignalP (Emanuelsson
221 et al., 2007; Almagro Armenteros et al., 2019) and WoLF PSORT (Horton et al., 2007)
222 software. This prediction was confirmed using the HECTAR method designed for
223 stramenopiles (Gschloessl et al., 2008). An ASAFAP-like motif was tentatively identified in
224 MGD α Sp-Ctp, with a 'SAAFSP' sequence (Supplemental Figure 3). In MGD β , a Sp was
225 detected using SignalP and HECTAR; a Ctp could only be predicted by ChloroP and WoLF
226 PSORT tools, but not by HECTAR. A low similarity ASAFAP-like motif was tentatively
227 identified with a 'GSGFVL' sequence (Supplemental Figure 3). In MGD γ , no prediction of a
228 bipartite pre-sequence could be obtained with any of these tools, confirming a recent analysis
229 of this isoform (Shang et al., 2022).

230 Specific domains involved in the binding to the UDP-Gal sugar donor, the DAG
231 acceptor, and the lipid activator phosphatidylglycerol (PG) were characterized structurally and
232 functionally in AtMGD1 (Rocha et al., 2016). Supplemental Figure 3 shows the sequence
233 alignment of the *A. thaliana* and *P. tricornutum* MGDs and the conservation of these domains
234 and key residues. Two residues are particularly critical for catalysis in AtMGD1 (Botte et al.,
235 2005; Dubots et al., 2010; Rocha et al., 2016). Firstly, a histidine in position 159 (H159) of
236 the alignment (Supplemental Figure 3) is involved in the deprotonation of the nucleophile
237 hydroxyl (OH) group of DAG. Secondly, a lysine (K576) binds to the sugar donor, stabilizing

238 the phosphates of the departing UDP during the reaction. The presence of anionic lipids such
239 as PA, PG and SQDG is critical for AtMGD1 activity. A PG-H catalytic dyad was proposed
240 to form an acid/base relay system facilitating the deprotonation of the acceptor substrate by
241 H159 (Makshakova et al., 2020; Nitenberg et al., 2020). It was further shown that PG
242 develops interactions with H159 and the neighboring arginine (R160), thus placing this lipid
243 close to the active site (Makshakova et al., 2020). R160 is also conserved in all *P. tricornutum*
244 sequences (Supplemental Figure 3).

245 MGDs belong to the Glycosyltransferase B (GT-B) superfamily (Rocha et al., 2016).
246 The conserved GT-B fold consists of two Rossmann-type $\alpha/\beta/\alpha$ domains of comparable size,
247 named the N- and C-domains, separated by a large cleft, and stabilized by two long C-
248 terminal α -helices. The N-domain of AtMGD1 corresponds to positions 141-338 and 655-
249 672, the C-domain to positions 362-654. They are connected by a loop at positions 339-361
250 (Supplemental Figure 3). Secondary and tertiary structures of *P. tricornutum* MGDs were
251 predicted using Phyre² (Kelley et al., 2015), based on the 3D-structure generated from the
252 crystal of AtMGD1 in complex with UDP (template 4X1T) (Rocha et al., 2016) and
253 AlphaFold (Jumper et al., 2021; Jumper and Hassabis, 2022; Mirdita et al., 2022). All *P.*
254 *tricornutum* MGDs, starting approximately at position 145 of the alignment in Supplemental
255 Figure 3 and ending around position 670, are typical GT-B fold structures, highly similar to
256 AtMGD1 (Supplemental Figure 4B-E). The MGD α C-domain is approximately the same size
257 as that of AtMGD1 (~170 amino acids), whereas it is longer in MGD β (~223 amino acids)
258 and MGD γ (~290 amino acids). The hinge connecting N- and C-domains is longer in all *P.*
259 *tricornutum* MGDs, possibly providing some flexibility in the structure. Additional segments
260 in the C-domain of MGD β and MGD γ , at the level of an unresolved loop in AtMGD1
261 structure, were predicted with a lower confidence level (50% and 70%, respectively) than the
262 rest of the structure (100%). The last amino acids of the C-terminal stabilizing helix were also
263 modeled with a lower confidence for MGD α and MGD γ (50%). Overall, these analyses show
264 the prediction of the canonical double-Rossmann fold structure defined in AtMGD1 in all *P.*
265 *tricornutum* MGDs, with conserved domains and residues that are critical for enzymatic
266 activity, and highlight some additional features that may provide specific functionalities.
267 These models have allowed designing guide RNA (gRNA) sequences in structured and
268 conserved regions, or within the enzymes active sites for the CRISPR-Cas9 editing
269 experiments.

270 **Functional analysis of MGDG synthase activity by heterologous expression in yeast.**

271 Sequences encoding the three mature MGDs were cloned under the control of the
272 galactose-inducible promoter pGAL1 and introduced into *Saccharomyces cerevisiae* for
273 heterologous expression. A transformed yeast expressing AtMGD2 was used as a positive
274 control. The lipid profile of three independent expressing lines were analyzed per transgene.
275 The expression of all *P. tricornutum* MGD isoforms led to the biosynthesis of MGDG
276 (Supplemental Figure 5), a lipid class absent in yeast. It is important to note that DAG
277 substrates produced in yeast are distinct from those found in *P. tricornutum*. In particular,
278 yeast lacks 20:5, which is abundant in diatom MGDG. Therefore, the observed MGDG
279 profiles do not reflect the enzymatic specificity *per se*, and lower MGDG level may simply
280 reflect a lack of favored DAG substrate for the considered enzyme. Nevertheless, the pattern
281 of MGDG species produced by the different lines highlighted some divergence among
282 isoforms. MGD α -expressing yeast showed the most distinct MGDG pattern compared to
283 yeast expressing AtMGD2, with higher proportions of MGDG-36:1 and MGDG-36:2
284 compared to any other line. MGD β -expressing yeast showed higher proportions of MGDG-
285 32:1 and 32:2, while no traces of MGDG 36:1 and MGDG-36:2 could be detected. Finally,
286 MGD γ was the only isoform able to produce MGDG-28:0 and MGDG-30:0, *i.e.* with
287 saturated acyl species. Although caution is needed when interpreting heterologous expression,
288 results presented here support different substrate specificities in the three MGD isoforms of *P.*
289 *tricornutum*.

290 ***Phaeodactylum tricornutum* MGD α , MGD β and MGD γ subcellular localization.**

291 *P. tricornutum* lines overexpressing MGD genes fused to *eGFP* were generated. For
292 each isoform, two independent overexpressing strains were analyzed and gave identical
293 results, namely MGD α -eGFP-A and -B, MGD β -eGFP-A and -B, and MGD γ -eGFP-A and -B.
294 Identical GFP signals were also obtained with a third series of overexpressing lines, used for
295 co-localization studies.

296 We first performed a immunoblot analysis on total protein extracts from each strain
297 using anti-GFP-HRP (horseradish peroxidase coupled) antibodies to assess the size and
298 expression level of each construct (Supplemental Figure 6). The molecular weight expected
299 for each MGD are 51, 61, and 66 kDa for MGD α , β and γ , respectively, therefore, we
300 expected to detect proteins around 78, 89 and 93 kDa for MGD α -, MGD β - and MGD γ -eGFP,
301 respectively. Immunoblot analyses revealed a band at \sim 70 and \sim 80 kDa for MGD α - and
302 MGD β -eGFP, respectively, which are in the expected range (Supplemental Figure 6),

303 coherent with the cleavage of a targeting pre-sequence in these proteins. A broad signal was
304 detected for MGD γ -eGFP, around 90-100 kDa. These analyses indicate that the transformed
305 lines expressed MGD-eGFP fusions, and could be used for subcellular localization studies as
306 well as phenotypic analyses of MGD overexpression. Overexpressing lines highlighted
307 distinct subcellular localizations of MGD isoforms (Figure 2).

308 MGD α -eGFP showed an intraplastidial signal visible as dots inside the plastid (Figure
309 2, Supplemental Figure 7A), corresponding to a localization in thylakoids (Huang et al.,
310 2024). This localization is consistent with the predicted bipartite Sp-Ctp, assuming that the
311 detected SAAFSP sequence has the same targeting function as the canonical ASAFAP motif
312 (Apt et al., 2002). By contrast, MGD β -eGFP fluorescence was detected in a single spot
313 corresponding to the blob-like structure (Figure 2, Supplemental Figure 7B). A membrane
314 vesicular network (VN) was shown to protrude inside the blob from the PPM but not the oEM
315 (Flori et al., 2016). Therefore, MGD β localization might correspond to the VN and extend to
316 the PPM. This localization is consistent with the predicted Sp-Ctp, and suggests that the
317 GSGFVL motif may be a specific alteration of the ASAFAP motif leading to a retention in
318 the blob and PPM (Apt et al., 2002). MGD γ -eGFP signal appeared in a reticulated membrane
319 structure corresponding to the ER, and at the periphery of the plastid (EpM) (Figure 2,
320 Supplemental Figure 7C). This localization is consistent with the absence of any predicted
321 targeting sequence in MGD γ .

322 We sought to confirm the observed localization of the three isoforms. We co-
323 expressed MGD α -eGFP with a gene coding for the photosynthetic electron transport C
324 subunit (PetC), addressed to thylakoids (Liu et al., 2016), fused to a monomeric red
325 fluorescent protein (mRFP). MGD α -eGFP and PetC-mRFP signals showed an overlapping
326 localization in thylakoids, with slightly different distribution patterns (Supplemental Figure
327 8A). We co-expressed MGD β -eGFP with the symbiont Derlin1-2 (sDer1-2), a component of
328 the protein translocation channel located at the PPM (Hempel et al., 2009), fused to mRFP.
329 The MGD β -eGFP and sDer1-2-mRFP signals showed identical localization in the PPM,
330 concentrated at the level of the blob-like structure (Supplemental Figure 8B). Finally, we co-
331 expressed MGD γ -eGFP with the mRFP, highlighting the difference between the MGD γ
332 localization at the periphery of the plastid and in the ER network and the soluble form of
333 mRFP in the cytosol (Supplemental Figure 8C).

334 The MGD β -eGFP signal always appeared at a location corresponding to the center of
335 plastids (Figure 2 and 3A, Supplemental Figures 7 and 8). The blob-associated fluorescence

336 was also observed in dividing plastids, which could suggest a role of MGD β during plastid
337 division. In dividing plastids observed by confocal microscopy, the blob was localized at the
338 interface between the two forming plastids (Figure 3A, Supplemental Figure 7B). This
339 observation was supported by scanning transmission electron microscopy (STEM)
340 observation of the blob VN in a zone connecting dividing plastids (Figure 3). After separation
341 of the daughter cells, blobs appeared facing each other. When MGD β -eGFP was observed in a
342 non-dividing plastid, the signal was much more diffuse. The blob-like structure may either
343 become more exposed as plastids undergo divisions, making the signal more visible, or it may
344 become more compacted during this particular time. Thus, it seems that the blob may play an
345 important role during cytokinesis, possibly as a glycerolipid synthesis machinery, and that this
346 machinery might be scattered in the PPM in non-dividing cells.

347 **Generation of MGD α , MGD β and MGD γ edited lines.**

348 We generated KO lines by CRISPR-Cas9 editing, using single guide RNA sequences
349 (sgRNA) selected to interrupt protein function. Since random genomic insertions and off-
350 target mutations can lead to unrelated phenotypes, we selected several independent strains for
351 each gene, obtained from different sgRNAs (distinct potential off-targets), and not originating
352 from the same initial transformed colony (distinct vector insertions in the genome) (Primers
353 and sgRNAs are shown in Supplemental Table 3).

354 For MGD α , we selected three pure KO lines with frameshift disruptions *i.e.* one KO
355 obtained with sgRNA 125i (*mgd α i1*), and two independent KO lines obtained with sgRNA
356 125j (*mgd α j1* and *mgd α j3*) (Supplemental Figure 9). Two *mgd α* mutants conserved for at
357 least one year highlighted homozygous profiles and were stable after repeated rounds of
358 cultivation in liquid medium (Supplemental Figure 10).

359 For MGD β , we selected one KO line with sgRNA 168b (*mgd β b1*), and three mutants
360 obtained with sgRNA 168c (*mgd β c1*, *mgd β c2* and *mgd β c3*) (Supplemental Figure 9). The
361 three lines *mgd β c1*, *mgd β c2* and *mgd β c3* derive from the same initially transformed cell. They
362 harbor the same vector insertion, but different editions at the targeted locus. The *mgd β c2*
363 mutant has a deletion of 6 nucleotides (DEL6), which does not generate a frameshift in the
364 sequence (Supplemental Figure 9). This mutant lacks two amino acids between the predicted
365 signal peptide and the first N-terminal conserved region. This deletion was not expected to
366 impair the enzyme function and thus *mgd β c2* was kept as a possible control for *mgd β c1* and

367 *mgdβc3* KO lines. All *mgdβ* mutants highlighted homozygous profiles and were stable after
368 repeated rounds of cultivation in liquid medium (Supplemental Figure 10).

369 For *MGDγ*, we selected two KO lines obtained with sgRNA 619h (*mgdγh1* and
370 *mgdγh2*) deriving from the same initially transformed cell, and one obtained with sgRNA 619i
371 (*mgdγi1*) (Supplemental Figure 9). For this gene, the generation of mutant colonies was
372 difficult, as transformed lines were greatly affected in their ability to form colonies on plates.
373 Only one mosaic colony was obtained after transformations with 619h and 619i sgRNAs. We
374 managed to isolate three strains from the mosaic colony obtained with 619h sgRNA. For the
375 mosaic colony obtained with 619i sgRNA, the mutant profile corresponds to a mixed
376 population, without any WT; all indels detected led to either a frameshift in the sequence or
377 the deletion of crucial amino acids in the active site region (Supplemental Figure 9). We used
378 this third heterogeneous KO line, without further purification. The growth phenotype
379 observed on plates may indicate some important role in stressful environmental conditions. In
380 addition, all *mgdγ* mutants evolved, as observed after one year of repeated rounds of
381 cultivation in liquid medium (Supplemental Figure 10), without any trace of WT *MGDγ*
382 sequence, and in all cases in the direction of larger deletions and functional loss. They were
383 therefore considered as valid KO lines.

384 **Impact of MGD knockout and overexpression on *Phaeodactylum* growth and** 385 **photosynthesis efficiency in optimal growth conditions.**

386 A first key observation is that none of the *MGD* isoform KO was lethal. The growth of
387 each KO line was monitored in optimal condition, in a rich medium (Supplemental Figure
388 11A). Most of the KO mutants had a slightly slower growth than the WT strain.
389 Overexpression of *MGDα* and *MGDγ* but not *MGDβ* led to a slight increase in growth
390 (Supplemental Figure 11B). The moderate impact of *MGD* KO on growth in liquid medium
391 suggests possible mutual compensating effects between isoforms.

392 As *MGDG* plays an important role in the stabilization and function of photosystems,
393 and in the xanthophyll cycle (Azadi-Chegeni et al., 2022; Garab et al., 2022), we sought
394 whether any effect on photosynthesis and/or photoprotection could be detected. We measured
395 the effective photochemical quantum yield of photosystem II (Ψ_{II}) and non-photochemical
396 quenching (NPQ) during two different kinds of light stress: a high light ($700 \mu\text{mol photons}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$)
397 and a moderate light stress (2-steps increase in light intensity, at 55 then 335
398 $\mu\text{mol}\cdot\text{photons m}^{-2}\cdot\text{s}^{-1}$) (Supplemental Figure 12). Ψ_{II} recovery and NPQ relaxation were

399 monitored at low light intensity ($20 \mu\text{mol photons.m}^{-2}.\text{s}^{-1}$) (Ruban et al., 2004; Goss and
400 Jakob, 2010). Under all tested conditions, no significant changes in photosynthetic efficiency
401 or non-photochemical quenching (NPQ) were observed, except in the case of *mgdgy1*. This
402 line showed lower photosynthetic efficiency under 55 and 335 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ and
403 reduced recovery under low light intensity following high light intensity measurements.
404 Additionally, *mgdgy1* exhibited slower relaxation of NPQ under the same low light
405 conditions (Supplemental Figure 12C and F).

406 **Impact of MGD knockout and overexpression on *Phaeodactylum* cell morphology and** 407 **plastid architecture.**

408 Cell morphology of MGD KO and overexpressing lines was examined by light and
409 epifluorescence microscopy (Supplemental Figure 13). All mutant cells were fusiform with a
410 normally sized plastid compared to the WT. No change in cell size was detected. Potential
411 impact of MGD KO on plastid membranes was checked by STEM (Figure 4), with cells from
412 *mgd α 3*, *mgd β 3* and *mgd γ 2* strains. No modifications could be observed in plastids nor in
413 any other parts of the cell in any MGD KO line. In MGD β null mutants, blob-like structures
414 were still present (Figure 4). Likewise, no change in cell ultrastructure could be noticed in
415 overexpressing lines (Supplemental Figure 14).

416 **Impact of MGD knockout and overexpression on *Phaeodactylum* total fatty acid and** 417 **glycerolipid class profiles.**

418 We cultured WT, KO and overexpressing lines in parallel in a nutrient-replete medium
419 (10N10P). The total amount of fatty acids (FA) in all strains was similar to the WT analyzed
420 in parallel (Supplemental Figure 15A and 16A) and FA distributions were similar to those
421 already reported in the literature in the same conditions (Abida et al., 2015). As observed
422 previously, *P. tricornutum* is rich in 20:5 (about 25-30 % of total FAs) and C16 molecular
423 species (mainly 16:0, 16:1, 16:3), while poor in C18 FA. The distribution of the different lipid
424 classes was analyzed by LC-MS/MS (Supplemental Figure 15B and 16B). In all strains, the
425 glycerolipid profiles were dominated by MGDG, PC and SQDG, which accounted for about
426 70 % of total glycerolipid content, as previously reported (Abida et al., 2015).

427 Only minor changes were observed in the relative proportion of membrane lipid
428 classes (Supplemental Figure 15B). Most importantly, MGD KO had no significant impact on
429 the levels of galactolipids. Concerning the other lipid classes, MGD α KO strains showed
430 slightly lower levels of SQDG (16.0% vs 19.6% in the WT) and more PC (17.8% vs 14.10%

431 in the WT) (Supplemental Figure 15B). In MGD β KO strains, the proportion of PC was
432 higher (14.10%) compared to both WT and silent *mgd β c2* controls (12.7%) (Supplemental
433 Figure 15B), consistently with an opposite trend in the MGD β overexpressing strains
434 (Supplemental Figure 16B). All MGD γ KO strains had slightly lower proportions of PG
435 (5.1% vs 6.3% in the WT) and slightly higher proportions of DAG (1.23% vs 1.05% in the
436 WT) (Supplemental Figure 15B). DAG tended to decrease in MGD γ overexpressing strains
437 (Supplemental Figure 16B), in line with a more active utilization of this MGD substrate.

438 **Impact of MGD KO and overexpression on acyl profiles in each glycerolipid class.**

439 Since KO or overexpression of MGD isoforms had minor impacts on cell division,
440 photosynthesis, morphology, total FAs and glycerolipid class profiles, one may wonder
441 whether the three isoforms might be fully interchangeable. Expression of the different MGD
442 isoforms in yeast suggested that they have distinct substrate specificities (Supplemental
443 Figure 5). We looked therefore in more detail at the variation of each lipid molecular species
444 in each lipid class (Abida et al., 2015; Jouhet et al., 2017).

445 We based our analyses on the following assumptions: (1) a decrease in a molecular
446 species should reflect the KO of the considered MGD isoform, and (2) an opposite trend
447 should be observed in the corresponding overexpressing lines (shown in Supplemental
448 Figures 17, 18). An additional principle was to assume that (3) an increase of a molecular
449 species in a KO line might be determined by a compensation by the other intact isoforms.
450 Finally, (4) we focused our analyses on the four glycerolipids known to be conserved in
451 photosynthetic membranes, *i.e.* MGDG, DGDG, SQDG and PG (Boudiere et al., 2014)
452 (Figures 5-7), and on other phospho- and betaine lipids, supposed to be related to the
453 endomembrane system, and acting as molecular platforms for the synthesis of 20:5.

454 In MGD α KO lines, most molecular species containing a C16 in position *sn-1*
455 decreased in MGDG (Figure 5). MGDG 16:1-16:2, 16:1-16:3 and 16:2-16:3 species increased
456 in MGD α -eGFP-B overexpressing strain (Supplemental Figure 17). By contrast, the
457 proportion of MGDG molecular species containing a 20:5 FA increased, except the 20:5-16:4.
458 MGDG 20:5-16:3 decreased in MGD α -eGFP-B overexpressor. The proportion of all
459 molecular species containing a 14:0 FA also decreased in MGD α KO lines, but did not change
460 in overexpressing lines. Concerning DGDG, the change in molecular species profile did not
461 reflect that of MGDG. An increase in species with a 16:1 in position *sn-1* was observed in the
462 MGD α KO lines, while species containing a 20:5 decreased, except for 20:5-16:0 and 20:5-

463 16:4 (Figure 5). In the MGD α -eGFP-B overexpressor, DGDG 16:1-16:0 and 16:1-16:1
464 decreased, while 20:5-16:2 species increased (Supplemental Figure 17). Concerning SQDG,
465 in the KO lines, the *sn-1* C16-SQDG tended to increase, while species containing a 20:5 FA
466 tended to decrease. SQDG 14:0-16:0 and 14:0-16:1 species also decreased. In PG, the
467 proportion of 16:1-16:0 molecular species decreased while that of 16:1-16:1 species increased
468 (Figure 5).

469 The acyl profile of other membrane glycerolipids, *i.e.* PC, PE, DGTA and DAG
470 changed in MGD α KO and overexpressing lines (Supplemental Figures 18 and 19). In brief,
471 all 14:0-containing molecular species decreased, and they consistently increased in MGD α -
472 eGFP-B strain (Supplemental Figure 18). All C18-containing species decreased, with the
473 exception of 18:2-18:1 and 18:2-18:2 in PC, and 20:5-18:4 in PE. All species containing a
474 C16 in positions *sn-1* and *sn-2* increased in PC and DGTA. However, 16:0-16:0 and 16:0-
475 16:1 molecular species decreased in DAG while 16:1-16:1 increased. Species containing a
476 20:5 at position *sn-1* and a C16 at position *sn-2* increased or tended to increase in PC, DGTA
477 and PE. In MGD α KO, the 20:5-22:6 molecular species increased in PC (Supplemental Figure
478 19), while it consistently decreased in MGD α -eGFP-B (Supplemental Figure 18).

479 For the analysis of MGD β KO lines, we used the silent *mgd β c2* strain as a negative
480 control in addition to the WT. MGDG molecular species containing a 20:5 in position *sn-1*
481 decreased (Figure 6), with the exception of 20:5-16:0 and 20:5-16:4. MGDG 16:1-16:0 and
482 MGDG 16:1-16:1 decreased, while other MGDG species containing a C16 FA in positions
483 *sn-1* and *sn-2* and two or more unsaturations on a C16 increased. The opposite trend was
484 observed in the MGD β overexpressing strains (Supplemental Figure 17). By comparison,
485 DGDG molecular species were little affected (Figure 6). DGDG 14:0-16:1 tended to decrease
486 and 16:1-16:1 strongly decreased, in line with similar changes in MGDG. The DGDG species
487 containing 20:5 were almost not affected, with only a slight increase in 20:5-16:0 compared to
488 the controls. By contrast with MGD α mutants, only a few changes were observed in other
489 glycerolipids in MGD β KO lines: for PC, DGTA and PE, 16:1-16:2, 16:1-16:3, and 20:5-16:2
490 species tended to increase in the mutants, while 20:5-18:1, 20:5-18:2 and 20:5-18:3 tended to
491 decrease (Supplemental Figure 20). PE 20:5-20:5 species increased in MGD β KO, while it
492 decreased consistently in the MGD β overexpressing strains (Supplemental Figure 18).

493 Finally, the levels of 16:0-16:0, 16:1-16:0 and 16:1-16:1 strongly decreased in both
494 MGDG and DGDG in MGD γ KO lines, (Figure 7) and increased in MGD γ overexpressors
495 (Supplemental Figure 17). The proportion of all 14:0 FA-containing species were decreased

496 compared to the WT. For several molecular species, the *mgdyh1* and *mgdyh2* mutants behaved
 497 slightly differently compared to *mgdyi1* (Figure 7). Nevertheless, the common increase in
 498 MGDG 16:1-16:3, 16:2-16:3 and 16:3-16:3 species was coherent with an opposite trend in the
 499 overexpressors (Supplemental Figure 17). No change was observed in 20:5-containing
 500 MGDG species, except a decrease in 20:5-16:4 mirrored by an increase in the overexpressing
 501 lines. In DGDG and SQDG, 20:5-containing species tended to increase except for DGDG
 502 20:5-16:4. PG molecular species distribution was globally unaffected by MGD γ mutations.

503 In other glycerolipid classes linked with the endomembrane system, many changes in
 504 species proportions were observed in MGD γ mutants (Supplemental Figure 21). In PC and
 505 DGTA, 16:0-16:1, 16:1-16:0, 16:1-16:1, 16:1-18:1 and 16:1-18:2 species increased, while
 506 DGTA 16:1-18:3 and PC 16:1-18:4 decreased. There was a striking decrease in 20:5-20:4 and
 507 20:5-20:5 in PC, DGTA and PE. By comparison with MGD α and MGD β mutations, 14:0-
 508 containing species were not impacted in MGD γ mutants. In DAG, changes were limited to
 509 16:0-16:1 and 20:5-16:0, which tended to increase and decrease, respectively.

510 **Impact of each MGD knockout mutation on the expression level of other MGD isoforms**

511 We performed RT-qPCR analyses on MGD KO strains to check for possible
 512 compensations at the transcriptional level. In replete conditions, the WT, MGD β and MGD γ
 513 were expressed at similar levels, while MGD α was ~4 times more expressed (Supplemental
 514 Figure 22). We observed a decrease in MGD β gene expression in MGD β KO lines (2.25-fold
 515 decrease on average), which was not observed in the silent mutant *mgd β c2*. In MGD γ
 516 mutants, an increase of MGD α and MGD β expression (1.37- and 1.69-fold increase,
 517 respectively), suggests the possible activation of a compensation mechanism at the
 518 transcriptional level, when MGD γ is affected. No apparent transcriptional compensation could
 519 be detected when MGD α or MGD β were affected (Supplemental Figure 22). This does not
 520 exclude some other compensation mechanisms at the post-transcriptional level.

521 **Impact of MGD α , MGD β and MGD γ mutations on the response of *P. tricornutum* to a** 522 **nitrogen starvation**

523 The diversification of the MGD isoforms in diatoms might play a role in response to
 524 environmental variations. We have previously reported that nitrogen deprivation led to a
 525 change in MGD protein proportions (Lupette et al., 2022): the protein levels of MGD α and
 526 MGD γ were reduced to 6.1% and 26.3% of control level grown in replete conditions,

527 respectively. By contrast, MGD β protein level showed an increase to 127.0% of its level in
528 control culture.

529 Two KO mutants per MGD were used for the analysis of the impact of nitrogen
530 deprivation: *mgdai1*, *mgdaj3*, *mgd β b1*, *mgd β c1*, *mgd γ h2*, and *mgd γ i1*. Cells grown in
531 nutrient-replete medium (enriched ESAW 10N10P) were shifted to nitrogen-deprived medium
532 (00N10P) at a starting cell concentration of 3 to 3.5×10^6 cells.mL⁻¹. Growth was monitored
533 after 3, 4 and 5 days of nitrogen deprivation (Supplemental Figure 23). Nitrogen deprivation
534 condition led to a slowing down of cell division. Cultures kept growing between day 3 and
535 day 5, with WT cells reaching a concentration of $\sim 9 \times 10^6$ cells.mL⁻¹ at day 5, which is about
536 half of the concentration that would be expected in 10N10P condition. The growth of MGD α
537 and MGD β KO lines were similar to that of the WT, whereas growth was affected in both
538 MGD γ KO lines (Supplemental Figure 23), suggesting a function of this isoform in this
539 condition.

540 The TAG accumulation in response to nitrogen deprivation was first monitored by
541 Nile Red staining at days 3 and 4 following deprivation. We observed that TAG accumulation
542 was faster in MGD β KO lines compared to the WT (Supplemental Figure 24). Cells stained
543 with Nile Red at day 5 of nitrogen deprivation were observed by epifluorescence microscopy
544 (Supplemental Figure 25). All strains presented a fusiform morphotype and showed large lipid
545 droplets. Plastids appeared shrunken compared to cells grown in 10N10P condition
546 (Supplemental Figure 13), consistent with the senescence of plastid membranes previously
547 reported during nitrogen starvation (Abida et al., 2015). Besides a few lipid droplets detected
548 in the medium in the *mgd γ i1* KO mutant, no difference in cell, plastid and lipid droplet size
549 could be noticed between WT and KO lines (Supplemental Figure 25).

550 After 5 days of nitrogen deprivation, cells were harvested for glycerolipid analysis.
551 The response of *P. tricornutum* to nitrogen shortage was consistent with previous reports
552 (Abida et al., 2015) with a total FA profile becoming close to the composition of TAG
553 molecular species (Supplemental Figure 26A), due to the increase in TAG, reaching 4.10
554 nmol.10⁻⁶ cells after 5 days of nitrogen deprivation, and representing $\sim 60\%$ of total lipid
555 content (Supplemental Figure 26B).

556 Consistent with the very low MGD α protein level in the WT in low nitrogen (Lupette
557 et al., 2022), no change was detected in the proportions of membrane glycerolipids in MGD α
558 KO lines compared to the WT. For MGD β KO lines, a slight increase in PE and PG, and a

559 very small decrease in PI were observed. The changes of membrane glycerolipid distribution
560 in MGD γ KO lines were more striking, with lower proportions in DGDG (~22% decrease)
561 and DAG (~10% decrease), and an increase in PC (~27 % increase) and DGTA (~57%
562 increase) compared to the WT (Supplemental Figure 26B). Quantitatively, the decrease in
563 DGDG was roughly equivalent to the increase in PC.

564 Concerning MGDG acyl composition, the impact was barely detected in MGD α KO
565 mutants (Figure 8), in line with the expected lack of involvement of this isoform in nitrogen
566 shortage. By contrast, MGDG molecular species composition was severely affected in MGD β
567 and MGD γ KO lines (Figure 8) suggesting that these two isoforms play a role under nitrogen
568 stress. In MGD β KO lines, a strong decrease in 14:0-16:2 and 14:0-16:3 was observed (Figure
569 8). Although MGDG 14:0-16:1 seemed stable, a decrease in DGDG 14:0-16:1 was observed.
570 MGDG 16:1-16:2, 16:1-16:3 and 16:2-16:3 strongly decreased while 16:0-16:0, 16:1-16:1,
571 16:3-16:3 and 16:4-16:3 increased. All 20:5-containing MGDG species decreased, except for
572 20:5-16:4 MGDG that strongly increased. In DGDG, 16:1-16:1, 16:1-16:2, and only 20:5-
573 16:2 and 20:5-16:4 species showed a moderate increase, while other 20:5-containing species
574 were stable.

575 Changes in galactolipid species distribution were the most spectacular in MGD γ
576 mutants (Figure 8): MGDG and DGDG species containing a 16:0 or a 16:1 in position *sn*-2
577 and either a 14:0 or a saturated or mono-unsaturated C16 in position *sn*-1 decreased. On the
578 contrary, species with a 16:2 or 16:3 in position *sn*-2 increased. Most 20:5-containing
579 galactolipid species were mainly increased, showing that compensatory mechanisms most
580 likely relying on MGD β , channeled 20:5 to galactolipids.

581 In other lipids, the impact of MGD α mutation was mild (Figure 8, Supplemental
582 Figure 27). In MGD β KO lines, the proportion of 14:0-16:0, 14:0-16:1, 20:5-16:0 in SQDG
583 and 14:0-16:0, 20:5-16:0 and 20:5-16:1 in PG decreased. The proportion of 16:1-16:0, 16:1-
584 16:2, 16:2-16:0 and 16:2-24:0 in SQDG, and that of 16:0-16:0, 16:1-16:0 and 16:1-16:1 in PG
585 increased. MGD γ mutation had no effect on SQDG, whereas the proportion of PG species
586 containing a C16 FA in position *sn*-1 decreased while 20:5-containing PG increased. It seems
587 that PG species were affected similarly to the galactolipid species with the same FA
588 composition, suggesting that they shared some substrates in their biosynthetic pathways.

589 The MGD γ KO lines were the only ones to show a striking impact on cytosolic TAG
590 upon nitrogen shortage, verified by opposite effects in overexpressing lines. MGD γ

591 inactivation led in particular to an increase of 14:0-16:1-16:1, 16:1-16:1-16:1 and 20:5-16:1-
592 16:1 TAG species that were consistent with an opposite trend in the overexpressing lines in
593 optimal conditions (Supplemental Figures 18 and 27). Additionally, KO of MGD γ in nitrogen
594 starvation led to a decrease of 14:0-16:0-16:0, 16:0-16:0-16:0, 16:0-16:0-16:1, 16:0-18:1-
595 16:0, 16:1-18:1-16:0, and 16:1-18:1-16:1 TAG species, consistent with an increase of these
596 species in the MGD γ overexpressing lines (Supplemental Figures 18 and 27).

597 Considering that MGD α is naturally absent under nitrogen shortage, the knockout of
598 either MGD β or MGD γ could be seen as a double invalidation (*mgda/mgd β* or *mgda/mgd γ* ,
599 respectively), resulting in stronger phenotypes compared to those observed in replete medium.
600 MGD β looked important to keep a high level of polyunsaturated galactolipids in the absence
601 of MGD α . MGD γ seems to play a specific role in response to nitrogen shortage, in particular
602 on galactolipids with a low unsaturation level. MGD γ KO had a stronger effect on non-
603 plastidial lipids than MGD β KO. The increase in DGTA 16:0-16:1 and 16:1-16:1 in MGD γ
604 KO lines was particularly striking and highlights a role for this isoform outside of the plastid,
605 consistent with its ER/EpM localization. Modification of TAG species distribution suggests a
606 coupling between MGD γ products and TAG synthesis.

607 **MGD β /MGD γ double mutations**

608 To better understand the role of MGD α , three MGD β /MGD γ double KO mutants were
609 generated (2KO4, 2KO6, and 2KO9; Supplemental Figures 28). Their growth in optimal
610 conditions was slightly affected compared to the WT (Supplemental Figures 29), nevertheless
611 the *mgd β /mgd γ* KO proved non-lethal. The double KO mutants showed no variation in the
612 MGDG levels compared to the WT, demonstrating that MGD α alone was capable of
613 maintaining MGDG production in the cell at a level equivalent to the WT, at least
614 quantitatively. A mild decrease of DGDG (Supplemental Figures 30) contrasted with the
615 results of the MGD β and MGD γ single KO mutants. The weakly unsaturated C16-containing
616 MGDG species showed a decrease whereas the highly unsaturated ones increased drastically
617 (Supplemental Figures 31), confirming that these MGDG species are more specifically
618 produced by MGD α . The vicinity of plastidial desaturases (Dolch and Marechal, 2015) and
619 their strong activity may be at the origin of both variations, unbalancing the ratio between
620 weakly and strongly unsaturated C16-containing MGDG. A reduction of the 20:5-containing
621 MGDG species, but not their elision, demonstrates that MGD α is able to process DAG species
622 containing 20:5 at the *sn-1* position. The acyl profile of DGDG is peculiar and very
623 informative as well (Supplemental Figures 31). A dramatic decrease of 14:0-16:1, 16:1-16:0

624 and double 16:1 was recorded, accompanied with a slight increase in 20:5-16:0 and 20:5-16:1.
625 The lack of reduction of the 16:1-16:2 DGDG may imply that this species is produced by a
626 DGDG synthase unable to process MGD α products. The increase in 20:5-containing DGDG
627 species supports the hypothesis that a specific DGDG synthase uses 20:5-containing MGDG
628 as preferential substrate and is highly active because, although its substrate is reduced in the
629 double KO mutants, its product is either unvaried or increased. Considering all other lipid
630 classes (Supplemental Figures 31 and 32), acyl profiles show some opposite trends to those
631 observed in MGD α KO lines (Figure 5 and Supplemental 19).

632

633

ACCEPTED MANUSCRIPT

634 **Discussion.**

635 **The emergence of a multigenic family of MGDs likely occurred before the radiation of**
636 **diatoms and pelagophyceae from other photosynthetic stramenopiles**

637 Phylogenetic analyses of MGD proteins reported earlier (Hori et al., 2016; Shang et
638 al., 2022) and completed here (Supplemental Figure 2) revealed the conservation of three
639 major isoforms in diatoms, originating from gene duplications and specialization, similar but
640 distinct to the process that led to the emergence of type A and type B in angiosperms. Based
641 on parsimonious assumptions, an ancestral MGD α type might have been vertically inherited
642 from a red alga during the secondary endosymbiosis event at the base of all stramenopiles.
643 Then, a first gene duplication occurred, before the radiation of diatoms and phaeophyceae
644 from other stramenopiles. This first gene duplication is likely to have led to the emergence of
645 a common ancestor to MGD β and MGD γ . The second duplication potentially followed the
646 separation of pelagophyceae from diatoms. Gene or genome duplications are crucial drivers in
647 the evolution of stramenopiles (Martens et al., 2008; Parks et al., 2018). In pennate diatoms
648 for instance, up to six whole genome duplications (WGD) have been identified (Parks et al.,
649 2018). It is important to stress that the genomes of some non-diatom stramenopiles contain
650 only one MGD -such as the eustigmatophytes *Nannochloropsis* or *Microchloropsis* (Vieler et
651 al., 2012; Corteggiani Carpinelli et al., 2014). A single isoform seems therefore sufficient for
652 the biogenesis of a secondary plastid bounded by four membranes.

653 ***P. tricornutum* MGD α , MGD β and MGD γ are active MGDG synthases and have distinct**
654 **substrate specificities.**

655 After assessing the accurate sequences of MGD proteins encoded by the genome of *P.*
656 *tricornutum*, structural modelling highlighted a typical double Rossmann fold (Supplemental
657 Figure 4) characterizing the three-dimensional structure of MGD determined in angiosperms
658 (Rocha et al., 2016). The critical residues for activity and involved in a PG-H catalytic dyad
659 are conserved supporting that diatom proteins are functional MGDs. A noticeable distinctive
660 feature lied in the relative size of the N- and C-domains among isoforms. Overall, the hinge
661 separating the N- and C-Rossmann folds was longer in *P. tricornutum* proteins than in plant
662 sequences. Compared to AtMGD1 and MGD α , the C-domain of MGD β and MGD γ contains
663 additional stretches of 50 and 110 amino acids, respectively. These two regions appeared to
664 be essentially hydrophilic and could influence protein function. In addition, the N-domain of
665 MGD β and MGD γ contains 82 and 81 additional amino acids, respectively. MGD β and

666 MGD γ mature proteins are therefore bigger; the role of the detected protein extensions is
667 unknown. The enzymatic function was confirmed by (1) heterologous expression of MGD
668 isoforms in yeast and (2) lipid changes in *P. tricornutum* MGD KO lines and overexpressors.
669 These two sets of experiments showed that substrate specificity was distinct for each isoform,
670 with consistent results.

671 ***P. tricornutum* MGD α , MGD β and MGD γ localize to distinct subcellular compartments** 672 **inside and outside of the plastid**

673 MGD α is predicted to contain a bipartite Sp-Ctp presequence suggesting its
674 localization inside the plastid. A putative cleavage site is predicted at the level of a SAAFSP
675 sequence, matching partly with the profile of the 'ASAFAP motif' characterized in nuclear-
676 encoded plastid proteins (Kilian and Kroth, 2005). In *P. tricornutum* expressing MGD α -eGFP
677 fusion proteins, localization was consistently observed in thylakoid membranes (Figure 2).
678 The localization was confirmed by co-expression with a thylakoid marker (Supplemental
679 Figure 8).

680 In MGD β , a Sp and a GSGFVL sequence with weak similarity to the ASAFAP motif
681 were predicted. MGD β -eGFP fusion proteins localize to the blob-like structure (Figures 2 and
682 3) and was confirmed by coexpression with a blob/PPM membrane marker (Supplemental
683 Figure 8). These results indicate that the precursor of MGD β is targeted to the secondary
684 plastid, successfully crosses the EpM, but, at least in over-expressing system, is not
685 transported further through the oEM and iEM. Arrest in the blob-like structure might be
686 determined by specific changes in the ASAFAP motif, corresponding to a different amino
687 acid pattern than previously characterized (Apt et al., 2002; Kilian and Kroth, 2005). This
688 may explain why MGD β was not detected by bioinformatics methods developed to predict the
689 secondary plastid proteome in *P. tricornutum* (Gruber et al., 2015). Additionally, MGD β
690 localization could extend to the PPM, as the vesicular network present in the blob-like
691 structure emerges from the PPM (Flori et al., 2016). Since the oEM and iEM correspond to
692 the chloroplast envelope in primary plastids, in rigorous terms, the suborganellar localization
693 of MGD β represents a case of MGD outside a membrane system vertically related to the
694 primary chloroplast. This evolution may have been determined by relevant mutations of the
695 ASAFAP motif of an MGD α -like ancestral protein.

696 Finally, no targeting sequence could be predicted in MGD γ , which was consistent with
697 an MGD γ -eGFP fluorescence detected outside of the secondary plastid, in the ER and

698 possibly the EpM (Figure 2 and Supplemental Figure 8). In a recent report, MGD γ -GFP
699 expression in *P. tricornutum* showed a signal partly overlapping that of chlorophyll, and
700 outside of the plastid (Shang et al., 2022). Authors suggested that such pattern was that of an
701 intra plastidial localization, whereas presented images rather show a localization at the
702 periphery of the plastid consistent with results presented here. Localization of MGD γ in the
703 ER and EpM is consistent with its detection in two replicates of the proteome of cytosolic
704 lipid droplets isolated from *P. tricornutum* (Leyland et al., 2020). Overall, these results
705 support that MGD γ is distributed between the ER and the EpM, and possibly the cytosol. In
706 *Arabidopsis*, although MGDs are classically considered bound to chloroplast envelope
707 membranes, the initial study of AtMGD2 and AtMGD3 fused to GFP showed some labelling
708 diffusing in the cytosol (Awai et al., 2001) and AtMGD2 was shown to localize to the cytosol
709 of elongating pollen tubes (Billey et al., 2021a). Based on this localization, MGD γ represents
710 an evolutionary innovation, performing MGDG synthesis outside of the plastid, for membrane
711 lipid synthesis involving other cell compartments. This is in line with the lipid phenotype we
712 observed. Results suggest therefore an evolution by gene duplication and acquisition of new
713 cell function, in the following order, MGD α \rightarrow MGD β \rightarrow MGD γ , at the base of diatoms'
714 radiation.

715 None of the isoforms seems to localize to the iEM nor the oEM, as in primary
716 chloroplasts of plants and green algae (Petroustos et al., 2014). The concentric localization
717 (MGD α / thylakoid; MGD β / PPM; MGD γ / ER-EpM) suggests distinct roles for each of
718 membrane compartment, and some kind of sophisticated interactions among glycerolipid
719 pathways occurring inside and outside the secondary plastid. This localization also contradicts
720 the assumption that MGDG, and its derivative DGDG, would be restricted to thylakoid, iEM
721 and oEM membranes. Multiple pools of MGDG are likely to occur in diatom cells, an
722 important feature to take into account in the interpretation of the lipid phenotypes in MGD
723 KO and overexpressing lines.

724 **The blob-like-structure, a lipid-synthesis machinery playing a role during cytokinesis?**

725 Observation of MGD β -eGFP in the blob-like structure between the PPM and oEM
726 also gave insightful information about the positioning of this structure, particularly during
727 plastid division (Figure 3). Our observations of MGD β -eGFP fluorescence by confocal
728 microscopy and of the WT by STEM suggest that the VN might converge and take a blob-like
729 structure mostly during cytokinesis, from the moment the plastid starts dividing until the
730 whole cell division has been completed. When the plastid does not show any sign of division,

731 the MGD β -eGFP fluorescence signal was more diffuse, either suggesting that the VN itself
 732 loses the blob structure and become more scattered, or that MGD β protein relocates to the
 733 PPM near the blob. During plastid division, the blob-like structure was located at the central
 734 constriction site of the dividing plastid. Following division, the two generated plastids seem to
 735 remain attached by the blob-like structure, meaning that the PPM and EpM are not fully
 736 divided, probably until the cleavage furrow comes through the blob-like structure. This is
 737 particularly important as it supports that the oNE-EpM continuum is never broken during
 738 plastid division, a point that was not addressed yet in studies of diatoms' plastid division
 739 (Tanaka et al., 2015). These observations suggest that MGD β could be involved in the
 740 provision of MGDG to the growing and dividing plastid membranes. Nevertheless, such
 741 possible role appears as non-essential, since MGD β KO lines grow nearly as well as WT.

742 **Functional compensation mechanisms between MGD isoforms**

743 For phenotypic analyses of KO lines, a minimum of two CRISPR-Cas9 KO mutants,
 744 each obtained with a different sgRNA, were used. None of the isoforms proved strictly
 745 necessary for cell survival. In spite of careful scrutiny, neither cell growth and morphology,
 746 nor membrane integrity, nor photosynthetic activity were affected, suggesting either a
 747 dispensable role of the targeted isoforms or, more likely, a performant complementation by
 748 the action of intact isoforms. Under optimal culture condition, moderate differences could be
 749 observed in the relative proportions of lipid classes of all mutant strains. No change in
 750 transcript level was observed in MGD α and MGD β KO lines. By contrast, in MGD γ mutants,
 751 an increase in *MGD α* and *MGD β* transcripts is one of the possible compensation mechanisms
 752 involved in addition to the control of MGD protein stability and enzymatic regulation (*i.e.*
 753 activation of remaining isoforms *via* metabolic regulatory loops). In addition, to establish
 754 appropriate MGDG levels at the different subcellular compartments, intense fluxes of MGDG
 755 between membranes need to be postulated.

756 **Specialization of MGD α in thylakoid membrane synthesis**

757 Expression analysis in *P. tricornutum* WT strain showed that *MGD α* was the most
 758 expressed isoform under our optimal laboratory conditions, with relative expressions ~4 times
 759 higher than that of *MGD β* and *MGD γ* . In a recent study, *MGD γ* was reported to be relatively
 760 more expressed than other isoforms, when cells were grown in a medium containing less
 761 nitrogen (0.88 mM instead of 5.5 mM used here) (Shang et al., 2022), a nutrient controlling
 762 the level of *MGD α* expression. MGD α basal role under optimal conditions is consistent with

763 its higher similarity to MGDs from other phylogenetic lineages, and the conserved role in
764 MGDG production for the massive demand of lipids for photosynthetic membranes. The
765 relative protein level of MGD α was shown to drop to 6 % in a whole-proteome study of
766 nitrogen starved *P. tricornutum* (Lupette et al., 2022), coinciding with a decrease of thylakoid
767 membranes observed in this stress condition (Abida et al., 2015). It also reveals that MGD β
768 and MGD γ might be recruited to sustain the cell in MGDG during nitrogen starvation. The
769 complementary role of MGD β and MGD γ for an accurate biogenesis of thylakoids is further
770 demonstrated in MGD α KO lines, showing only a moderate phenotype, with apparently well-
771 structured plastids and healthy photosynthesis.

772 **MGD α uses a DAG pool deriving from plastid 16:0- and 16:1-ACP and produces**
773 **MGDG with 16 carbon-FA at position *sn*-1, unsaturated up to 16:4 by plastid acyl-**
774 **desaturases**

775 Through combined observations of MGD α , MGD β and MGD γ and MGD β /MGD γ KO
776 effects, we identified the main DAG substrates used by each isoform and the fate of their
777 products (Figure 9). It is important to note that this reasoning is based on whole cell lipidomic
778 changes, considering both the location of an MGD isoform and the DAG substrate at its
779 vicinity. Substrate specificity was further confirmed for medium chain fatty acids (C16) by
780 the analysis of yeasts after heterologous expression.

781 MGD α substrate specificity is consistent with its localization close to *de novo* FA
782 synthesis, as well as to the plastid desaturases. Under optimal condition, MGD α KO most
783 strongly affected MGDG species containing a C16-FA with two or more unsaturations, except
784 those containing a C20:5 in *sn*-1 position. These species are highly reduced in nitrogen
785 deprivation, when MGD α is not expressed. In the total FA distribution in MGD α mutants, a
786 small increase in 16:1 was observed, while 16:2, 16:3, and 16:4 FAs tended to decrease.
787 Given that in *P. tricornutum* 16:0-ACP is produced by the FA synthase in the stroma of the
788 plastid, desaturated into 16:1-ACP by a plastid palmitoyl-ACP-desaturase (Smith et al., 2021),
789 the pool of DAG used by MGD α is likely generated inside the plastid. In addition, C16-
790 desaturations are performed by desaturases acting mostly on MGDG, which were predicted to
791 localize to the stromal side of the plastid issued from the detection of Sp-Ctp and ASAFAP
792 motifs (Dolch and Marechal, 2015). Based on observed phenotypes, MGD α could mainly use
793 DAG with the following compositions: 14:0-16:0, 14:0-16:1, 16:0-16:0, 16:1-16:0 and 16:1-
794 16:1 (the main DAG species detected by LC-MS), corresponding to a *de novo* production of
795 FA. The heterologous expression of MGD α in yeast confirms this capacity to use DAG with

796 C16 unsaturated substrates, but MGD α appears to be also able to use C18 species in this non-
797 native context. Once the corresponding MGDG products are formed, they do not accumulate,
798 as plastid desaturases rapidly generate MGDG species with higher unsaturations, *i.e.* 16:2,
799 16:3, and 16:4 (Figure 9 and Table 1).

800 The contrasting effects on 20:5-containing DGDG molecular species (decrease in the
801 MGD α KO and increase in the MGD β /MGD γ double KO) along with the corresponding
802 increase or decrease in DGDG species containing 16:1 at the sn-1 position, provide additional
803 insight into the fate of the MGDG pool produced by MGD α . It is well established in *A.*
804 *thaliana* that AtMGD1 and AtDGD1 operate together for the bulk of thylakoid galactolipids,
805 whereas AtMGD2/3 and AtDGD2 operate together when galactolipids contribute to whole-
806 cell membrane lipid remodeling processes (Benning and Ohta, 2005). A reasonable
807 hypothesis is that MGD α operates in close association with a specific DGDG synthase
808 (MGD α -associated DGD), very active on any 20:5-containing MGDG produced. In the
809 absence of MGD α , the unused DAG pool is available for the other MGD isoforms. The
810 MGDG thus produced close to MGD β and/or MGD γ would fuel the syntheses of distinct
811 DGDG pools (Figure 9).

812 **MGD β uses a DAG pool enriched in 20:5 imported from the cytosol, and contributes to** 813 **the production of highly unsaturated MGDG species in nitrogen deprivation**

814 Under optimal conditions, MGD β mutation strongly decreased the accumulation of
815 20:5-containing MGDG species, except MGDG 20:5-16:4. Interestingly, MGDG 14:0-16:1,
816 16:1-16:0 and 16:1-16:1 species also decreased. This firstly suggests that under normal
817 conditions, MGD β mostly uses 20:5-16:0 and 20:5-16:1 DAG, and that the corresponding
818 MGDG products are rapidly desaturated, leading to the accumulation of 20:5-containing
819 MGDG species with a higher unsaturation level of the C16 at position *sn*-2 (Figure 9 and
820 Table 1). Given the localization of this isoform in the blob/PPM, the 20:5-containing MGDG
821 it produces need to be rapidly transported to the thylakoid membranes where C16-desaturases
822 are located. The role of MGD β in the production of 20:5-containing MGDG species is also
823 consistent with its subcellular localization, since 20:5 FA is produced outside or at the
824 periphery of the plastid following elongation and desaturation steps (Figure 9). For the
825 massive incorporation of this FA into the plastid, 20:5 and/or 20:5-containing precursor need
826 to cross the four limiting membranes, via the PPM where MGD β is strategically localized.

827 The decrease of MGDG 14:0-16:1, 16:1-16:0 and 16:1-16:1 species suggests that these
 828 three species are also direct products of MGD β , but that most of them are not directly
 829 available for desaturation. The heterologous expression of MGD β in yeast highlights clearly
 830 32:1 and 32:2 MGDG as major products. Therefore, these saturated/monounsaturated MGDG
 831 species might derive from another pool of DAG containing FAs synthesized *de novo* in the
 832 stroma. MGDG 16:1-16:0 and 16:1-16:1 proportions increase in the MGD β overexpressing
 833 lines, while the MGDG 16:2-16:3 and 16:3-16:3 decreased. By this mean, MGD β can produce
 834 more saturated forms of MGDG, remaining away from stromal C16-desaturases, thus
 835 compensating for the loss of MGD γ .

836 Thus, on the one hand, MGD β appears able to use a DAG pool deriving from the *de*
 837 *novo* synthesis of FA in the stroma, but forming MGDG accumulating in non-photosynthetic
 838 membranes. On the other hand, MGD β uses a DAG pool enriched in 20:5 imported from the
 839 ER. This isoform in the blob-like structure seems therefore to facilitate bidirectional fluxes of
 840 MGDG molecular species deriving from internal and external DAG pools.

841 In this case also, changes in the DGDG profile highlight a possible association of
 842 MGD β isoform with a specific digalactosyldiacylglycerol synthase (DGD) enzyme. Under
 843 normal conditions, only DGDG 16:1-16:1 and 16:1-16:2 proportions decreased in MGD β KO
 844 lines, while DGDG 14:0-16:1 tended to decrease. This would be consistent with the role of
 845 MGD β in the production of MGDG 14:0-16:1 and MGDG 16:1-16:1 for the formation of a
 846 specific pool of DGDG. The decrease in DGDG 16:1-16:2 is likely related to that of DGDG
 847 16:1-16:1 from which it can be produced through desaturation by the plastid FAD6 desaturase
 848 (Dolch and Marechal, 2015). The stable proportion of 20:5-containing DGDG species might
 849 be due to compensation effects by MGD α and MGD γ . In line with this, 20:5-containing
 850 MGDG are still produced, and could be used in priority to maintain 20:5-containing DGDG
 851 levels, most importantly by the MGD α -associated DGD pointed above.

852 Finally, under nitrogen deprivation, at low MGD α level and in the absence of MGD β ,
 853 MGD γ could compensate for the lack of MGDG, and for the provision of substrates for
 854 DGDG species.

855 **MGD γ uses a DAG pool enriched in 16:0 and 16:1 exported to the EpM/ER, for the**
 856 **production of MGDG species with low unsaturation level in non-platidial membranes.**

857 Under optimal conditions, MGD γ mutation led to a strong decrease in all 14:0-
 858 containing MGDG species and in MGDG 16:0-16:0, 16:1-16:0, 16:1-16:1 and 20:5-16:4.

859 Therefore, the main DAG substrates for MGD γ under normal conditions are likely those
860 directly deriving from *de novo* FA synthesis in the stroma, *i.e.* 14:0-16:0, 14:0-16:1, 16:0-
861 16:0, 16:0-16:1 and 16:1-16:1, requiring that these DAG are exported by specific systems, at
862 least to the EpM (Figure 9). Heterologous expression in yeast is consistent with this substrate
863 preference. Given that MGD γ is located in the ER and at the outermost membrane of the
864 plastid, its MGDG products are not directly available for plastid desaturases. Coherently,
865 MGD γ mutation did not have any strong impact on MGDG species containing a C16 with two
866 or more unsaturations, except for 14:0-16:2 and 14:0-16:3 species. The strong impact on
867 MGDG 14:0-16:0, 14:0-16:1, 16:0-16:0, 16:1-16:0, 16:1-16:1 species indicates that MGD γ is
868 responsible for their production and accumulation (Figures 9 and 10).

869 Under nitrogen deprivation, MGD γ also led to a very strong decrease of the same
870 molecular species. Coherently, MGDG 14:0-16:1, 16:0-16:0, 16:1-16:0 and 16:1-16:1 species
871 increased in MGD γ overexpressing lines, while MGDG 16:1-16:3, 16:2-16:3 and 16:3-16:3
872 decreased. The products of MGD γ are therefore able to accumulate in cell membranes where
873 plastid C16-desaturases are not present.

874 The nature of the membranes where these MGDG species accumulate is intriguing,
875 because this represents an observation that MGDG synthesis in a non-plastidial membrane
876 system would occur in a photosynthetic eukaryote. The MGD γ subcellular localization
877 suggests that MGDG might be synthesized in the EpM and the ER. By extension, this pool of
878 MGDG could virtually be transferred to any other compartment of the endomembrane system,
879 directly or indirectly connected to the ER via membrane continuity (like the outer nuclear
880 envelope) or via vesicular trafficking. MGDG presence outside of the plastid is further
881 supported by the comparison of lipidomes of purified plastids and whole cells of *P.*
882 *tricornutum* (Huang et al., 2024). We have no element to comprehend the role of MGDG in
883 extraplastidial membranes, and this represents therefore a critical question to address in
884 diatoms. The role of MGDG in membranes, a lipid class usually characterized by the high
885 unsaturation level of its esterified FAs, is often associated to its capacity to form Hexagonal II
886 phase. However, it is uncertain whether the highly saturated products of MGD γ would
887 organize in a similar way or not. Therefore, both the localization and the distinct nature of
888 MGD γ products should be studied in the future.

889 MGD γ KO led to a decrease of DGDG 14:0-16:1, 16:0-16:0, 16:1-16:0, and 16:1-16:1,
890 and to a proportional increase in 20:5-containing DGDG species. These changes were
891 moderate in optimal conditions, and very strong in nitrogen deprivation condition. Therefore,

892 we can suppose that a particular DGD enzyme might be dependent on MGD γ for the
893 production of DGDG species with a 14:0 or a C16 in position *sn-1*. This hypothetical DGD
894 enzyme should localize at or close to the EpM. Nevertheless, it must be noted that under
895 nitrogen deprivation tuning down MGD α expression, in the MGD β KO, MGD γ was able to
896 produce 20:5-containing MGDG species, and maintained a sufficient provision of substrates
897 for the production of 20:5-containing DGDG species. Therefore, MGD γ appears to have
898 flexibility in its roles, able to also use DAG 20:5-16:0 and 20:5-16:1 for the production of
899 MGDG species (Figure 9). This role of MGD γ in the production of 20:5-containing MGDG is
900 however minor, when other MGDs are active.

901 **MGD isoforms and extraplastidial lipid homeostasis**

902 Mutation and overexpression of the three MGD isoforms led to changes in non-
903 plastidial lipids, showing a high level of integration between membrane glycerolipid
904 syntheses. The observed phenotypes support the existence of a flux of 16:0/16:1-rich DAG
905 going outward from the stromal side of the plastid to the EpM and ER, where this pool is
906 available for extraplastidial lipids formation by MGD γ . They also highlighted that 20:5-DAG
907 species could be imported inward or produced at the PPM, forming a DAG pool available for
908 MGD β , and further transported to the stroma, where it joined the 16:0/16:1-rich DAG pool
909 available for plastidial lipid syntheses by MGD α . The results presented here show that the
910 impact of MGD α and MGD β on extraplastidial lipids rather reflect the establishment of a new
911 equilibrium compensating alterations in DAG pools. By contrast, alterations of MGD γ reveal
912 more spectacular changes, supporting a specific role of this isoform in whole-cell lipid
913 remodeling occurring in response to environmental stress conditions.

914 MGD γ mutation had a stronger impact on non-plastidial lipid composition. In
915 particular, 16:1 and 16:0 containing PC and DGTA increased both under optimal and nitrogen
916 deprived conditions. MGD γ being located at the periphery and outside of the plastid, it is
917 possible that it sits near to PC and DGTA synthesis sites. Therefore, it is coherent that the
918 DAG species not used by MGD γ to produce MGDG are used as substrate for the synthesis of
919 PC and DGTA (Figure 9). Following MGD γ KO, 16:0 amounts available for elongation into
920 18:0 by Δ^0 -elongases (Dolch et al., 2017a), the committing step for further desaturations up to
921 18:4 and 20:5 in phospholipids and DGTA increased. The amounts of PC, PE and DGTA
922 containing an 18:1 or an 18:2 increased. It seems however that the rates of 18:2-to-18:3 and
923 18:3-to-18:4 desaturations and 18:4-to-20:4 elongation were slower than the 16:0 export rate.

924 In *P. tricornutum* and other photosynthetic stramenopiles, an excess of 16:0 or 16:1 in
 925 the cytosol is known to be ‘absorbed’ by an increased production of the storage lipid TAG
 926 (Dolch et al., 2017a; Smith et al., 2021). Under optimal growth conditions, MGD γ KO led to
 927 strong differences in the profile of several TAG species, with a decrease in 16:0-16:0-16:0
 928 and 16:0-16:0-16:1, and an increase in 16:1-16:1-16:1. This impact on TAG was more
 929 pronounced when cells were subjected to a nitrogen starvation. This coupling between MGD γ
 930 activity and TAG production highlights a contribution of this MGD isoform to the intense
 931 membrane and storage glycerolipid remodeling occurring in response to a variety of
 932 environmental factors. The control of a MGDG/TAG balance is known to occur in other algal
 933 systems, such as *Chlamydomonas reinhardtii* containing only one *MGD* gene (Li et al., 2012;
 934 Gu et al., 2021; Iwai et al., 2021). Diatoms appear to contain an isoform of MGD in their
 935 extraplastidial membranes specialized in controlling part of this equilibrium.

936 **Comparison of the emergence of specialized MGD isoforms in diatoms and in** 937 **angiosperms.**

938 The emergence of specialized MGD isoforms in diatoms and angiosperm is a
 939 fascinating example of evolutionary convergence, as important similarities can be drawn. In
 940 both cases, MGD isoforms derived from an ancestral form dedicated to the production of the
 941 bulk of MGDG for photosynthetic membranes, following gene duplications and speciation
 942 occurring before the main radiation of the phylum. Whereas in angiosperms, type B MGD
 943 could gain function distinct from type A bound to the iEM, following a relocation to the oEM
 944 after a loss of the Ctp, MGD β/γ could gain additional function following a similar topological
 945 relocation. A plausible mechanism lies in the evolution of the original Sp-Ctp bipartite
 946 sequence characterizing MGD α , with mutations of the ASAFAP motif arresting the import of
 947 the protein precursor into the PPM (MGD β), followed by a loss of the targeting sequence,
 948 leading to a localization in the EpM/ER (MGD γ). In both angiosperms and diatoms, the
 949 topological separation of MGD isoforms correlated with the use of DAG pools with distinct
 950 acyl-profiles, the production of MGDG with distinct acyl-profiles, and the specific association
 951 with DGD enzymes, also producing DGDG with distinct acyl-profiles. In both cases, MGDG
 952 dedicated to photosynthetic membrane highlighted the highest level of saturations (16:3 and
 953 18:3 in angiosperms, up to 16:4 and 20:5 in diatoms).

954 In both cases, the emergence of ‘novel’ MGD isoforms (type B MGDs in angiosperm
 955 and MGD β/γ in diatoms) strengthens the integration of the organelle within the cell and
 956 implies intense trafficking of lipids. On the one hand, these novel isoforms facilitate the

957 incorporation of DAG molecular species deriving from extraplastidial precursors, via the so-
958 called eukaryotic pathway in angiosperms (Browse et al., 1986), and the still unresolved Ω -
959 pathway in diatoms (Petroustos et al., 2014). The nature of the imported precursor(s) may
960 differ, as well as the mechanism involved. On the other hand, without sparing MGDG
961 required for photosynthesis, novel MGD isoforms could be involved in the specific
962 production of galactolipids, for novel function, like lipid remodeling occurring under
963 environmental stress. An example is the role of AtMGD2/3 in *Arabidopsis* upon phosphate
964 shortage (Kobayashi et al., 2009b) and that of MGD γ and to a lesser extent, MGD β , under a
965 lack of nitrogen. In both cases, the emergence of specialized isoforms correlated with a
966 remarkable ecological success of each phylum populating a variety of terrestrial and marine
967 habitats, marked by constant environmental changes and stress conditions.

968 Among the differences, the roles played by type A and type B do not overlap in
969 angiosperms: a KO of type A is lethal, and cannot be compensated by type B MGDs
970 (Kobayashi et al., 2007a, b). Furthermore, type B MGDs appear as specific to some
971 environmental conditions, like phosphate starvation, performing function that type A cannot
972 compensate. By contrast, the multigenic family of MGDs in diatoms showed strong
973 overlapping activities, preventing the loss of integrity of photosynthetic membranes. Any
974 defect of one isoform is nearly completely compensated by the two others. Even under
975 nitrogen shortage, when MGD α level is drastically reduced, MGD γ compensates for the loss
976 of MGD β and *vice versa*. This does not make the dissection of respective roles easy, but this
977 highlights the powerful flexibility of membrane glycerolipid metabolism in diatoms, despite
978 their very sophisticated membrane compartmentation.

979 Eventually, diatoms introduced an important milestone in the evolution of MGD
980 proteins, unreported to date in any other phylum, marked by the targeting in systems
981 ontogenetically unrelated to membranes deriving from a primary chloroplast, *i.e.* in the PPM
982 (MGD β) supposed to derive from the red alga endosymbiont plasma membrane, and the EpM
983 and ER (MGD γ). MGDG has been assumed to be the hallmark of chloroplastic membranes,
984 and in textbooks it is referred to as such. We do not have yet the conceptual framework to
985 address MGDG role outside of the plastid, but this work provides genetic tools to address this
986 fascinating question. It is important to note, that whereas the biogenesis of the iEM and oEM
987 could be supposed to follow that of the primary chloroplast envelope in angiosperms and
988 green algae, and whereas the biogenesis of the EpM could be considered to derive from the
989 ER, no lipid-synthesis protein could be located to the PPM until now. This work provided

990 therefore clues on the biogenesis of all membrane systems of a secondary plastid, with
991 biological processes unsuspected until now, paving the way for future works.

992

ACCEPTED MANUSCRIPT

993 **Materials and Methods.**

994 **Primary sequence analyses.**

995 *Phaeodactylum tricornutum* MGD genes were identified based on protein sequence
 996 homology using BLAST, with Arabidopsis (*Arabidopsis thaliana*) main isoform, *i.e.*
 997 AtMGD1 (NP_194906), as query. The identified locus tags of *P. tricornutum* were named
 998 *MGD* α (Phatr3_J14125, Uniprot; XM_002181649.1, RefSeq; J14125, EnsemblProtists),
 999 *MGD* β (Phatr3_J43938; XM_002186319; J54168) and *MGD* γ (Phatr3_J9619;
 1000 XM_002176764.1; J9619) respectively. Target sequence predictions were performed
 1001 considering multiple START/Methionine codons present in the N-terminal part of putative
 1002 reading frames. Signal peptides (Sp) were predicted with the SignalP-6.0 tool, with default
 1003 parameters (Emanuelsson et al., 2007; Almagro Armenteros et al., 2019) Chloroplast transit
 1004 peptide (Ctp) were predicted using Wolf PSORT (Horton et al., 2007) and ChloroP - 1.1
 1005 (Emanuelsson et al., 2007), with default parameters. Specific bipartite pre-sequence were
 1006 predicted using the HECTAR software specifically designed to detect Sp-Ctp in stramenopiles
 1007 (Gschloessl et al., 2008).

1008 **Protein structure predictions.**

1009 The software Phyre² (Protein Homology/analogY Recognition Engine V 2.0) (Kelley
 1010 et al., 2015), and AlphaFold (Jumper et al., 2021; Jumper and Hassabis, 2022) run from
 1011 ChimeraX (Mirdita et al., 2022) were used to model *P. tricornutum* MGD sequences, which
 1012 were then viewed on PyMol (Janson et al., 2017). AtMGD1 structure (4X1T- wwPDB
 1013 consortium, 2019) (Rocha et al., 2016) was used to build Phyre² PtMGDs models. For each *P.*
 1014 *tricornutum* MGD, the peptide sequences used as input were those deduced after cDNA
 1015 sequencing and signal peptide prediction, as represented in Figure 3E. These correspond to
 1016 the following sequences: for *MGD* α , the sequence retrieved from Phatr3_J14125.t1 starting
 1017 from the second methionine (MCKL to TRTS, 460 amino acid); for *MGD* β , the whole
 1018 sequence from Phatr3_J54168.t1 (MVWS to LNNK, 559 amino acid); for *MGD* γ , the whole
 1019 sequence from Phatr3_J9619.t1 (MATG to EPSR, 618 amino acids).

1020 **MGD phylogenetic analysis.**

1021 All MGD protein sequences used for phylogenetic analysis were retrieved from the
 1022 National Centre for Biotechnology Information (NCBI) (Supplemental Table 1) after a protein
 1023 BLAST search using *MGD* α (XM_002181649.1), *MGD* β (XM_002186319) and *MGD* γ

1024 (XM_002176764.1) as queries. Sequences from different phyla covering the biodiversity of
1025 plastid-containing eukaryotes, including Archaeplastida and Stramenopiles were retained,
1026 with at least two species per phylum. The dataset was manually curated to validate the
1027 presence of full-length sequences with conserved MGD domains including residues at the
1028 active site. Alignment was performed with a customized pipeline in NGphylogeny.fr
1029 (Lemoine et al., 2019) using the MUSCLE v3.8.1551 software (Edgar, 2004). The
1030 ambiguously aligned regions were curated using the Block Mapping and Gathering with
1031 Entropy (BMGE v1.12_1) software (Criscuolo and Gribaldo, 2010) implemented in
1032 NGphylogeny.fr using default settings. Preliminary phylogenetic trees were inferred using
1033 FastME with 2,500 bootstrap pseudoreplicates. MEGA X v10.0.5 software (Kumar et al.,
1034 2018) was fed with the aligned and curated data set. The best evolutionary model was
1035 evaluated, and a Maximum Likelihood phylogenetic analysis was performed. To define the
1036 best evolutionary model, MEGA X was used to compare the 56 models implemented. The
1037 LG+G model (Le et al., 2008) was chosen (lowest Bayesian Information Criterion (BIC)
1038 score). The phylogeny was inferred by Maximum Likelihood with 5,000 bootstrap
1039 pseudoreplicates. The tree with the highest log likelihood (-12414.63) was retained. Initial
1040 trees for the heuristic search were obtained by applying the BioNJ method (Gascuel, 1997) to
1041 a matrix of pairwise distances estimated using a JTT model. A discrete Gamma distribution
1042 was used to model evolutionary rate differences among sites (five categories (+G, parameter =
1043 0.7081)). The best phylogenetic tree was then retrieved in Newick format and visualised by
1044 iTOL (v5) (Letunic and Bork, 2021).

1045 **Cultivation of *Phaeodactylum tricornerutum* cells**

1046 *Phaeodactylum tricornerutum* (Pt1) Strain 8.6 CCAP 1055/1 (CCMP2561- Culture
1047 Collection of Marine Phytoplankton, now known as NCMA: National Center for Marine
1048 Algae and Microbiota) was used to generate overexpressing and KO lines. Cells were
1049 maintained and grown in a modified ESAW (Enriched Seawater, Artificial Water) medium
1050 (NaCl 362.7 mM; Na₂SO₄ 25 mM; KCl 8.03 mM; NaHCO₃ 2.067 mM; KBr 0.725 mM;
1051 H₃BO₃ 0.372 mM; NaF 0.0657 mM; MgCl₂ 47.18 mM; CaCl₂ 9.134 mM; SrCl₂ 0.082 mM;
1052 Na₂-glycerophosphate 21.8 μM; Na₂SiO₃ 105.6 μM; disodium ethylenediaminetetraacetate
1053 dehydrate (Na₂EDTA) 14.86 μM; Fe(NH₄)₂(SO₄)₂ 5.97 μM; FeCl₃ 0.592 μM; MnSO₄ 2.42
1054 μM; ZnSO₄ 0.254 μM; CoSO₄ 0.0569 μM; Na₂MoO₄ 0.52 μM; H₃BO₃ 61.46 μM; Na₂SeO₃
1055 10 nM; biotin (vitamin H) 8.18 nM; cobalamin (vitamin B₁₂) 2.94 nM; thiamine (vitamin B₁)
1056 0.594 μM) (Falciatore et al., 2000) using ten times enriched nitrogen and phosphate sources

1057 (10N10P containing 5.49 mM NaNO₃ and 0.224 mM NaH₃PO₄) (Abida et al., 2015). Culture
1058 in liquid media was performed at 20°C, on a 12:12 light (40 μmol photons.m⁻².sec⁻¹) dark
1059 cycle, under gentle agitation, in an INFORS HT Multitron Pro incubator. Cell concentration
1060 was kept between 0.5×10⁶ and 10×10⁶ cell.mL⁻¹. For nitrogen starvation, *P. tricornutum* cells
1061 were cultivated in triplicates in 50 mL of ESAW 10N10P in 250-mL flasks until a
1062 concentration of 3-4×10⁶ cell.mL⁻¹. After centrifugation at 1,500 g for 10 minutes and
1063 removal of supernatant, cells were washed with 50 mL of ESAW 0N10P medium, and
1064 transferred into to 250-mL flasks containing 50 mL of ESAW 0N10P, at a concentration of
1065 2.5-3×10⁶ cell.mL⁻¹. Cultures on solid medium were performed with ESAW 10N10P
1066 complemented with Agar (1% m/v), at 20°C under continuous light (50 μmol photons.m⁻².s⁻¹),
1067 in a MLR-352-PE Climate Chamber. To guarantee axenic conditions, all media were
1068 supplemented with carbenicillin disodium salt (0.237 μM). For culture of mutant lines
1069 transformed with CRISPR-Cas9 vectors and of overexpressing lines, zeocin (Invitrogen; 0.07
1070 μM) was added to the media. For each analysis, three independent lines were cultivated.
1071 Growth was evaluated by cell counting using a TECAN infinite M1000Pro plate reader and
1072 determined following the equation $y = 0.01834 x + 0.03758$ with y the absorbance at 730 nm
1073 and x the number of cells in million per milliliter (Conte et al., 2018).

1074 **DNA and RNA extractions.**

1075 Genomic DNA extraction from *P. tricornutum* was performed on a cell pellet
1076 corresponding to 100×10⁶ cells, suspended in 500 μL of lysis buffer (250 mM Tris pH 8.2,
1077 100 mM EDTA, 2% SDS v/w, 350 mM NaCl) and transferred to a 1.5-mL tube. The sample
1078 was incubated 15 seconds at 60°C under agitation (450 rpm) using a thermobloc. 500 μL of
1079 cold phenol:chloroform:isoamyl alcohol (25:24:1) were added to the sample and
1080 homogenized gently. The sample was centrifuged at 13,000×g at 4 °C for 10 minutes. 300 μL
1081 of the upper aqueous phase was collected. A second extraction was performed by addition of
1082 one volume of chloroform:isoamyl alcohol (24:1), and 200 μL of the upper aqueous phase
1083 was collected. DNA was precipitated with 30 μL of sodium acetate 3 M pH 5 and three
1084 volumes of ethanol. After a 20-minute incubation at -20°C and a 10-minute centrifugation at
1085 13,000×g at 4°C the precipitated DNA was washed with 700 μL of ethanol 70% v/v,
1086 centrifuged at 13,000×g at 4°C for 5 minutes and air dried at room temperature. DNA was
1087 resuspended in DNase-free water. Concentration and purity were evaluated using a
1088 NanoDrop2000 (ThermoFisher Scientific). For RNA extraction from *P. tricornutum* cells, 1.5
1089 mL of TRI Reagent (Sigma-Aldrich) were added to a frozen cell pellet corresponding to

1090 150×10^6 cells. The sample was first vigorously mixed and incubated 5 minutes at 60°C , a
1091 process repeated twice for a thorough cell lysis. $300 \mu\text{L}$ of chloroform were added to the
1092 sample then vigorously mixed by tube inversion. The sample was incubated at room
1093 temperature for 15 minutes before centrifugation at $11,000 \times g$ at 4°C for 15 minutes for phase
1094 separation. About $600 \mu\text{L}$ of the upper aqueous phase was transferred to a Phasemaker tube
1095 (ThermoFisher Scientific). A volume of 1-bromo-3-chloropropane corresponding to one fifth
1096 of the transferred aqueous phase was added and mixed vigorously. The sample was incubated
1097 for 3 minutes at room temperature before centrifugation at $16,000 \times g$ at 4°C for 5 minutes.
1098 After centrifugation, 1 mL of the aqueous phase was collected. RNA was precipitated by
1099 addition of one volume of isopropanol and a 30-min incubation at room temperature. RNA
1100 was pelleted down by centrifugation at $11,000 \times g$ at 4°C for 10 minutes and further washed in
1101 1 mL of ethanol 75% v/v, followed by a centrifugation at $7,500 g$ at 4°C for 10 minutes. The
1102 pellet was left to air dry for 30 minutes at room temperature. RNA was resuspended in $35 \mu\text{L}$
1103 of RNA-free water. A first DNase treatment was applied using the Invitrogen Ambion
1104 TURBO DNA-free kit following manufacturer's instructions. Cleaning and purification of the
1105 sample was then achieved using the RNeasy MinElute Cleanup kit (Qiagen) following
1106 manufacturer's instructions. RNA concentration and quality was assessed using a
1107 NanoDrop2000 (ThermoFisher Scientific) and by electrophoresis. RNA was stored at -80°C
1108 until use.

1109 ***MGD* gene expression analysis by reverse transcription quantitative polymerase chain** 1110 **reaction (RT-qPCR).**

1111 This method is detailed in supplemental material.

1112 **Heterologous expression of *MGD* isoforms in yeast**

1113 This method is detailed in supplemental material.

1114 **Construction of *MGD* overexpressing lines.**

1115 Full-length coding sequence of *MGD α* , *MGD β* and *MGD γ* genes were amplified by
1116 PCR using cDNA derived from *P. tricornutum* as template and primers detailed in
1117 Supplemental Table 4. The PCR products were cloned into pEASY-T1 simple (TransGen,
1118 Beijing) vector for DNA sequencing. Sequence-confirmed *MGD α* , *MGD β* and *MGD γ* genes
1119 were excised from T-cloning vector with BamHI and XhoI in the case of *MGD α* , BamHI and
1120 Sall for *MGD β* , and NheI and Sall for *MGD γ* . Each gene was in-frame inserted upstream the

1121 eGFP coding sequence in the multi cloning site of pPha-CG vector (GenBank AF219942)
1122 (Zaslavskaja et al., 2000). The obtained pPhaT1 vectors harboring *MGD α -eGFP*, *MGD β -*
1123 *eGFP* and *MGD γ -eGFP* were linearized with ScaI and introduced in *P. tricornutum* cells by
1124 electroporation as previously described (Zhang and Hu, 2014). *Phaeodactylum tricornutum*
1125 PetC (Liu et al., 2016) and symbiont Derlin1-2 (sDer1-2) (Hempel et al., 2009) were used as
1126 markers of thylakoid and PPM, respectively. For co-localization studies, three pairs of
1127 linearized plasmids, pPhaT1-MGD α -eGFP and pPhaT1-PetC-mRFP, pPhaT1-MGD β -eGFP
1128 and pPhaT1-sDer1-2-mRFP, pPhaT1-MGD γ -eGFP and pPhaT1-mRFP were co-transformed
1129 into *P. tricornutum*, respectively. Fluorescence of eGFP and plastid auto-fluorescence were
1130 excited at 488 nm, and were detected with two photomultiplier tubes at 500–520 nm and 625–
1131 720 nm respectively using a Leica TCS SP8 laser scanning confocal microscope. Gain was set
1132 at 800. The mRFP fluorescence was excited at 552 nm and detected at 580–600 nm.

1133 **Construction of CRISPR-Cas9 single and double knockout lines.**

1134 Single guide RNAs (sgRNA) were designed using the PhytoCrispex online tool
1135 (Rastogi et al., 2016), choosing NGG as Protospacer Adjacent Motif (PAM) sequence, and
1136 allowing CRISPR targets to start with any nucleotide. sgRNA were selected based on their
1137 proximity to START codons and active site residues. sgRNA forward and reverse sequences
1138 preceded with the short nucleotide sequences TCGA and AAAC respectively are shown in
1139 Supplemental Table 3. The sgRNA sequences were inserted in the pKSdiaCas9_sgRNA-zeo
1140 vector, kindly provided by Cécile Giustini (LPCV, Grenoble), derived from pKS
1141 diaCas9_sgRNA vector (Addgene). *P. tricornutum* cells were transformed with corresponding
1142 vectors by particle bombardment as described previously (Falciatore et al., 1999; Dolch et al.,
1143 2017b). For MGD β /MGD γ double KO, *mgd β 1* single KO was chosen as genetic
1144 background. The sgRNA 619h (Supplemental Table 3) was cloned between the two BsaI sites
1145 of the pYS31 episome, derived from PtPuc3_diaCas9_sgRNA episome (Addgene) with a
1146 blasticidin resistance cassette instead of a zeocin. The PtPuc3_diaCas9_sgRNA was digested
1147 by SalI and StuI in order to excise the zeocin resistance gene. The Blasticidin S deaminase
1148 (BSD) coding sequence, conferring the resistance to blasticidin, was amplified using the
1149 primers oIYS40 (5'-GAAAAATTAACCAAGATGGCCAAGCCTTTGTCTCA-3') and
1150 oIYS41 (5'-CCCAGATCTCCGAGGTCAGCCCTCCACACATAAC-3'). 50 ng of the
1151 vector were used for Gibson assembly with a 2-fold molar excess of the BSD fragment. The
1152 episome was transformed by bacterial conjugation (Karas et al., 2015). Colony PCR (Phire
1153 Plant Direct PCR Master Mix, F160, ThermoFisher Scientific) were conducted to analyse the

1154 mutation profile by sequencing. Primers were designed for each sgRNA to amplify a region of
1155 about 500 nucleotides surrounding the cutting site of the Cas9 protein (Supplemental Table
1156 3). In some rare cases, no amplification was obtained and primers amplifying a larger region
1157 were used to reveal a possible very large deletion. Three different kinds of edition profiles
1158 were usually obtained: either no edition (pure WT colony), several different editions (mosaic
1159 colony), or one edition only (pure mutant colony). For the analysis of mosaic colonies, TIDE
1160 (Tracking of Indels by Decomposition v3.3.0) (Brinkman et al., 2014) and ICE (Inference of
1161 CRISPR Edits v2, Synthego.com) online software were used to decompose the
1162 chromatograms and estimate the edition profile of the colony. Whenever a fraction of cells
1163 within the colony presented insertions/deletions (INDELS), a strain isolation was attempted.
1164 The equivalent of 200 cells from a mosaic colony were spread on a plate with carbenicillin
1165 and zeocin. After three to six weeks, the same steps of PCR and sequencing analyses were
1166 conducted. If pure edited colonies were obtained, the screening process was stopped.
1167 Otherwise, subsequent isolation attempts were performed.

1168 **Bacterial conjugation**

1169 This method is detailed in supplemental material.

1170 **Glycerolipid analyses.**

1171 Glycerolipids were extracted according to the Folch method (Folch et al., 1957) as
1172 described previously (Abida et al., 2015). One tenth of total glycerolipid extract was used to
1173 quantify the total FA content, after methanolysis. Lipids were solubilised with 1 mL of
1174 chloroform and a 100- μ L aliquot was transferred to a 10-mL crim cap vial (Gerstel).
1175 Methanolysis was performed by a MultiPurpose Sampler (MPS, Gerstel). Briefly, 5 μ L of
1176 15:0 FA (a FA with a 15-carbon chain length) at 1 mg.mL⁻¹ was added as internal standard
1177 and 3 mL of methanolysis buffer (methanol/sulphuric acid; 40:1 v/v) were added, vigorously
1178 mixed, and incubated at 80°C for one hour, yielding FAs methyl esters (FAME). Reaction
1179 was stopped with the addition of 3 mL of distilled water and 3 mL of hexane. FAME retrieved
1180 from the upper phase were suspended in hexane and analyzed by Gas Chromatography
1181 coupled to Flame Ionization Detection (GC-FID, Perkin Elmer Clarus 580 equipped with a
1182 30-m long cyanopropyl polysilphenesiloxane column, 0.22 mm diameter). FAME were
1183 identified by comparison of their retention times with those of standards (Sigma) and
1184 quantified by the surface peak method using 15:0 for calibration. Extraction and
1185 quantification were performed with three biological replicates. Glycerolipids were then

1186 analyzed and quantified by high-pressure liquid chromatography-tandem mass spectrometry
1187 (HPLC-MS/MS) as previously described (Dolch et al., 2017b), with appropriate standard
1188 lipids (Jouhet et al., 2017). Lipid extracts corresponding to 25 nmol of total FAs were
1189 dissolved in 100 μ L of chloroform/methanol [2/1, v/v] containing 125 pmol of each internal
1190 standard. Phosphatidylethanolamine (PE) 18:0-18:0 and diacylglycerol (DAG) 18:0-22:6 from
1191 Avanti Polar Lipid, and sulfoquinovosyldiacylglycerol (SQDG) 16:0-18:0 extracted from
1192 spinach thylakoid (Deme et al., 2014) and hydrogenated (Buseman et al., 2006) were used as
1193 internal standards. Lipids were then separated by HPLC and quantified by MS/MS. Lipid
1194 classes were separated using an Agilent 1200 HPLC system using a 150 mm \times 3 mm (length
1195 \times internal diameter) 5 μ m diol column (Macherey-Nagel), at 40°C. The mobile phases
1196 consisted of hexane/isopropanol/water/1 M ammonium acetate, pH 5.3 [625/350/24/1, v/v]
1197 (A) and isopropanol/water/1 M ammonium acetate, pH 5.3 [850/149/1, v/v] (B). The injection
1198 volume was 20 μ L. After 5 minutes, the percentage of B was increased linearly from 0 to
1199 100% in 30 minutes and kept at 100% for 15 min. This elution sequence was followed by a
1200 return to 100% A in 5 minutes and an equilibration for 20 minutes with 100% A before the
1201 next injection, leading to a total runtime of 70 minutes. The flow rate of the mobile phase was
1202 200 μ L.min⁻¹. The distinct glycerophospholipid classes were eluted successively as a function
1203 of the polar head group. Mass spectrometric analysis was performed on a 6460 triple
1204 quadrupole mass spectrometer (Agilent) equipped with a Jet stream electrospray ion source
1205 under following settings: drying gas heater at 260°C, drying gas flow at 13 L.min⁻¹, sheath gas
1206 heater at 300°C, sheath gas flow at 11 L.min⁻¹, nebulizer pressure at 25 psi, capillary voltage
1207 at \pm 5000 V and nozzle voltage at \pm 1,000 V. Nitrogen was used as collision gas. The
1208 quadrupoles Q1 and Q3 were operated at widest and unit resolution respectively.
1209 Phosphatidylcholine (PC) and diacylglyceryl hydroxymethyltrimethyl- β -alanine (DGTA)
1210 analyses were carried out in positive ion mode by scanning for precursors of m/z 184 and 236
1211 respectively at a collision energy (CE) of 34 and 52 eV. SQDG analysis was carried out in
1212 negative ion mode by scanning for precursors of m/z -225 at a CE of -56 eV.
1213 Phosphatidylethanolamine (PE), phosphatidylinositol (PI), phosphatidylglycerol (PG),
1214 monogalactosyldiacylglycerol (MGDG) and digalactosyldiacylglycerol (DGDG)
1215 measurements were performed in positive ion mode by scanning for neutral losses of 141 Da,
1216 277 Da, 189 Da, 179 Da and 341 Da at CEs of 20 eV, 12 eV, 16 eV, 8 eV and 8 eV,
1217 respectively. DAG and triacylglycerol (TAG) species were identified and quantified by
1218 multiple reaction monitoring (MRM) as singly charged ions [M+NH₄]⁺ at a CE of 16 and 22

1219 eV respectively. Quantification was done for each lipid species by multiple reaction
1220 monitoring (MRM) with 50 ms dwell time with the various transitions previously recorded
1221 (Abida et al., 2015). Mass spectra were processed using the MassHunter Workstation software
1222 (Agilent) for identification and quantification of lipids. Lipid amounts (pmol) were corrected
1223 for response differences between internal standards and endogenous lipids and by comparison
1224 with a qualified control (Jouhet et al., 2017).

1225 **Epifluorescence microscopy.**

1226 Cells were observed using an epifluorescence microscope (Zeiss AxioScope A1)
1227 equipped with a Zeiss AxioCam MRc. Images were captured using a Zeiss EC Plan
1228 Neofluar 100×/1.3 oil immersion objective. Chlorophyll autofluorescence and Nile Red
1229 fluorescence in lipid droplets were visualized with Zeiss filter set 16 (BP 485/20, FT510,
1230 LP515) as described previously (Dolch et al., 2017b).

1231 **Laser scanning confocal microscopy.**

1232 Confocal microscopy was performed with a microscope Zeiss LSM880 equipped with
1233 a 63×/1.4 oil-immersed Plan-Apochromat objective, running Zen 2.3 SP1 acquisition software
1234 (Platform μ Life, IRIG, LPCV). Chlorophyll autofluorescence and eGFP fluorescence were
1235 excited using the 488 nm ray of an Argon Multiline laser, and were detected at 600-730 nm
1236 and 500-531 nm, respectively, using a GaAsP detector. Laser intensity was set at 5% for
1237 visualization of chlorophyll, and at 2%, 3%, and 5% for eGFP visualisation according to the
1238 eGFP signal intensity in each cell. On all images, gain was set at 800 for chlorophyll and for
1239 MGD β , 850 for MGD γ and 900 for MGD α . 'Pseudo brightfield' images were acquired in
1240 parallel by differential interference contrast (DIC), using the 488 nm laser ray at 0.6% as light
1241 source and a photomultiplier tube detector for transmitted light (T-PMT), with gain set at 500.
1242 Z-stacks containing consecutive images with a distance of 0.47 nm were obtained for each
1243 cell. For co-localization analyses, fluorescence of eGFP and plastid autofluorescence were
1244 excited at 488 nm, and were detected with two photomultiplier tubes at 500–520 nm and 625–
1245 720 nm respectively using a Leica TCS SP8 laser scanning confocal microscope. The mRFP
1246 fluorescence was excited at 552 nm and detected at 580–600 nm.

1247 **Transmission electron microscopy.**

1248 Samples were prepared as previously described (Flori et al., 2018). Ultrathin sections
1249 (50-70 nm) were prepared with a diamond knife on a PowerTome ultramicrotome (RMC
1250 products, Tucson, AZ, USA) and collected on 200- μ m carbon-coated gold grids. Samples

1251 were visualized by scanning transmission electron microscopy (STEM) using a MERLIN
1252 microscope (Zeiss, Oberkochen, Germany) set up at 30 kV and 240 pA or using a FEI tecnai
1253 OSIRIS microscope (Hillsboro, Oregon, USA) set up at 200 kV and ~300 pA.

1254 **Accession numbers.**

1255 Species and associated MGD protein sequences used for phylogeny (NCBI accession
1256 numbers): *Vitrella brassicaformis* (CEM14063.1, CEM34604.1); *Gregarina niphandrodes*
1257 (XP_011130274.5); *Pseudo-nitzschia multistriata* (VEU37191.1, VEU39505.1);
1258 *Fragilariopsis cylindrus* (OEU16730.1, OEU06991.1); *Fistulifera solaris* (GAX23913.1,
1259 GAX10983.1, GAX09439.1); *Phaeodactylum tricorutum* (XP_002186355, XP_002176800,
1260 XP_002181685), *Thalassiosira pseudonana* (XP_002295865, XP_002293576.1,
1261 XP_002294242); *Ectocarpus siliculosus* (CBJ28381.1, CBJ28372.1, CBN79326.1);
1262 *Aureococcus anophagefferens* (XP_009033523.1, XP_009038839.1, XP_009035780.1);
1263 *Microchloropsis/Nannochloropsis gaditana* (XP_005855249.1); *Coccomyxa subellipsoidea*
1264 (XP_005651388.1); *Chlorella sorokiniana* (PRW45643.1); *Ostreococcus tauri*
1265 (OUS42062.1); *Chlamydomonas reinhardtii* (PNW74102.1); *Monoraphidium neglectum*
1266 (XP_013903204.1); *Raphidocelis subcapitata* (GBF88428.1); *Arabidopsis thaliana*
1267 (NP_194906.1, NP_565352.1, NP_568394.2); *Coffea arabica* (XP_027069738.1,
1268 XP_027084801.1); *Spinacia oleracea* (XP_021867203.1, XP_021852153.1); *Nymphaea*
1269 *colorata* (XP_031495579.1, XP_031480803.1); *Amborella trichopoda* (XP_006852865.1,
1270 XP_006845407.1); *Brachypodium distachyon* (XP_010238179.1, XP_003573675.2,
1271 XP_003570331.1); *Oryza sativa* (XP_015611851.1, XP_015649135.1, XP_015627712.1);
1272 *Physcomitrium patens* (XP_024403181.1, XP_024404076.1); *Selaginella moellendorffii*
1273 (XP_024533920.1); *Cyanidioschyzon merolae* (XP_005536420.1); *Cyanidiococcus*
1274 *yangmingshanensis* (KAF6002292.1); *Gracilariopsis chorda* (PXF49423.1, PXF42956.1,
1275 PXF46603.1); *Blastochloris viridis* (WP_055036643.1). *Phaeodactylum tricorutum* MGD
1276 gene accession numbers: *MGD α* (PHATR DRAFT_14125, NCBI; Phatr3_J14125, Ensembl);
1277 *MGD β* (PHATR_43938, NCBI; Phatr3_J54168, Ensembl); *MGD γ* (PHATR DRAFT_9619,
1278 NCBI; Phatr3_J9619, Ensembl).

1279

1280 **Supplementary Data**

1281 Supplementary Figure S1. Schematic view of *MGD* genes in *P. tricornutum*.

1282 Supplementary Figure S2. Phylogenetic analysis of MGDs from plastid-containing
1283 eukaryotes.

1284 Supplementary Figure S3. Multiple sequence alignment of *Arabidopsis thaliana* AtMGD1,
1285 AtMGD2 and AtMGD3 and *Phaeodactylum tricornutum* MGD α , MGD β and MGD γ proteins

1286 Supplementary Figure S4. Protein models of *P. tricornutum* MGDs.

1287 Supplementary Figure S5. Heterologous expression of *Phaeodactylum tricornutum* MGD
1288 isoforms in yeast.

1289 Supplementary Figure S6. Visualization of MGD-eGFP proteins overexpression by
1290 immunoblot.

1291 Supplementary Figure S7. Multiple observations of MGD isoforms fused to eGFP in *P.*
1292 *tricornutum* cells.

1293 Supplementary Figure S8. Colocalization of MGD isoforms fused to eGFP with subcellular
1294 protein makers fused to RFP.

1295 Supplementary Figure S9. Selected knock-out MGD α , MGD β and MGD γ mutants, generated
1296 by CRISPR-Cas9 editing.

1297 Supplementary Figure S10. Analysis of loss of heterozygosity and stability of MGD KO lines.

1298 Supplementary Figure S11. Growth curves of MGD mutant and overexpressing strains
1299 compared to the WT.

1300 Supplementary Figure S12. Effect of MGDs mutations on photosynthetic properties under
1301 high and moderate light stresses.

1302 Supplementary Figure S13. Cell morphology of KO and overexpressing lines.

1303 Supplementary Figure S14. *P. tricornutum* cell ultrastructure in MGD overexpressing lines.

1304 Supplementary Figure S15. Quantitative analysis of fatty acid and glycerolipid content in
1305 MGD KO mutants.

1306 Supplementary Figure S16. Quantitative analysis of fatty acid and glycerolipid content in
1307 MGD overexpressing lines.

1308 Supplementary Figure S17. Impact of the overexpression of MGD isoforms on the molecular
1309 species constituting MGDG, DGDG, SQDG and PG in *P. tricornutum*.

1310 Supplementary Figure S18. Impact of the overexpression of MGD isoforms on the molecular
1311 species of endomembrane glycerolipids in *P. tricornutum*.

1312 Supplementary Figure S19. Impact of MGD α mutations on the molecular species constituting
1313 PC, PE, DGTA and DAG in *P. tricornutum*.

1314 Supplementary Figure S20. Impact of MGD β mutations on the molecular species constituting
1315 PC, PE, DGTA and DAG in *P. tricornutum*.

1316 Supplementary Figure S21. Impact of MGD γ mutations on the molecular species constituting
1317 PC, PE, DGTA and DAG in *P. tricornutum*.

1318 Supplementary Figure S22. *MGD* relative gene expression in mutant lines compared to the
1319 WT, grown three days in 10N10P medium.

1320 Supplementary Figure S23. Wild type and mutant *P. tricornutum* growth during nitrogen
1321 limitation.

1322 Supplementary Figure S24. Accumulation of non-polar lipids during nitrogen limitation.

1323 Supplementary Figure S25. Observation of cell morphology and TAG accumulation by
1324 epifluorescence microscopy.

1325 Supplementary Figure S26. Quantitative analysis of FA and glycerolipid content in MGD
1326 mutant lines upon nitrogen starvation.

1327 Supplementary Figure S27. Impact of MGD mutations on the molecular species constituting
1328 PC, PE, DGTA, DAG and TAG in nitrogen-deprived *P. tricornutum* cells.

1329 Supplementary Figure S28. Selected *MGD β /MGD γ* double knock-out mutants, generated by
1330 CRISPR-Cas9 editing.

1331 Supplementary Figure S29. Growth curves of *MGD β /MGD γ* double mutants compared to the
1332 WT.

1333 Supplementary Figure S30. Quantitative analysis of glycerolipid content in *MGD β /MGD γ*
1334 double mutant overexpressing lines

1335 Supplementary Figure S31. Impact of *MGD β /MGD γ* double mutations on MGDG, DGDG,
1336 SQDG and PG molecular species.

1337 Supplementary Figure S32. Impact of MGD β /MGD γ double mutations on PC, DGTA, PE and
1338 TAG molecular species.

1339 Supplementary Table S1. *Phaeodactylum tricornutum* MGD entries in NCBI, EnsemblProtist
1340 and Uniprot databases.

1341 Supplementary Table S2. Taxonomic classification of the MGD proteins used for
1342 phylogenetic analyses.

1343 Supplementary Table S3. Single guide RNA list.

1344 Supplementary Table S4. Primers list.

1345 Supplementary Methods

1346

1347 **Funding**

1348 N.G. was supported by a PhD grant from INRAE. A.A., F.C., E.D., N.G., S.D.G., V.G, J.J.,
1349 D.P., G.S.L., Y.S. and E.M were supported by the French National Research Agency (GRAL
1350 Labex ANR-10-LABEX-04, EUR CBS ANR-17-EURE-0003, ANR AlpAlga ANR-20-CE02-
1351 0020, ANR DIM ANR-21-CE02-0021, PEPR Algadvance A-22-PEBB-0002) and IDEX
1352 Université Grenoble-Alpes (Glyco@Alps Cross-Disciplinary Program; Grant ANR-15-IDEX-
1353 02). H.H. and Y.G. were supported by the National Natural Science Foundation of China
1354 (41976119, 91751117). A.A., E.M., H.H. and Y.G. were supported by a CEA-CAS bilateral
1355 program. Imaging was performed at the μ Life Platform of IRIG. Lipid analyzes were
1356 performed at the LIPANG platform supported by the Rhône-Alpes Region, the FEDER funds,
1357 and ANR (GRAL Labex ANR-10-LABEX-04, EUR CBS ANR-17-EURE-0003).

1358 **Acknowledgements.**

1359 Authors wish to thank Guillaume Allorent, Baptiste Doussot and Cécile Giustini for technical
1360 assistance and Sylvaine Roy for lipid data analysis software development.

1361 **Author contributions.**

1362 N.G. and Y.S. contributed equally to this work. N.G. and Y.S have contributed equally to this
1363 work. N.G. has performed most experimental works. Y.S. has generated multiple KO lines
1364 and contributed to imaging and lipidomic analyses. F.C. has provided guidance on CRISPR-
1365 Cas9 strategy. V.G. and G.S.L. have provided technical assistance on lipidomic analyses and
1366 STEM, respectively. D.P. has contributed to the design of photosynthetic measurements and

1367 their analyses. A.A., E.D., S.D.G. and M.M. have provided assistance on heterologous
1368 expression in yeast. H.H. and Y.G. have developed overexpressing lines. D.L.M. and J.S.
1369 have contributed to confocal imaging. J.J. has provided guidance in the analysis of lipidomic
1370 profiles. M.S and M.C. have performed lipid analyses. J.S., A.A. and E.M. have contributed
1371 to the design of experiments and their analyses. All authors have contributed to the writing of
1372 the manuscript.

1373

1374 **Tables**

1375

1376 **Figure Legends**

1377 **Figure 1: Chimeric origin of the secondary plastid in diatom.** The scheme shows a
1378 fusiform cell of *Phaeodactylum tricornutum*. The plastid is limited by four membranes. The
1379 epiplastidial membrane (EpM), shown in blue, is continuous with the outer nuclear envelope
1380 (oNE). The periplastidial membrane (PPM), shown in red, is in tight contact with the inner
1381 nuclear envelope (iNE). The outer and inner envelope membranes (iEM and oEM), shown in
1382 light green, are tightly apposed. A specific periplastidial compartment relies on the detection
1383 of a “blob”-like structure observed by confocal and transmission electron microscopy. In
1384 confocal microscopy, the precursors of blob-residing proteins fused to GFP cross only the
1385 EpM and PPM and form fluorescent spots in the middle of the plastid (Lang et al., 1998;
1386 Kilian and Kroth, 2005; Huang et al., 2024). In electron microscopy, the blob is marked by
1387 the development of a vesicular network (VN) budding from the PPM (Flori et al., 2016). Each
1388 membrane surrounding the plastid derive from distinct evolutionary origin. The oEM and iEM
1389 stem from the chloroplast of the engulfed red algae. The PPM potentially derive from the
1390 endomembrane systems of the red algal symbiont and the EpM from that of the host cell,
1391 respectively. M, mitochondrion; N, nucleus; P, plastid; Thyl: thylakoids.

1392 **Figure 2: Subcellular localization of MGD isoforms fused to eGFP in *P. tricornutum*.**
1393 Detection of MGD-eGFP expressed in transgenic *P. tricornutum* lines with wild type as a
1394 negative control. For each cell, bright field, chlorophyll fluorescence and eGFP signal are
1395 shown, with a merge of all images. eGFP excitation was achieved at either 5 % or 3 % of laser
1396 power, as indicated. The WT control shows low unspecific detection of fluorescence from
1397 thylakoid photosystems. An intraplastidial localization was obtained when MGD α -eGFP is
1398 expressed. With MGD β -eGFP, the eGFP signal corresponds to the blob-like structure,

1399 indicating a localization of MGD β at the level of the periplastidal compartment, and possibly
1400 the PPM. MGD γ -eGFP is detected mainly in the ER, and at lower level in the EpM. Scale bar:
1401 5 μ m. Similar independent observations are provided in Supplemental Figure 7. MGD,
1402 monogalactosyldiacylglycerol synthase; WT, wild type.

1403 **Figure 3: Subcellular localization of the blob-like structure, associated to MGD β -eGFP**
1404 **fluorescence.** Observation of blob-like structure localization by confocal microscopy (A) and
1405 transmission electron microscopy (B). *P. tricornutum* transgenic lines overexpressing MGD β -
1406 eGFP were used to monitor the appearance of blob-like structures in the cell (A). Cells 1 and
1407 2 show a scattered eGFP signal, cells 3-6 show a blob-like structure at the central constriction
1408 of dividing plastids, cell 7 shows a blob-like structure extending from one plastid to the other
1409 inside a dividing cell, and cell 8 shows two blob-like structures facing each other on distinct
1410 plastids during cytokinesis. Cell 6 was also shown in Figure 2. Observation of a wild type *P.*
1411 *tricornutum* cell containing two plastids following plastid division inside the cell (B).
1412 Cleavage furrows are visible at both end of the cell. The blob-like structure (outlined in blue)
1413 was observed connecting the two plastids at two magnifications. BLS, blob-like structure (or
1414 blob); G, Golgi apparatus; M, mitochondrion; MGD, monogalactosyldiacylglycerol synthase;
1415 N, nucleus; P and P', plastids.

1416 **Figure 4: *P. tricornutum* cell ultrastructure in MGD KO lines determined by scanning**
1417 **transmission electron microscopy** A WT and KO *mgdaj3* and *mgdyh2* lines were cultured in
1418 parallel. In a separate experiment, a WT and a *mgd β c3* mutant were cultured in parallel. Cell
1419 ultrastructure is shown in each mutant with corresponding WT control on the left. No impact
1420 of MGD α , MGD β or MGD γ KO could be observed at the level of membrane compartments,
1421 including plastids. BLS outlined in yellow was observed at two magnifications. BLS, blob-
1422 like structure; G, Golgi, KO, knock-out; M, mitochondria; MGD,
1423 monogalactosyldiacylglycerol synthase; N, nucleus, P, plastid; WT, wild type.

1424 **Figure 5: Impact of MGD α mutations on MGDG, DGDG, SQDG and PG molecular**
1425 **species.** Lipid molecular species in MGDG, DGDG, SQDG and PG were analyzed by LC-
1426 MS/MS. Each result is the median of six biological replicates \pm min and max values. (*), *P*-
1427 *value* $< 5.10^{-2}$; (**), *P*-*value* $< 1.10^{-2}$; (***), *P*-*value* $< 1.10^{-3}$, based on an unpaired multiple t
1428 test. DGDG, digalactosyldiacylglycerol, MGDG, monogalactosyldiacylglycerol, PG,
1429 phosphatidylglycerol, SQDG, sulfoquinovosyldiacylglycerol.

1430 **Figure 6: Impact of MGD β mutations on MGDG, DGDG, SQDG and PG molecular**
1431 **species.** Lipid molecular species were analyzed by LC-MS/MS. The *mgd β c2* mutant contains
1432 a silent mutation and was used as a control. Each result is the median of six biological
1433 replicates \pm min and max values. (*), *P*-value $< 5.10^{-2}$; (**), *P*-value $< 1.10^{-2}$; (***), *P*-value
1434 $< 1.10^{-3}$, based on an unpaired multiple t test. DGDG, digalactosyldiacylglycerol, MGDG,
1435 monogalactosyldiacylglycerol, PG, phosphatidylglycerol, SQDG,
1436 sulfoquinovosyldiacylglycerol.

1437 **Figure 7: Impact of MGD γ mutations on MGDG, DGDG, SQDG and PG molecular**
1438 **species.** Lipid molecular species were analyzed by LC-MS/MS. Each result is the median of
1439 six biological replicates \pm min and max values. (*), *P*-value $< 5.10^{-2}$; (**), *P*-value $< 1.10^{-2}$;
1440 (***), *P*-value $< 1.10^{-3}$, based on an unpaired multiple t test. DGDG,
1441 digalactosyldiacylglycerol, MGDG, monogalactosyldiacylglycerol, PG, phosphatidylglycerol,
1442 SQDG, sulfoquinovosyldiacylglycerol.

1443 **Figure 8: Impact of MGD mutations on the molecular species constituting MGDG,**
1444 **DGDG, SQDG and PG in nitrogen-deprived *P. tricornutum* cells.** Lipid molecular species
1445 in MGDG, DGDG, SQDG and PG were analyzed by LC-MS/MS. Each result is the median
1446 of six biological replicates \pm min and max values. (*), *P*-value $< 5.10^{-2}$; (**), *P*-value $< 1.10^{-2}$;
1447 (***), *P*-value $< 1.10^{-3}$, based on an unpaired multiple t test. DGDG,
1448 digalactosyldiacylglycerol, MGDG, monogalactosyldiacylglycerol, PG, phosphatidylglycerol,
1449 SQDG, sulfoquinovosyldiacylglycerol.

1450 **Figure 9. Function of MGD α , MGD β and MGD γ in *P. tricornutum* cell.** The model
1451 summarizes all compartmentalization, structural and functional data obtained in this study. FA
1452 are synthesized in the stroma as 16:0- and 16:1-ACP (1), where they can be used for the
1453 synthesis of DAG (2a). Alternatively, FA are exported to extraplastidial membranes where
1454 they can serve for the synthesis of DAG (2b) and phospholipids such as PC and PE, or the
1455 betaine lipid DGTA (2c). PC and DGTA serve as platforms to produce 20:5 (2c), used inside
1456 the plastid (at an unknown membrane site) for the production of 20:5-containing DAG (2d).
1457 MGD α is localized at the level of thylakoids, and functional studies reported here are
1458 consistent with a utilization of DAG available inside the plastid (3 α), to form MGDG species
1459 rapidly desaturated (16:1 \rightarrow 16:4) by plastid desaturases (4 α). Based on obtained phenotype
1460 of KO and overexpressing lines, MGDG produced by MGD α is likely used by a specific DGD
1461 enzyme (5 α) to form DGDG. MGD β is localized at the level of the blob-like structure at the
1462 PPM, and functional studies reported here are consistent with a utilization of DAG either

1463 produced inside the plastid or with 20:5 acyls imported from the ER by a still uncharacterized
1464 Ω -pathway (3β). MGDG produced can be desaturated (16:1 \rightarrow 16:4) by plastid desaturases
1465 (4β). Based on obtained phenotype of KO and overexpressing lines, MGDG produced by
1466 MGD β is likely used by a specific DGD enzyme (5β) to form DGDG. MGD γ is localized at
1467 the periphery of the plastid, possibly partitioned between the ER and the EpM. Functional
1468 studies are consistent with a utilization of DAG containing FA freshly exported from the
1469 plastid (3γ), producing MGDG at a location far from plastid desaturases. Based on obtained
1470 phenotype of KO and overexpressing lines, MGDG produced by MGD γ is also likely
1471 associated to a specific DGD enzyme (5γ). In nitrogen shortage, MGD α protein level
1472 decreases drastically, and MGDG production relies on MGD β and MGD γ . When one MGD is
1473 knocked-out, compensation mechanisms rely on intact isoforms located in distinct membrane
1474 systems, requiring rapid inward and outward trafficking of glycerolipids including MGDG,
1475 through the four membranes limiting the plastid. KO, knock-out; DAG, diacylglycerol;
1476 DGTA, diacylglyceryl hydroxymethyltrimethyl- β -alanine; DGD, digalactosyldiacylglycerol
1477 synthase; DGDG, digalactosyldiacylglycerol; ER, endoplasmic reticulum; FA; fatty acid;
1478 MGD, monogalactosyldiacylglycerol synthase; MGDG, monogalactosyldiacylglycerol; PC,
1479 phosphatidylcholine.

1480 **Figure 10: Contribution of each MGD isoform to the accumulation of different pools of**
1481 **MGDG species.** The level of contribution of each MGD to the accumulation of each MGDG
1482 species is indicated in three different intensities of grey. Diacyl species corresponding to
1483 diacylglycerol (DAG) substrates deriving from fatty acids synthesized *de novo* in the stroma
1484 of the plastid are shown in green. Diacyl species corresponding to DAG substrates containing
1485 eicosapentaenoic acid (20:5) imported from the endoplasmic reticulum or endomembranes are
1486 shown in light brown. Diacyl species corresponding to MGDG molecular species desaturated
1487 inside the plastid are shown in dark green and dark brown. This table summarizes information
1488 we deduced from (1) the lipidome analysis of *P. tricornutum* knock-out and overexpressing
1489 lines, and (2) the range of medium chain MGDG fatty acyl species obtained after
1490 heterologous expression of isoforms in yeast. MGD, monogalactosyldiacylglycerol synthase;
1491 MGDG, monogalactosyldiacylglycerol.

1492

1493 **References.**

- 1494 Abida, H., Dolch, L.J., Mei, C., Villanova, V., Conte, M., Block, M.A., Finazzi, G., Bastien,
1495 O., Tirichine, L., Bowler, C., Rebeille, F., Petroustos, D., Jouhet, J., and Marechal, E.
1496 (2015). Membrane glycerolipid remodeling triggered by nitrogen and phosphorus
1497 starvation in *Phaeodactylum tricornutum*. *Plant Physiol* 167, 118-136.
- 1498 Almagro Armenteros, J.J., Tsirigos, K.D., Sonderby, C.K., Petersen, T.N., Winther, O.,
1499 Brunak, S., von Heijne, G., and Nielsen, H. (2019). SignalP 5.0 improves signal
1500 peptide predictions using deep neural networks. *Nat Biotechnol* 37, 420-423.
- 1501 Apt, K.E., Zaslavkaia, L., Lippmeier, J.C., Lang, M., Kilian, O., Wetherbee, R., Grossman,
1502 A.R., and Kroth, P.G. (2002). In vivo characterization of diatom multipartite plastid
1503 targeting signals. *J Cell Sci* 115, 4061-4069.
- 1504 Awai, K., Marechal, E., Block, M.A., Brun, D., Masuda, T., Shimada, H., Takamiya, K.,
1505 Ohta, H., and Joyard, J. (2001). Two types of MGDG synthase genes, found widely in
1506 both 16:3 and 18:3 plants, differentially mediate galactolipid syntheses in
1507 photosynthetic and nonphotosynthetic tissues in *Arabidopsis thaliana*. *Proc Natl Acad*
1508 *Sci U S A* 98, 10960-10965.
- 1509 Azadi-Chegeni, F., Thallmair, S., Ward, M.E., Perin, G., Marrink, S.J., Baldus, M.,
1510 Morosinotto, T., and Pandit, A. (2022). Protein dynamics and lipid affinity of
1511 monomeric, zeaxanthin-binding LHCII in thylakoid membranes. *Biophysical journal*
1512 121, 396-409.
- 1513 Benning, C., and Ohta, H. (2005). Three enzyme systems for galactoglycerolipid biosynthesis
1514 are coordinately regulated in plants. *Journal of Biological Chemistry* 280, 2397-2400.
- 1515 Benoiston, A.S., Ibarbalz, F.M., Bittner, L., Guidi, L., Jahn, O., Dutkiewicz, S., and Bowler,
1516 C. (2017). The evolution of diatoms and their biogeochemical functions. *Philos Trans*
1517 *R Soc Lond B Biol Sci* 372.
- 1518 Billey, E., Hafidh, S., Cruz-Gallardo, I., Litholdo, C.G., Jean, V., Carpentier, M.C., Picart, C.,
1519 Kumar, V., Kulichova, K., Marechal, E., Honys, D., Conte, M.R., Deragon, J.M., and
1520 Bousquet-Antonelli, C. (2021a). LARP6C orchestrates posttranscriptional
1521 reprogramming of gene expression during hydration to promote pollen tube guidance.
1522 *Plant Cell*.
- 1523 Billey, E., Magneschi, L., Leterme, S., Bedhomme, M., Andres-Robin, A., Poulet, L.,
1524 Michaud, M., Finazzi, G., Dumas, R., Crouzy, S., Laueffer, F., Fourage, L., Rebeille,
1525 F., Amato, A., Collin, S., Jouhet, J., and Marechal, E. (2021b). Characterization of the
1526 Bubblegum acyl-CoA synthetase of *Microchloropsis gaditana*. *Plant Physiol* 185, 815-
1527 835.
- 1528 Botte, C., Jeanneau, C., Snajdrova, L., Bastien, O., Imbert, A., Breton, C., and Marechal, E.
1529 (2005). Molecular modeling and site-directed mutagenesis of plant chloroplast
1530 monogalactosyldiacylglycerol synthase reveal critical residues for activity. *J Biol*
1531 *Chem* 280, 34691-34701.
- 1532 Boudiere, L., Michaud, M., Petroustos, D., Rebeille, F., Falconet, D., Bastien, O., Roy, S.,
1533 Finazzi, G., Rolland, N., Jouhet, J., Block, M.A., and Marechal, E. (2014).
1534 Glycerolipids in photosynthesis: composition, synthesis and trafficking. *Biochim*
1535 *Biophys Acta* 1837, 470-480.
- 1536 Bowler, C., Allen, A.E., Badger, J.H., Grimwood, J., Jabbari, K., Kuo, A., Maheswari, U.,
1537 Martens, C., Maumus, F., Otilar, R.P., Rayko, E., Salamov, A., Vandepoele, K.,
1538 Beszteri, B., Gruber, A., Heijde, M., Katinka, M., Mock, T., Valentin, K., Verret, F.,
1539 Berges, J.A., Brownlee, C., Cadoret, J.P., Chiovitti, A., Choi, C.J., Coesel, S., De
1540 Martino, A., Detter, J.C., Durkin, C., Falciatore, A., Fournet, J., Haruta, M., Huysman,
1541 M.J., Jenkins, B.D., Jiroutova, K., Jorgensen, R.E., Joubert, Y., Kaplan, A., Kroger,

- 1542 N., Kroth, P.G., La Roche, J., Lindquist, E., Lommer, M., Martin-Jezequel, V., Lopez,
1543 P.J., Lucas, S., Mangogna, M., McGinnis, K., Medlin, L.K., Montsant, A., Oudot-Le
1544 Secq, M.P., Napoli, C., Obornik, M., Parker, M.S., Petit, J.L., Porcel, B.M., Poulsen,
1545 N., Robison, M., Rychlewski, L., Rynearson, T.A., Schmutz, J., Shapiro, H., Siaut,
1546 M., Stanley, M., Sussman, M.R., Taylor, A.R., Vardi, A., von Dassow, P., Vyverman,
1547 W., Willis, A., Wyrwicz, L.S., Rokhsar, D.S., Weissenbach, J., Armbrust, E.V.,
1548 Green, B.R., Van de Peer, Y., and Grigoriev, I.V. (2008). The *Phaeodactylum* genome
1549 reveals the evolutionary history of diatom genomes. *Nature* 456, 239-244.
- 1550 Brinkman, E.K., Chen, T., Amendola, M., and van Steensel, B. (2014). Easy quantitative
1551 assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res*
1552 42, e168.
- 1553 Browse, J., Warwick, N., Somerville, C.R., and Slack, C.R. (1986). Fluxes through the
1554 prokaryotic and eukaryotic pathways of lipid synthesis in the 16-3 Plant *Arabidopsis*
1555 *thaliana*. *Biochemical Journal* 235, 25-31.
- 1556 Bullmann, L., Haarmann, R., Mirus, O., Bredemeier, R., Hempel, F., Maier, U.G., and
1557 Schleiff, E. (2010). Filling the gap, evolutionarily conserved Omp85 in plastids of
1558 chromalveolates. *J Biol Chem* 285, 6848-6856.
- 1559 Buseman, C.M., Tamura, P., Sparks, A.A., Baughman, E.J., Maatta, S., Zhao, J., Roth, M.R.,
1560 Esch, S.W., Shah, J., Williams, T.D., and Welti, R. (2006). Wounding stimulates the
1561 accumulation of glycerolipids containing oxophytodienoic acid and dinor-
1562 oxophytodienoic acid in *Arabidopsis* leaves. *Plant Physiol* 142, 28-39.
- 1563 Cavalier-Smith, T. (2018). Kingdom Chromista and its eight phyla: a new synthesis
1564 emphasising periplastid protein targeting, cytoskeletal and periplastid evolution, and
1565 ancient divergences. *Protoplasma* 255, 297-357.
- 1566 Conte, M., Lupette, J., Seddiki, K., Mei, C., Dolch, L.J., Gros, V., Barette, C., Rebeille, F.,
1567 Jouhet, J., and Marechal, E. (2018). Screening for biologically annotated drugs that
1568 trigger triacylglycerol accumulation in the diatom *Phaeodactylum*. *Plant Physiol*.
- 1569 Corteggiani Carpinelli, E., Telatin, A., Vitulo, N., Forcato, C., D'Angelo, M., Schiavon, R.,
1570 Vezzi, A., Giacometti, G.M., Morosinotto, T., and Valle, G. (2014). Chromosome
1571 scale genome assembly and transcriptome profiling of *Nannochloropsis gaditana* in
1572 nitrogen depletion. *Molecular plant* 7, 323-335.
- 1573 Criscuolo, A., and Gribaldo, S. (2010). BMGE (Block Mapping and Gathering with Entropy):
1574 a new software for selection of phylogenetic informative regions from multiple
1575 sequence alignments. *BMC evolutionary biology* 10, 210.
- 1576 Daboussi, F., Leduc, S., Marechal, A., Dubois, G., Guyot, V., Perez-Michaut, C., Amato, A.,
1577 Falciatore, A., Juillerat, A., Beurdeley, M., Voytas, D.F., Cavarec, L., and Duchateau,
1578 P. (2014). Genome engineering empowers the diatom *Phaeodactylum tricornutum* for
1579 biotechnology. *Nat Commun* 5, 3831.
- 1580 De Riso, V., Raniello, R., Maumus, F., Rogato, A., Bowler, C., and Falciatore, A. (2009).
1581 Gene silencing in the marine diatom *Phaeodactylum tricornutum*. *Nucleic Acids Res*
1582 37, e96.
- 1583 de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahe, F., Logares, R., Lara, E., Berney, C.,
1584 Le Bescot, N., Probert, I., Carmichael, M., Poulain, J., Romac, S., Colin, S., Aury,
1585 J.M., Bittner, L., Chaffron, S., Dunthorn, M., Engelen, S., Flegontova, O., Guidi, L.,
1586 Horak, A., Jaillon, O., Lima-Mendez, G., Lukes, J., Malviya, S., Morard, R., Mulot,
1587 M., Scalco, E., Siano, R., Vincent, F., Zingone, A., Dimier, C., Picheral, M., Searson,
1588 S., Kandels-Lewis, S., Tara Oceans, C., Acinas, S.G., Bork, P., Bowler, C., Gorsky,
1589 G., Grimsley, N., Hingamp, P., Iudicone, D., Not, F., Ogata, H., Pesant, S., Raes, J.,
1590 Sieracki, M.E., Speich, S., Stemmann, L., Sunagawa, S., Weissenbach, J., Wincker, P.,

- 1591 and Karsenti, E. (2015). Ocean plankton. Eukaryotic plankton diversity in the sunlit
1592 ocean. *Science* 348, 1261605.
- 1593 Deme, B., Cataye, C., Block, M.A., Marechal, E., and Jouhet, J. (2014). Contribution of
1594 galactoglycerolipids to the 3-dimensional architecture of thylakoids. *FASEB J* 28,
1595 3373-3383.
- 1596 Dolch, L.J., and Marechal, E. (2015). Inventory of fatty acid desaturases in the pennate
1597 diatom *Phaeodactylum tricornutum*. *Marine drugs* 13, 1317-1339.
- 1598 Dolch, L.J., Rak, C., Perin, G., Tourcier, G., Broughton, R., Leterrier, M., Morosinotto, T.,
1599 Tellier, F., Faure, J.D., Falconet, D., Jouhet, J., Sayanova, O., Beaudoin, F., and
1600 Marechal, E. (2017a). A Palmitic Acid Elongase Affects Eicosapentaenoic Acid and
1601 Plastidial Monogalactosyldiacylglycerol Levels in *Nannochloropsis*. *Plant Physiol*
1602 173, 742-759.
- 1603 Dolch, L.J., Lupette, J., Tourcier, G., Bedhomme, M., Collin, S., Magneschi, L., Conte, M.,
1604 Seddiki, K., Richard, C., Corre, E., Fourage, L., Laeuffer, F., Richards, R., Reith, M.,
1605 Rebeille, F., Jouhet, J., McGinn, P., and Marechal, E. (2017b). Nitric Oxide Mediates
1606 Nitrite-Sensing and Acclimation and Triggers a Remodeling of Lipids. *Plant Physiol*
1607 175, 1407-1423.
- 1608 Dubots, E., Audry, M., Yamaryo, Y., Bastien, O., Ohta, H., Breton, C., Marechal, E., and
1609 Block, M.A. (2010). Activation of the chloroplast monogalactosyldiacylglycerol
1610 synthase MGD1 by phosphatidic acid and phosphatidylglycerol. *J Biol Chem* 285,
1611 6003-6011.
- 1612 Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high
1613 throughput. *Nucleic Acids Res* 32, 1792-1797.
- 1614 Emanuelsson, O., Brunak, S., von Heijne, G., and Nielsen, H. (2007). Locating proteins in the
1615 cell using TargetP, SignalP and related tools. *Nature protocols* 2, 953-971.
- 1616 Falciatore, A., Alcalà, M.R., Croot, P., and Bowler, C. (2000). Perception of Environmental
1617 Signals by a Marine Diatom. *Science* 288, 2363.
- 1618 Falciatore, A., Casotti, R., Leblanc, C., Abrescia, C., and Bowler, C. (1999). Transformation
1619 of Nonselectable Reporter Genes in Marine Diatoms. *Mar Biotechnol (NY)* 1, 239-
1620 251.
- 1621 Fawley, M.W., Jameson, I., and Fawley, K.P. (2015). The phylogeny of the genus
1622 *Nannochloropsis* (Monodopsidaceae, Eustigmatophyceae), with descriptions of *N.*
1623 *australis* sp. nov. and *Microchloropsis* gen. nov. *Phycologia* 54, 545-552.
- 1624 Flori, S., Jouneau, P.H., Finazzi, G., Marechal, E., and Falconet, D. (2016). Ultrastructure of
1625 the Periplastidial Compartment of the Diatom *Phaeodactylum tricornutum*. *Protist* 167,
1626 254-267.
- 1627 Flori, S., Jouneau, P.H., Gallet, B., Estrozi, L.F., Moriscot, C., Schoehn, G., Finazzi, G., and
1628 Falconet, D. (2018). Imaging Plastids in 2D and 3D: Confocal and Electron
1629 Microscopy. *Methods Mol Biol* 1829, 113-122.
- 1630 Folch, J., Lees, M., and Sloane Stanley, G.H. (1957). A simple method for the isolation and
1631 purification of total lipides from animal tissues. *J Biol Chem* 226, 497-509.
- 1632 Garab, G., Yaguzhinsky, L.S., Dlouhy, O., Nesterov, S.V., Spunda, V., and Gasanoff, E.S.
1633 (2022). Structural and functional roles of non-bilayer lipid phases of chloroplast
1634 thylakoid membranes and mitochondrial inner membranes. *Prog Lipid Res* 86,
1635 101163.
- 1636 Gascuel, O. (1997). BIONJ: an improved version of the NJ algorithm based on a simple
1637 model of sequence data. *Mol Biol Evol* 14, 685-695.
- 1638 Goss, R., and Jakob, T. (2010). Regulation and function of xanthophyll cycle-dependent
1639 photoprotection in algae. *Photosynth Res* 106, 103-122.

- 1640 Gould, S.B., Maier, U.G., and Martin, W.F. (2015). Protein Import and the Origin of Red
1641 Complex Plastids. *Curr Biol* 25, R515-R521.
- 1642 Grosche, C., Hempel, F., Bolte, K., Zauner, S., and Maier, U.G. (2014). The periplastidal
1643 compartment: a naturally minimized eukaryotic cytoplasm. *Curr Opin Microbiol* 22,
1644 88-93.
- 1645 Gruber, A., Rocap, G., Kroth, P.G., Armbrust, E.V., and Mock, T. (2015). Plastid proteome
1646 prediction for diatoms and other algae with secondary plastids of the red lineage. *Plant*
1647 *J* 81, 519-528.
- 1648 Gruber, A., Vugrinec, S., Hempel, F., Gould, S.B., Maier, U.G., and Kroth, P.G. (2007).
1649 Protein targeting into complex diatom plastids: functional characterisation of a
1650 specific targeting motif. *Plant Mol Biol* 64, 519-530.
- 1651 Gschloessl, B., Guermeur, Y., and Cock, J.M. (2008). HECTAR: a method to predict
1652 subcellular targeting in heterokonts. *BMC Bioinformatics* 9, 393.
- 1653 Gu, X., Cao, L., Wu, X., Li, Y., Hu, Q., and Han, D. (2021). A Lipid Bodies-Associated
1654 Galactosyl Hydrolase Is Involved in Triacylglycerol Biosynthesis and Galactolipid
1655 Turnover in the Unicellular Green Alga *Chlamydomonas reinhardtii*. *Plants (Basel)*
1656 10.
- 1657 Guiry, M.D. (2012). How Many Species of Algae Are There? *J Phycol* 48, 1057-1063.
- 1658 Hempel, F., Bullmann, L., Lau, J., Zauner, S., and Maier, U.G. (2009). ERAD-derived
1659 preprotein transport across the second outermost plastid membrane of diatoms. *Mol*
1660 *Biol Evol* 26, 1781-1790.
- 1661 Hori, K., Nobusawa, T., Watanabe, T., Madoka, Y., Suzuki, H., Shibata, D., Shimojima, M.,
1662 and Ohta, H. (2016). Tangled evolutionary processes with commonality and diversity
1663 in plastidial glycolipid synthesis in photosynthetic organisms. *Biochim Biophys Acta*
1664 1861, 1294-1308.
- 1665 Horton, P., Park, K.J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., and Nakai,
1666 K. (2007). WoLF PSORT: protein localization predictor. *Nucleic Acids Res* 35,
1667 W585-587.
- 1668 Huang, T., Pan, Y., Marechal, E., and Hu, H. (2024). Proteomes reveal the lipid metabolic
1669 network in the complex plastid of *Phaeodactylum tricornutum*. *Plant J* 117, 385-403.
- 1670 Iwai, M., Yamada-Oshima, Y., Asami, K., Kanamori, T., Yuasa, H., Shimojima, M., and
1671 Ohta, H. (2021). Recycling of the major thylakoid lipid MGDG and its role in lipid
1672 homeostasis in *Chlamydomonas reinhardtii*. *Plant Physiol* 187, 1341-1356.
- 1673 Janson, G., Zhang, C., Prado, M.G., and Paiardini, A. (2017). PyMod 2.0: improvements in
1674 protein sequence-structure analysis and homology modeling within PyMOL.
1675 *Bioinformatics* 33, 444-446.
- 1676 Jensen, P.E., and Leister, D. (2014). Chloroplast evolution, structure and functions.
1677 *F1000Prime Rep* 6, 40.
- 1678 Jouhet, J., Lupette, J., Clerc, O., Magneschi, L., Bedhomme, M., Collin, S., Roy, S.,
1679 Maréchal, E., and Rébeillé, F. (2017). LC-MS/MS versus TLC plus GC methods:
1680 Consistency of glycerolipid and fatty acid profiles in microalgae and higher plant cells
1681 and effect of a nitrogen starvation. *PLOS ONE* 12, e0182423.
- 1682 Jumper, J., and Hassabis, D. (2022). Protein structure predictions to atomic accuracy with
1683 AlphaFold. *Nat Methods* 19, 11-12.
- 1684 Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O.,
1685 Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., Bridgland, A., Meyer, C.,
1686 Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R.,
1687 Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M.,
1688 Pacholska, M., Berghammer, T., Bodenstern, S., Silver, D., Vinyals, O., Senior, A.W.,

- 1689 Kavukcuoglu, K., Kohli, P., and Hassabis, D. (2021). Highly accurate protein structure
1690 prediction with AlphaFold. *Nature* 596, 583-589.
- 1691 Karas, B.J., Diner, R.E., Lefebvre, S.C., McQuaid, J., Phillips, A.P., Noddings, C.M.,
1692 Brunson, J.K., Valas, R.E., Deerinck, T.J., Jablanovic, J., Gillard, J.T., Beeri, K.,
1693 Ellisman, M.H., Glass, J.I., Hutchison, C.A., 3rd, Smith, H.O., Venter, J.C., Allen,
1694 A.E., Dupont, C.L., and Weyman, P.D. (2015). Designer diatom episomes delivered
1695 by bacterial conjugation. *Nat Commun* 6, 6925.
- 1696 Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J. (2015). The Phyre2
1697 web portal for protein modeling, prediction and analysis. *Nature protocols* 10, 845-
1698 858.
- 1699 Kilian, O., and Kroth, P.G. (2005). Identification and characterization of a new conserved
1700 motif within the presequence of proteins targeted into complex diatom plastids. *Plant J*
1701 41, 175-183.
- 1702 Kobayashi, K., Nakamura, Y., and Ohta, H. (2009a). Type A and type B
1703 monogalactosyldiacylglycerol synthases are spatially and functionally separated in the
1704 plastids of higher plants. *Plant Physiology and Biochemistry* 47, 518-525.
- 1705 Kobayashi, K., Kondo, M., Fukuda, H., Nishimura, M., and Ohta, H. (2007a). Galactolipid
1706 biosynthesis is essential for proper chloroplast biogenesis and embryogenesis.
1707 *Photosynthesis Research* 91, 216-216.
- 1708 Kobayashi, K., Kondo, M., Fukuda, H., Nishimura, M., and Ohta, H. (2007b). Galactolipid
1709 synthesis in chloroplast inner envelope is essential for proper thylakoid biogenesis,
1710 photosynthesis, and embryogenesis. *Proc Natl Acad Sci U S A* 104, 17216-17221.
- 1711 Kobayashi, K., Awai, K., Nakamura, M., Nagatani, A., Masuda, T., and Ohta, H. (2009b).
1712 Type-B monogalactosyldiacylglycerol synthases are involved in phosphate starvation-
1713 induced lipid remodeling, and are crucial for low-phosphate adaptation. *Plant Journal*
1714 57, 322-331.
- 1715 Kooistra, W.H., De Stefano, M., Mann, D.G., and Medlin, L.K. (2003). The phylogeny of the
1716 diatoms. *Prog Mol Subcell Biol* 33, 59-97.
- 1717 Kooistra, W.H., Gersonde, R., Medlin, L.K., and Mann, D.G. (2007). The Origin and
1718 Evolution of the Diatoms: Their Adaptation to a Planktonic Existence. In *Evolution of*
1719 *Primary Producers in the Sea* P.G. Falkowski and A.H. Knoll, eds (Cambridge,
1720 U.S.A.: Academic Press), pp. 207-249.
- 1721 Kroth, P.G., Bones, A.M., Daboussi, F., Ferrante, M.I., Jaubert, M., Kolot, M., Nymark, M.,
1722 Rio Bartulos, C., Ritter, A., Russo, M.T., Serif, M., Winge, P., and Falciatore, A.
1723 (2018). Genome editing in diatoms: achievements and goals. *Plant Cell Rep* 37, 1401-
1724 1408.
- 1725 Kumar, S., Stecher, G., Li, M., Nnyaz, C., and Tamura, K. (2018). MEGA X: Molecular
1726 Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol* 35, 1547-
1727 1549.
- 1728 Lang, M., Apt, K.E., and Kroth, P.G. (1998). Protein transport into "complex" diatom plastids
1729 utilizes two different targeting signals. *J Biol Chem* 273, 30973-30978.
- 1730 Le, S.Q., Lartillot, N., and Gascuel, O. (2008). Phylogenetic mixture models for proteins.
1731 *Philos Trans R Soc Lond B Biol Sci* 363, 3965-3976.
- 1732 Lemoine, F., Correia, D., Lefort, V., Doppelt-Azeroual, O., Mareuil, F., Cohen-Boulakia, S.,
1733 and Gascuel, O. (2019). NGPhylogeny.fr: new generation phylogenetic services for
1734 non-specialists. *Nucleic Acids Res* 47, W260-W265.
- 1735 Letunic, I., and Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for
1736 phylogenetic tree display and annotation. *Nucleic Acids Res* 49, W293-W296.
- 1737 Leyland, B., Zarka, A., Didi-Cohen, S., Boussiba, S., and Khozin-Goldberg, I. (2020). High
1738 Resolution Proteome of Lipid Droplets Isolated from the Pennate Diatom

- 1739 Phaeodactylum tricornutum (Bacillariophyceae) Strain pt4 provides mechanistic
 1740 insights into complex intracellular coordination during nitrogen deprivation. *J Phycol.*
 1741 Li-Beisson, Y., Shorosh, B., Beisson, F., Andersson, M.X., Arondel, V., Bates, P.D., Baud,
 1742 S., Bird, D., Debono, A., Durrett, T.P., Franke, R.B., Graham, I.A., Katayama, K.,
 1743 Kelly, A.A., Larson, T., Markham, J.E., Miquel, M., Molina, I., Nishida, I., Rowland,
 1744 O., Samuels, L., Schmid, K.M., Wada, H., Welti, R., Xu, C., Zallot, R., and Ohlrogge,
 1745 J. (2010). Acyl-lipid metabolism. In *The Arabidopsis Book*, pp. e0133.
- 1746 Li, X., Moellering, E.R., Liu, B., Johnny, C., Fedewa, M., Sears, B.B., Kuo, M.H., and
 1747 Benning, C. (2012). A galactoglycerolipid lipase is required for triacylglycerol
 1748 accumulation and survival following nitrogen deprivation in *Chlamydomonas*
 1749 *reinhardtii*. *Plant Cell* 24, 4670-4686.
- 1750 Liu, X.J., Hempel, F., Stork, S., Bolte, K., Moog, D., Heimerl, T., Maier, U.G., and Zauner, S.
 1751 (2016). Addressing various compartments of the diatom model organism via sub-
 1752 cellular marker proteins. *Algal Res* 20, 249-257.
- 1753 Lupette, J., Tardif, M., Brugiére, S., Coute, Y., Salvaing, J., and Marechal, E. (2022).
 1754 Quantitative proteomic analyses reveal the impact of nitrogen starvation on the
 1755 proteome of the model diatom *Phaeodactylum tricornutum*. *Proteomics* 22, e2200155.
- 1756 Lupette, J., Jaussaud, A., Seddiki, K., Morabito, C., Brugiére, S., Schaller, H., Kuntz, M.,
 1757 Putaux, J.L., Jouneau, P.H., Rebellle, F., Falconet, D., Coute, Y., Jouhet, J., Tardif,
 1758 M., Salvaing, J., and Marechal, E. (2019). The architecture of lipid droplets in the
 1759 diatom *Phaeodactylum tricornutum*. *Algal Res* 38.
- 1760 Makshakova, O., Breton, C., and Perez, S. (2020). Unraveling the complex enzymatic
 1761 machinery making a key galactolipid in chloroplast membrane: a multiscale computer
 1762 simulation. *Scientific reports* 10, 13514.
- 1763 Mann, D.G. (1999). The species concept in diatoms. *Phycologia* 38, 437-495.
- 1764 Mann, D.G., and Droop, S.J.M. (1996). Biodiversity, biogeography and conservation of
 1765 diatoms. *Hydrobiologia*, 19–32.
- 1766 Marechal, E. (2024). How Did Thylakoids Emerge in Cyanobacteria, and How Were the
 1767 Primary Chloroplast and Chromatophore Acquired? *Methods Mol Biol* 2776, 3-20.
- 1768 Martens, C., Vandepoele, K., and Van de Peer, Y. (2008). Whole-genome analysis reveals
 1769 molecular innovations and evolutionary transitions in chromalveolate species. *Proc*
 1770 *Natl Acad Sci U S A* 105, 3427-3432.
- 1771 Mirdita, M., Schutze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022).
 1772 ColabFold: making protein folding accessible to all. *Nature Methods* 19, 679-+.
- 1773 Moog, D., Stork, S., Zauner, S., and Maier, U.G. (2011). In silico and in vivo investigations
 1774 of proteins of a minimized eukaryotic cytoplasm. *Genome Biol Evol* 3, 375-382.
- 1775 Murakami, R., and Hashimoto, H. (2009). Unusual nuclear division in *Nannochloropsis*
 1776 *oculata* (Eustigmatophyceae, Heterokonta) which may ensure faithful transmission of
 1777 secondary plastids. *Protist* 160, 41-49.
- 1778 Nitenberg, M., Makshakova, O., Rocha, J., Perez, S., Marechal, E., Block, M.A., Girard-
 1779 Egrot, A., and Breton, C. (2020). Mechanism of activation of plant
 1780 monogalactosyldiacylglycerol synthase 1 (MGD1) by phosphatidylglycerol.
 1781 *Glycobiology* 30, 396-406.
- 1782 Parks, M.B., Nakov, T., Ruck, E.C., Wickett, N.J., and Alverson, A.J. (2018). Phylogenomics
 1783 reveals an extensive history of genome duplication in diatoms (Bacillariophyta). *Am J*
 1784 *Bot* 105, 330-347.
- 1785 Petroustos, D., Amiar, S., Abida, H., Dolch, L.J., Bastien, O., Rebeille, F., Jouhet, J.,
 1786 Falconet, D., Block, M.A., McFadden, G.I., Bowler, C., Botte, C., and Marechal, E.
 1787 (2014). Evolution of galactoglycerolipid biosynthetic pathways--from cyanobacteria to
 1788 primary plastids and from primary to secondary plastids. *Prog Lipid Res* 54, 68-85.

- 1789 Rastogi, A., Murik, O., Bowler, C., and Tirichine, L. (2016). PhytoCRISP-Ex: a web-based
1790 and stand-alone application to find specific target sequences for CRISPR/CAS editing.
1791 BMC Bioinformatics 17, 261.
- 1792 Reyes-Prieto, A., Weber, A.P., and Bhattacharya, D. (2007). The origin and establishment of
1793 the plastid in algae and plants. *Annu Rev Genet* 41, 147-168.
- 1794 Rocha, J., Sarkis, J., Thomas, A., Pitou, L., Radzimanowski, J., Audry, M., Chazalet, V., de
1795 Sanctis, D., Palcic, M.M., Block, M.A., Girard-Egrot, A., Marechal, E., and Breton, C.
1796 (2016). Structural insights and membrane binding properties of MGD1, the major
1797 galactolipid synthase in plants. *Plant J* 85, 622-633.
- 1798 Ruban, A., Lavaud, J., Rousseau, B., Guglielmi, G., Horton, P., and Etienne, A.L. (2004). The
1799 super-excess energy dissipation in diatom algae: comparative analysis with higher
1800 plants. *Photosynth Res* 82, 165-175.
- 1801 Sato, N., and Awai, K. (2017). "Prokaryotic Pathway" Is Not Prokaryotic: Noncyanobacterial
1802 Origin of the Chloroplast Lipid Biosynthetic Pathway Revealed by Comprehensive
1803 Phylogenomic Analysis. *Genome Biol Evol* 9, 3162-3178.
- 1804 Serôdio, J., and Lavaud, J. (2020). Diatoms and Their Ecological Importance. In *Life Below*
1805 *Water*, W. Leal Filho, A.M. Azul, L. Brandli, A. Lange Salvia, and T. Wall, eds
1806 (Cham: Springer International Publishing), pp. 1-9.
- 1807 Shang, S., Liu, R., Luo, L., Li, X., Zhang, S., Zhang, Y., Zheng, P., Chen, Z., and Wang, B.
1808 (2022). Functional Characterization of the Monogalactosyldiacylglycerol Synthase
1809 Gene ptMGD2 in the Diatom *Phaeodactylum tricornutum*. *Front. Mar. Sci.* 9,
1810 e874448.
- 1811 Siaut, M., Heijde, M., Mangogna, M., Montsant, A., Coesel, S., Allen, A., Manfredonia, A.,
1812 Falciatore, A., and Bowler, C. (2007). Molecular toolbox for studying diatom biology
1813 in *Phaeodactylum tricornutum*. *Gene* 406, 23-35.
- 1814 Smith, R., Jouhet, J., Gandini, C., Nekrasov, V., Marechal, E., Napier, J.A., and Sayanova, O.
1815 (2021). Plastidial acyl carrier protein Delta9-desaturase modulates eicosapentaenoic
1816 acid biosynthesis and triacylglycerol accumulation in *Phaeodactylum tricornutum*.
1817 *Plant J* 106, 1247-1259.
- 1818 Sorhannus, U. (2007). A nuclear-encoded small-subunit ribosomal RNA timescale for diatom
1819 evolution. *Mar Micropaleontol* 65, 1-12.
- 1820 Tanaka, A., De Martino, A., Amato, A., Montsant, A., Mathieu, B., Rostaing, P., Tirichine,
1821 L., and Bowler, C. (2015). Ultrastructure and Membrane Traffic During Cell Division
1822 in the Marine Pennate Diatom *Phaeodactylum tricornutum*. *Protist* 166, 506-521.
- 1823 Vieler, A., Wu, G., Tsai, C.H., Bullard, B., Cornish, A.J., Harvey, C., Reca, I.B., Thornburg,
1824 C., Achawanantakun, R., Buehl, C.J., Campbell, M.S., Cavalier, D., Childs, K.L.,
1825 Clark, T.J., Deshpande, R., Erickson, E., Armenia Ferguson, A., Handee, W., Kong,
1826 Q., Li, X., Liu, B., Lundback, S., Peng, C., Roston, R.L., Sanjaya, Simpson, J.P.,
1827 Terbush, A., Warakanont, J., Zauner, S., Farre, E.M., Hegg, E.L., Jiang, N., Kuo,
1828 M.H., Lu, Y., Niyogi, K.K., Ohlrogge, J., Osteryoung, K.W., Shachar-Hill, Y., Sears,
1829 B.B., Sun, Y., Takahashi, H., Yandell, M., Shiu, S.H., and Benning, C. (2012).
1830 Genome, functional gene annotation, and nuclear transformation of the heterokont
1831 oleaginous alga *Nannochloropsis oceanica* CCMP1779. *PLoS Genet* 8, e1003064.
- 1832 Yates, A.D., Allen, J., Amode, R.M., Azov, A.G., Barba, M., Becerra, A., Bhai, J., Campbell,
1833 L.I., Martinez, M.C., Chakiachvili, M., Chougule, K., Christensen, M., Contreras-
1834 Moreira, B., Cuzick, A., Fioretto, L.D., Davis, P., De Silva, N.H., Diamantakis, S.,
1835 Dyer, S., Elser, J., Filippi, C.V., Gall, A., Grigoriadis, D., Guijarro-Clarke, C., Gupta,
1836 P., Hammond-Kosack, K.E., Howe, K.L., Jaiswal, P., Kaikala, V., Kumar, V., Kumari,
1837 S., Langridge, N., Le, T., Luypaert, M., Maslen, G.L., Maurel, T., Moore, B., Muffato,
1838 M., Mushtaq, A., Naamati, G., Naithani, S., Olson, A., Parker, A., Paulini, M., Pedro,

1839 H., Perry, E., Preece, J., Quinton-Tulloch, M., Rodgers, F., Rosello, M., Ruffier, M.,
1840 Seager, J., Sitnik, V., Szpak, M., Tate, J., Tello-Ruiz, M.K., Trevanion, S.J., Urban,
1841 M., Ware, D., Wei, S., Williams, G., Winterbottom, A., Zarowiecki, M., Finn, R.D.,
1842 and Flicek, P. (2022). Ensembl Genomes 2022: an expanding genome resource for
1843 non-vertebrates. *Nucleic Acids Research* 50, D996-D1003.
1844 Zaslavskaja, L.A., Lippmeier, J.C., Kroth, P.G., Grossman, A.R., and Apt, K.E. (2000).
1845 Transformation of the diatom *Phaeodactylum tricornutum* (Bacillariophyceae) with a
1846 variety of selectable marker and reporter genes. *J Phycol* 36, 379-386.
1847 Zhang, C., and Hu, H. (2014). High-efficiency nuclear transformation of the diatom
1848 *Phaeodactylum tricornutum* by electroporation. *Mar Genomics* 16, 63-66.
1849
1850

ACCEPTED MANUSCRIPT

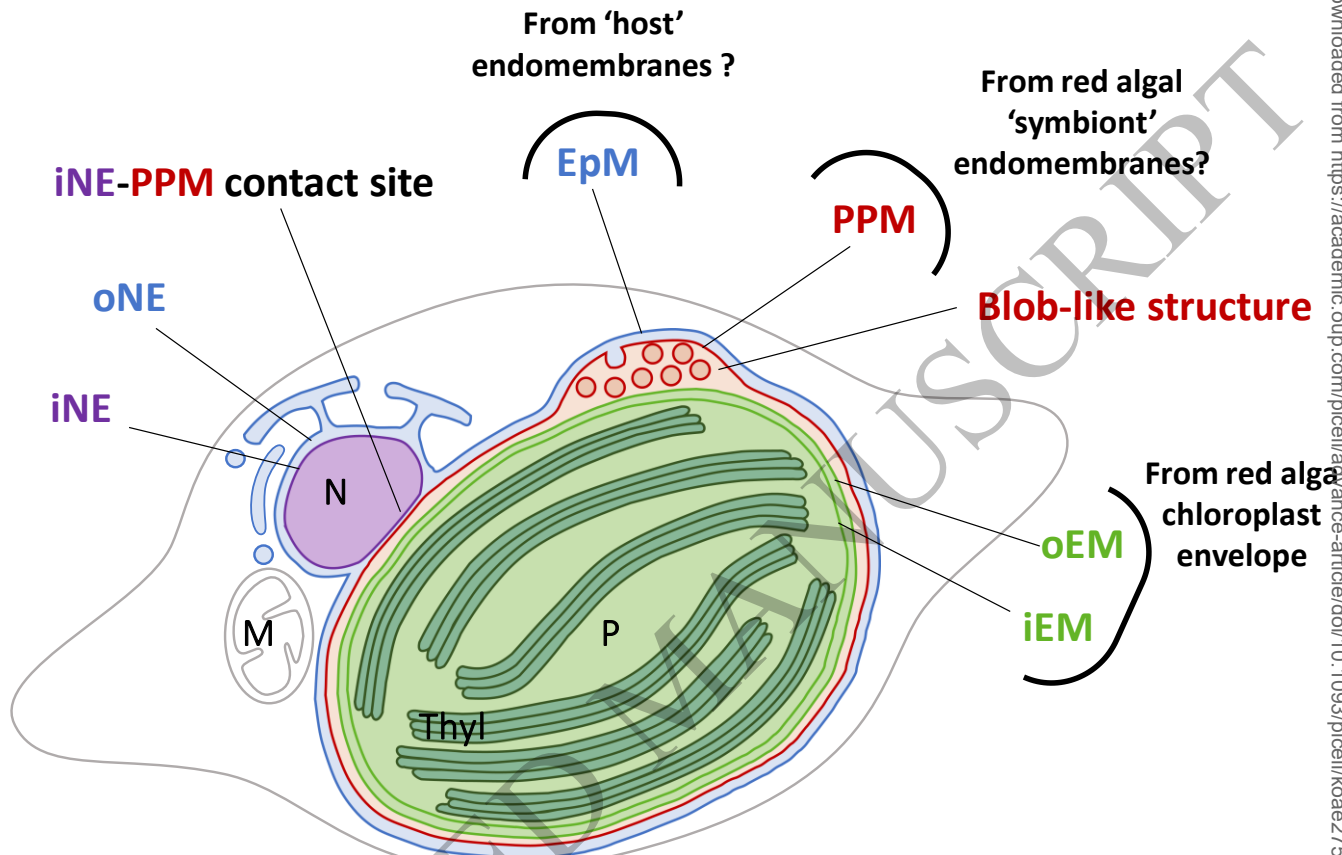


Figure 1: Chimeric origin of the secondary plastid in diatom. The scheme shows a fusiform cell of *Phaeodactylum tricornutum*. The plastid is limited by four membranes. The epiplastidial membrane (EpM), shown in blue, is continuous with the outer nuclear envelope (oNE). The periplastidial membrane (PPM), shown in red, is in tight contact with the inner nuclear envelope (iNE). The outer and inner envelope membranes (iEM and oEM), shown in light green, are tightly apposed. A specific periplastidial compartment relies on the detection of a “blob”-like structure observed by confocal and transmission electron microscopy. In confocal microscopy, the precursors of blob-residing proteins fused to GFP cross only the EpM and PPM and form fluorescent spots in the middle of the plastid (Lang et al., 1998; Kilian and Kroth, 2005; Huang et al., 2024). In electron microscopy, the blob is marked by the development of a vesicular network (VN) budding from the PPM (Flori et al., 2016). Each membrane surrounding the plastid derive from distinct evolutionary origin. The oEM and iEM stem from the chloroplast of the engulfed red algae. The PPM potentially derive from the endomembrane systems of the red algal symbiont and the EpM from that of the host cell, respectively. M, mitochondrion; N, nucleus; P, plastid; Thyl: thylakoids.

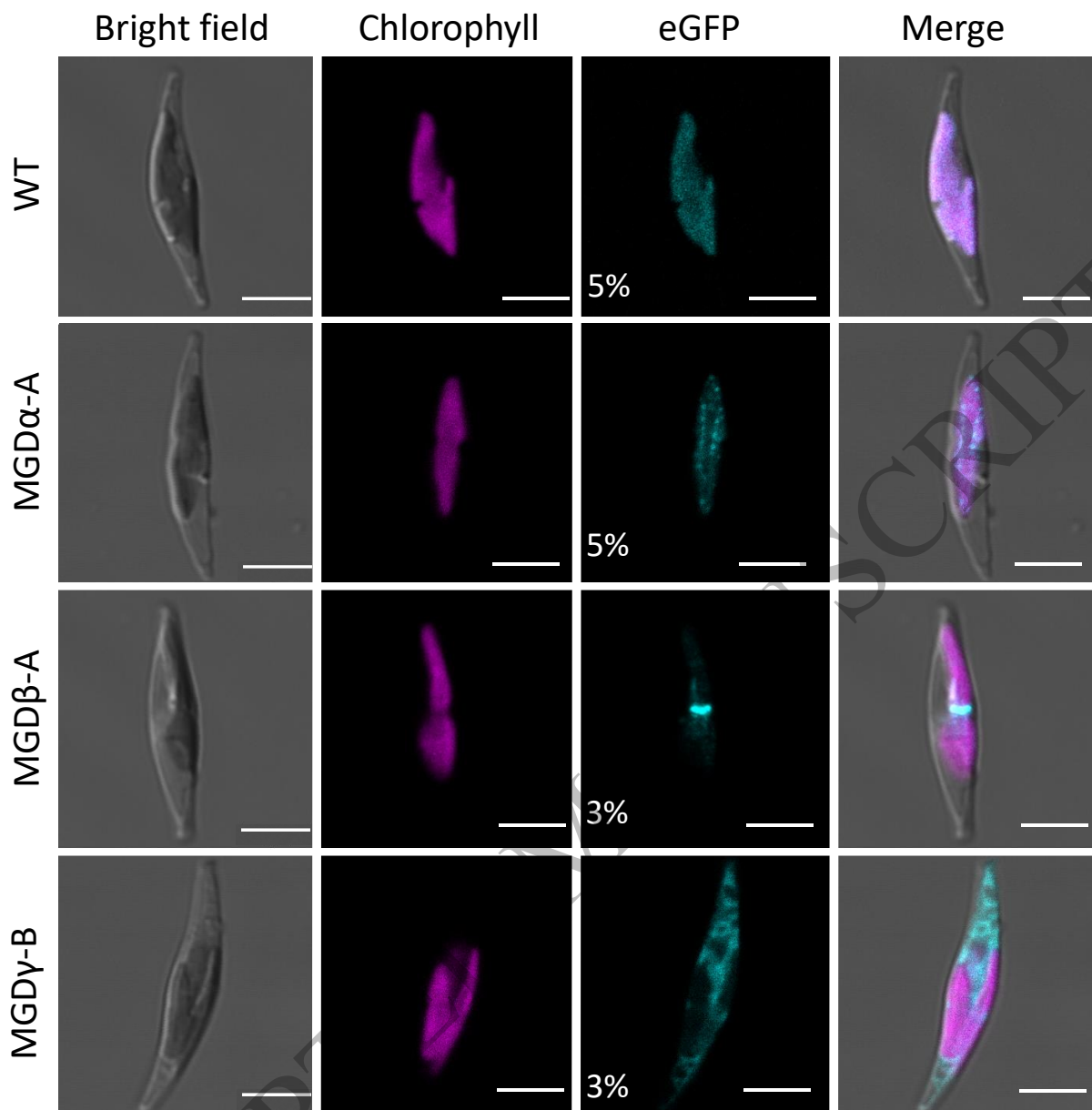


Figure 2: Subcellular localization of MGD isoforms fused to eGFP in *P. tricornutum*. Detection of MGD-eGFP expressed in transgenic *P. tricornutum* lines with wild type as a negative control. For each cell, bright field, chlorophyll fluorescence (magenta) and eGFP signal (light blue) are shown, with a merge of all images. eGFP excitation was achieved at either 5 % or 3 % of laser power, as indicated. The WT control shows low unspecific detection of fluorescence from thylakoid photosystems. An intraplasmidial localization was obtained when MGD α -eGFP is expressed. With MGD β -eGFP, the eGFP signal corresponds to the blob-like structure, indicating a localization of MGD β at the level of the periplasmic compartment, and possibly the PPM. MGD γ -eGFP is detected mainly in the ER, and at lower level in the EpM. Scale bar: 5 μ m. Similar independent observations are provided in Supplemental Figure 7. MGD, monogalactosyldiacylglycerol synthase; WT, wild type.

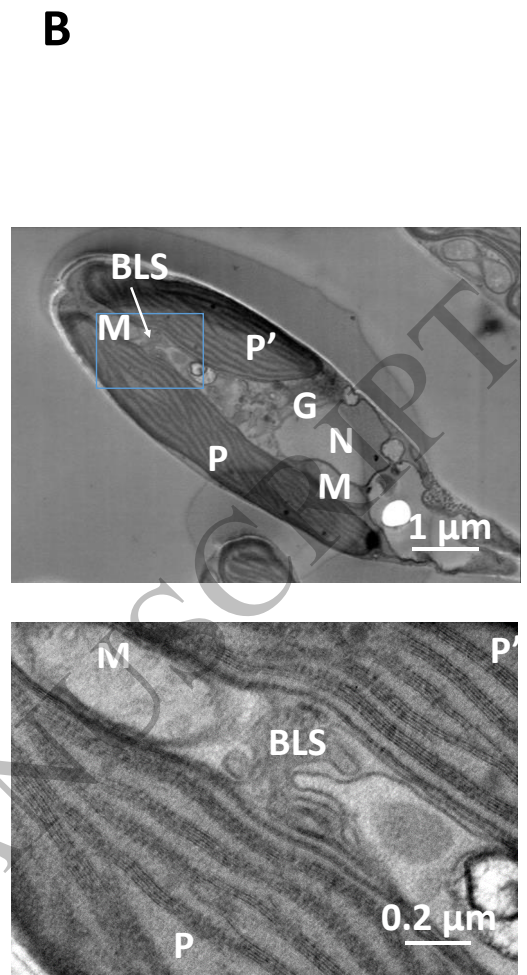
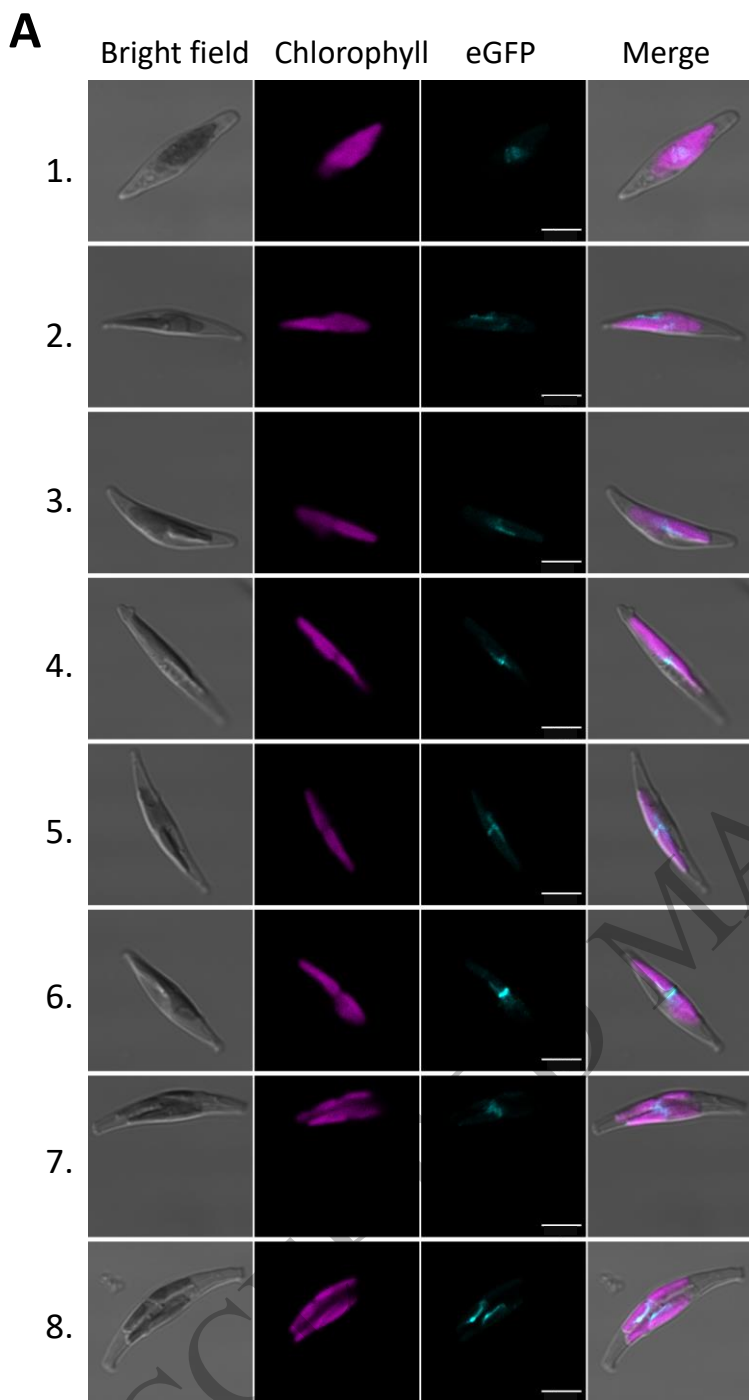


Figure 3: Subcellular localization of the blob-like structure, associated to MGD β -eGFP fluorescence. Observation of blob-like structure localization by confocal microscopy (A) and transmission electron microscopy (B). *P. tricornutum* transgenic lines overexpressing MGD β -eGFP were used to monitor the appearance of blob-like structures in the cell (A). Cells 1 and 2 show a scattered eGFP signal, cells 3–6 show a blob-like structure at the central constriction of dividing plastids, cell 7 shows a blob-like structure extending from one plastid to the other inside a dividing cell, and cell 8 shows two blob-like structures facing each other on distinct plastids during cytokinesis. Cell 6 was also shown in Figure 2. Observation of a wild type *P. tricornutum* cell containing two plastids following plastid division inside the cell (B). Cleavage furrows are visible at both end of the cell. The blob-like structure (outlined in blue) was observed connecting the two plastids at two magnifications. BLS, blob-like structure (or blob); G, Golgi apparatus; M, mitochondrion; MGD, monogalactosyldiacylglycerol synthase; N, nucleus; P and P', plastids.

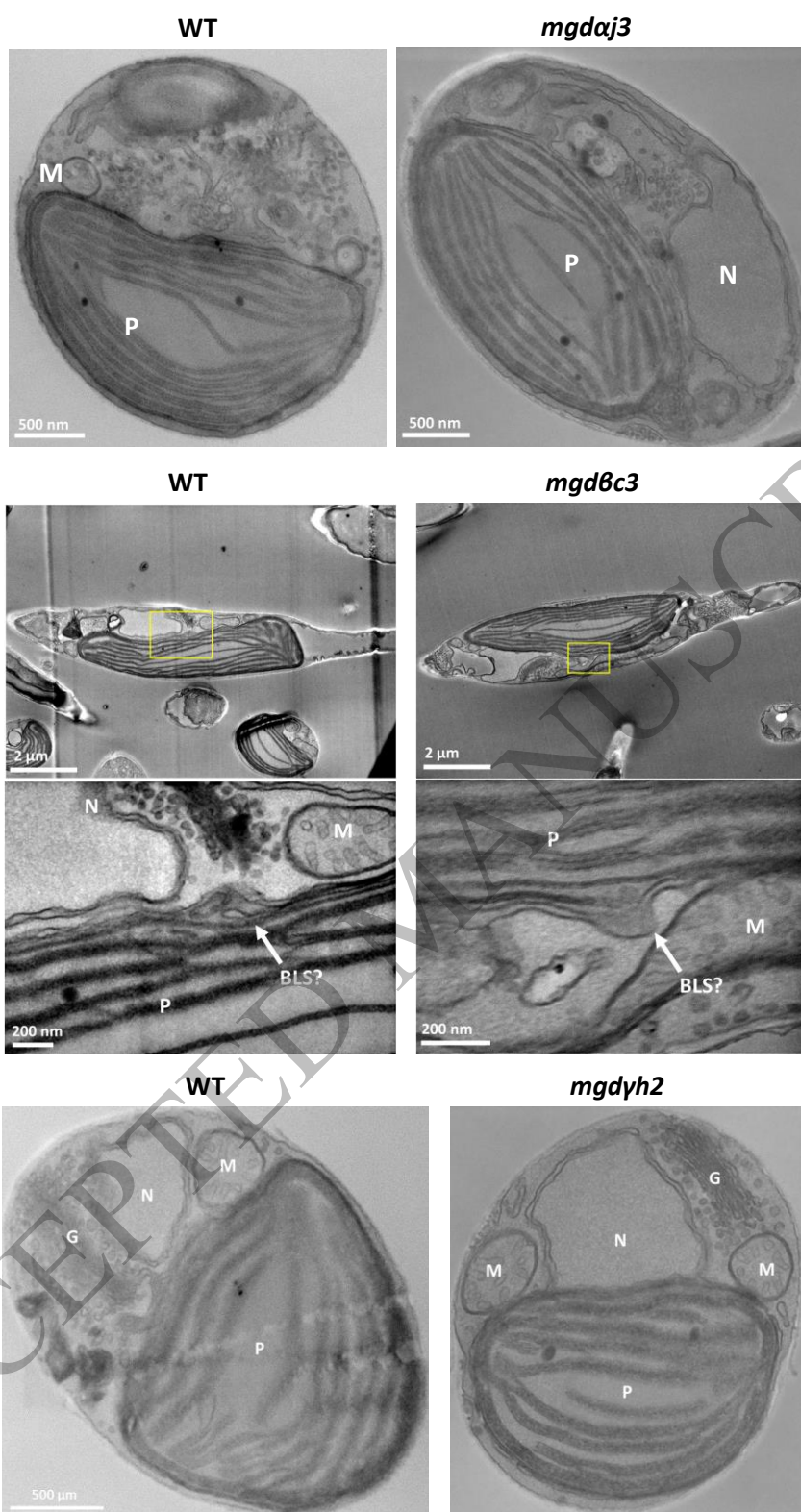


Figure 4: *P. tricornutum* cell ultrastructure in MGD KO lines determined by scanning transmission electron microscopy. A WT and KO *mgdα3* and *mgdγ2* lines were cultured in parallel. In a separate experiment, a WT and a *mgdβ3* mutant were cultured in parallel. Cell ultrastructure is shown in each mutant with corresponding WT control on the left. No impact of *MGDα*, *MGDβ* or *MGDγ* KO could be observed at the level of membrane compartments, including plastids. BLS outlined in yellow was observed at two magnifications. BLS, blob-like structure; G, Golgi, KO, knock-out; M, mitochondria; MGD, monogalactosyldiacylglycerol synthase; N, nucleus, P, plastid; WT, wild type.

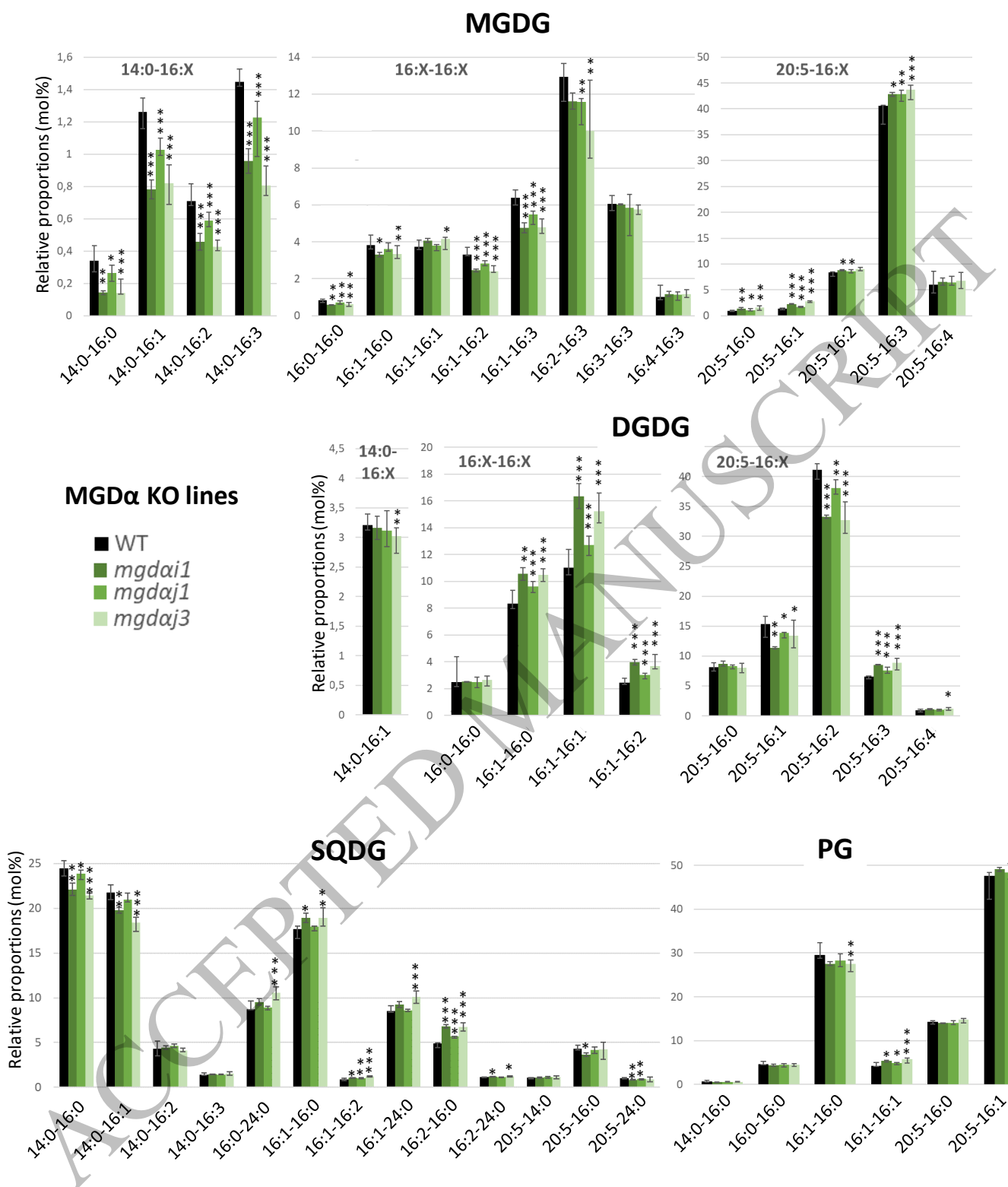


Figure 5: Impact of MGDα mutations on MGDG, DGDG, SQG and PG molecular species. Lipid molecular species in MGDG, DGDG, SQDG and PG were analysed by LC-MS/MS. Each result is the median of six biological replicates ± min and max values. (*), *P*-value < 5.10⁻²; (**), *P*-value < 1.10⁻²; (***), *P*-value < 1.10⁻³, based on an unpaired multiple t test. DGDG, digalactosyldiacylglycerol, MGDG, monogalactosyldiacylglycerol, PG, phosphatidylglycerol, SQDG, sulfoquinovosyldiacylglycerol.

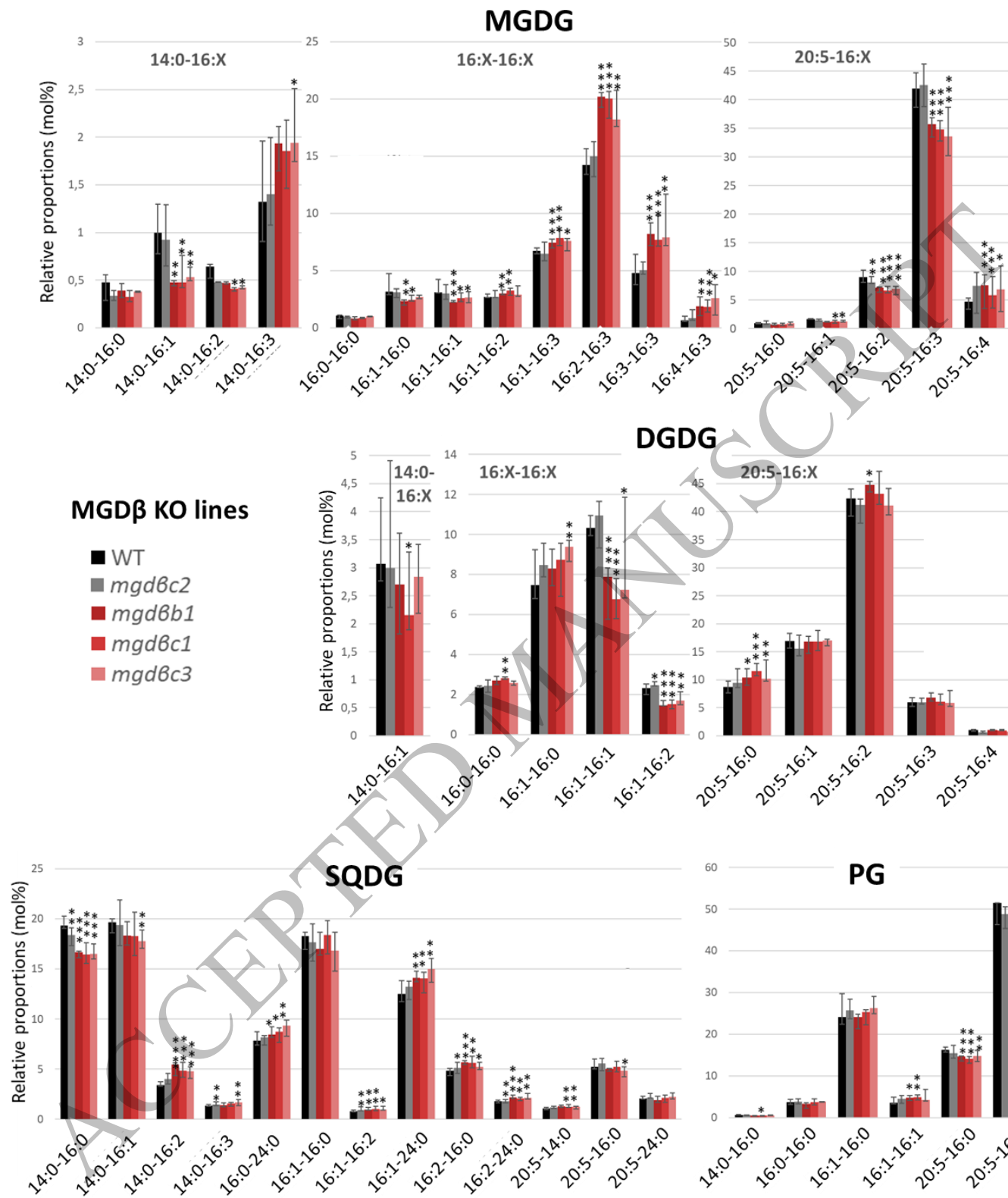


Figure 6: Impact of MGDβ mutations on MGDG, DGDG, SQG and PG molecular species. Lipid molecular species were analyzed by LC-MS/MS. The *mgdβc2* mutant contains a silent mutation and was used as a control. Each result is the median of six biological replicates \pm min and max values. (*), P -value $< 5.10^{-2}$; (**), P -value $< 1.10^{-2}$; (***), P -value $< 1.10^{-3}$, based on an unpaired multiple t test. DGDG, digalactosyldiacylglycerol, MGDG, monogalactosyldiacylglycerol, PG, phosphatidylglycerol, SQDG, sulfoquinovosyldiacylglycerol.

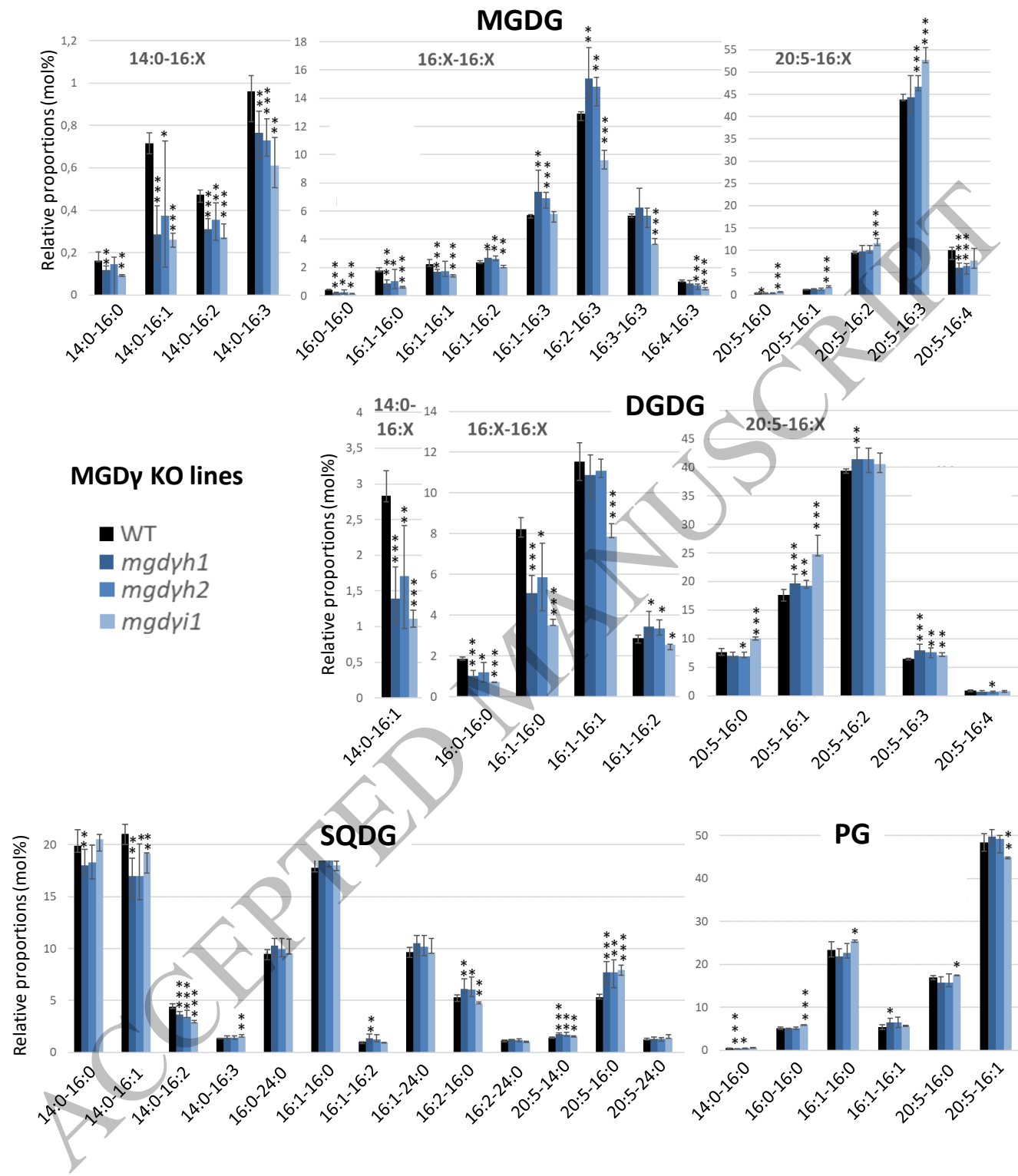


Figure 7: Impact of MGDy mutations on MGDG, DGDG, SQG and PG molecular species. Lipid molecular species were analysed by LC-MS/MS. Each result is the median of six biological replicates \pm min and max values. (*), *P*-value $< 5.10^{-2}$; (**), *P*-value $< 1.10^{-2}$; (***), *P*-value $< 1.10^{-3}$, based on an unpaired multiple t test. DGDG, digalactosyldiacylglycerol, MGDG, monogalactosyldiacylglycerol, PG, phosphatidylglycerol, SQDG, sulfoquinovosyldiacylglycerol.

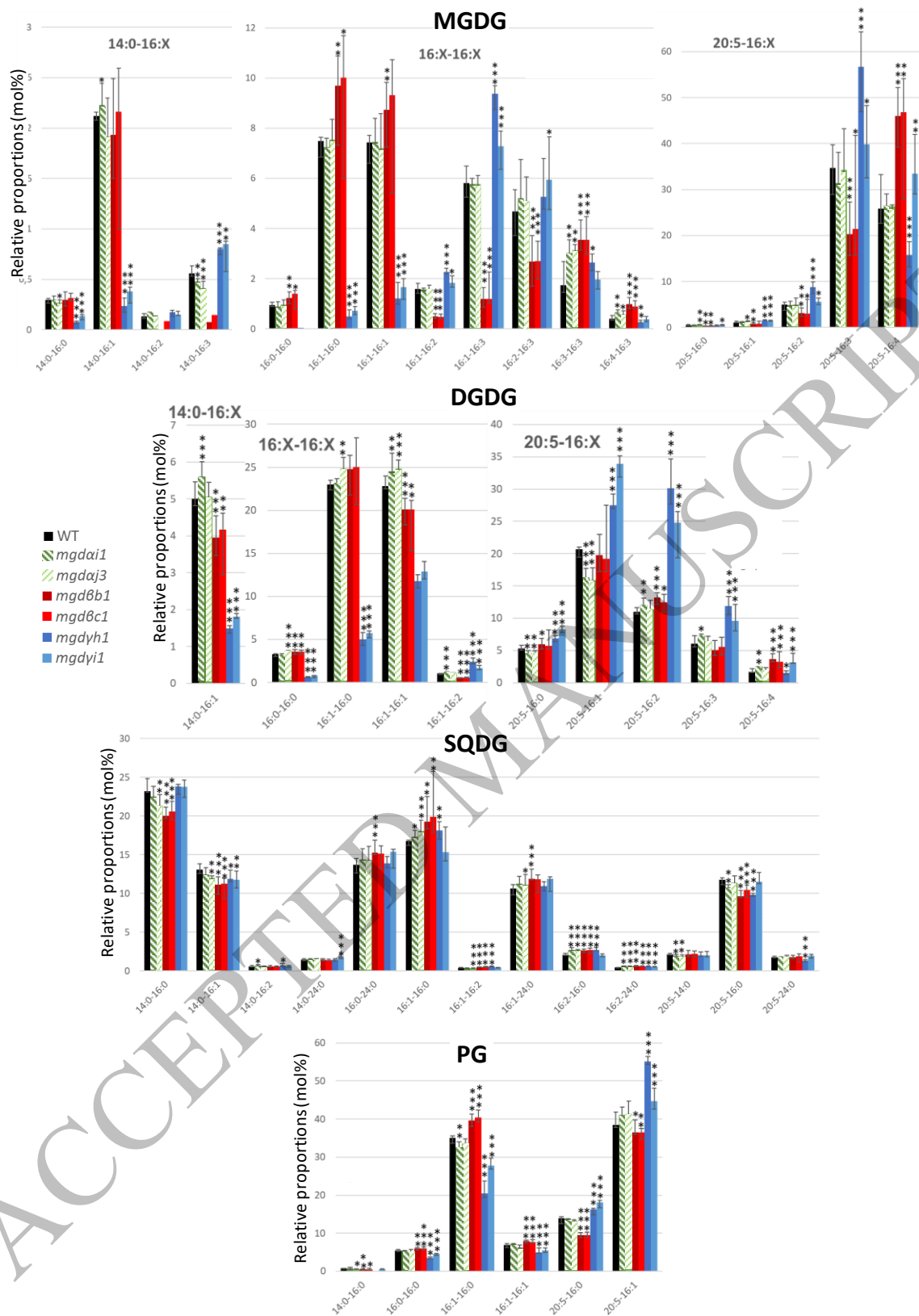


Figure 8: Impact of MGD mutations on the molecular species constituting MGDG, DGDG, SQDG and PG in nitrogen-deprived *P. tricornutum* cells. Lipid molecular species in MGDG, DGDG, SQDG and PG were analyzed by LC-MS/MS. Each result is the median of six biological replicates \pm min and max values. (*), P -value $< 5.10^{-2}$; (**), P -value $< 1.10^{-2}$; (***), P -value $< 1.10^{-3}$, based on an unpaired multiple t test. DGDG, digalactosyldiacylglycerol, MGDG, monogalactosyldiacylglycerol, PG, phosphatidylglycerol, SQDG, sulfoquinovosyldiacylglycerol.

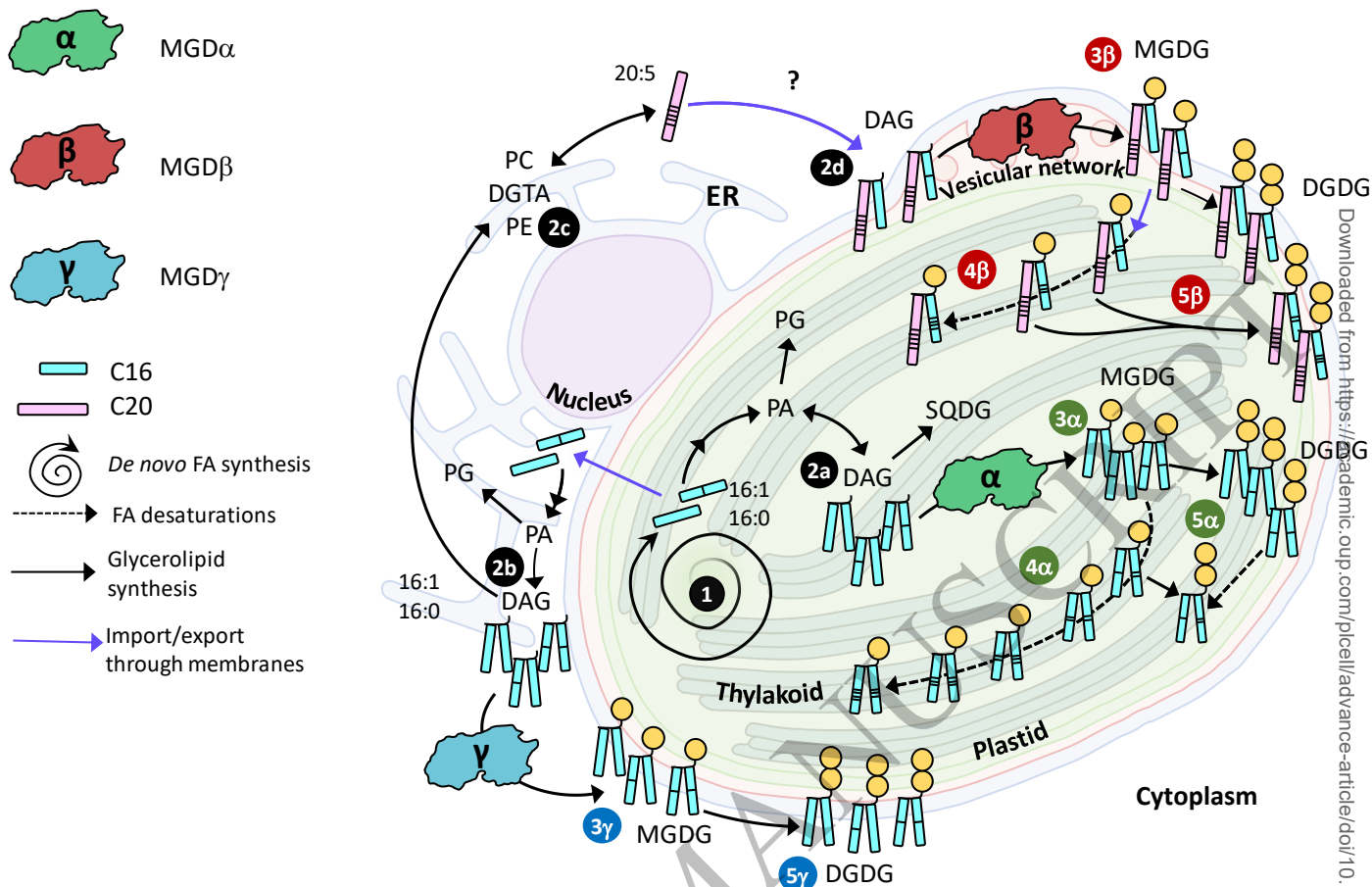


Figure 9. Function of MGD α , MGD β and MGD γ in *P. tricornutum* cell. The model summarizes all compartmentalization, structural and functional data obtained in this study. FA are synthesized in the stroma as 16:0- and 16:1-ACP (1), where they can be used for the synthesis of DAG (2a). Alternatively, FA are exported to extraplastidial membranes where they can serve for the synthesis of DAG (2b) and phospholipids such as PC and PE, or the betaine lipid DGTA (2c). PC and DGTA serve as platforms to produce 20:5 (2c), used inside the plastid (at an unknown membrane site) for the production of 20:5-containing DAG (2d). MGD α is localized at the level of thylakoids, and functional studies reported here are consistent with a utilization of DAG available inside the plastid (3 α), to form MGDG species rapidly desaturated (16:1 \square 16:4) by plastid desaturases (4 α). Based on obtained phenotype of KO and overexpressing lines, MGDG produced by MGD α is likely used by a specific DGD enzyme (5 α) to form DGDG. MGD β is localized at the level of the blob-like structure at the PPM, and functional studies reported here are consistent with a utilization of DAG either produced inside the plastid or with 20:5 acyls imported from the ER by a still uncharacterized Ω -pathway (3 β). MGDG produced can be desaturated (16:1 \square 16:4) by plastid desaturases (4 β). Based on obtained phenotype of KO and overexpressing lines, MGDG produced by MGD β is likely used by a specific DGD enzyme (5 β) to form DGDG. MGD γ is localized at the periphery of the plastid, possibly partitioned between the ER and the EpM. Functional studies are consistent with a utilization of DAG containing FA freshly exported from the plastid (3 γ), producing MGDG at a location far from plastid desaturases. Based on obtained phenotype of KO and overexpressing lines, MGDG produced by MGD γ is also likely associated to a specific DGD enzyme (5 γ). In nitrogen shortage, MGD α protein level decreases drastically, and MGDG production relies on MGD β and MGD γ . When one MGD is knocked-out, compensation mechanisms rely on intact isoforms located in distinct membrane systems, requiring rapid inward and outward trafficking of glycerolipids including MGDG, through the four membranes limiting the plastid. KO, knock-out; DAG, diacylglycerol; DGTA, diacylglyceryl hydroxymethyltrimethyl- β -alanine; DGD, digalactosyldiacylglycerol synthase; DGDG, digalactosyldiacylglycerol; ER, endoplasmic reticulum; FA; fatty acid; MGD, monogalactosyldiacylglycerol synthase; MGDG, monogalactosyldiacylglycerol; PC, phosphatidylcholine.

Downloaded from https://academic.oup.com/plcell/advance-article/doi/10.1093/plcell/ckab277/6293375 by University of Cambridge user on 02 July 2021

Parsed Citations

Abida, H., Dolch, L.J., Mei, C., Villanova, V., Conte, M., Block, M.A., Finazzi, G., Bastien, O., Tirichine, L., Bowler, C., Rebeille, F., Petroustos, D., Jouhet, J., and Marechal, E. (2015). Membrane glycerolipid remodeling triggered by nitrogen and phosphorus starvation in *Phaeodactylum tricornutum*. *Plant Physiol* 167, 118-136.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Almagro Armenteros, J.J., Tsirigos, K.D., Sonderby, C.K., Petersen, T.N., Winther, O., Brunak, S., von Heijne, G., and Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* 37, 420-423.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Apt, K.E., Zaslavkaia, L., Lippmeier, J.C., Lang, M., Kilian, O., Wetherbee, R., Grossman, A.R., and Kroth, P.G. (2002). In vivo characterization of diatom multipartite plastid targeting signals. *J Cell Sci* 115, 4061-4069.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Awai, K., Marechal, E., Block, M.A., Brun, D., Masuda, T., Shimada, H., Takamiya, K., Ohta, H., and Joyard, J. (2001). Two types of MGDG synthase genes, found widely in both 16:3 and 18:3 plants, differentially mediate galactolipid syntheses in photosynthetic and nonphotosynthetic tissues in *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* 98, 10960-10965.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Azadi-Chegeni, F., Thallmair, S., Ward, M.E., Perin, G., Marrink, S.J., Baldus, M., Morosinotto, T., and Pandit, A. (2022). Protein dynamics and lipid affinity of monomeric, zeaxanthin-binding LHCII in thylakoid membranes. *Biophysical Journal* 121, 396-409.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Benning, C., and Ohta, H. (2005). Three enzyme systems for galactoglycerolipid biosynthesis are coordinately regulated in plants. *Journal of Biological Chemistry* 280, 2397-2400.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Benoiston, A.S., Ibarbalz, F.M., Bittner, L., Guidi, L., Jahn, O., Dutkiewicz, S., and Bowler, C. (2017). The evolution of diatoms and their biogeochemical functions. *Philos Trans R Soc Lond B Biol Sci* 372.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Billey, E., Hafidh, S., Cruz-Gallardo, I., Litholdo, C.G., Jean, V., Carpentier, M.C., Picart, C., Kumar, V., Kulichova, K., Marechal, E., Honys, D., Conte, M.R., Deragon, J.M., and Bousquet-Antonelli, C. (2021a). LARP6C orchestrates posttranscriptional reprogramming of gene expression during hydration to promote pollen tube guidance. *Plant Cell*.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Billey, E., Magneschi, L., Leterme, S., Bedhomme, M., Andres-Robin, A., Poulet, L., Michaud, M., Finazzi, G., Dumas, R., Crouzy, S., Lueffer, F., Fourage, L., Rebeille, F., Amato, A., Collin, S., Jouhet, J., and Marechal, E. (2021b). Characterization of the Bubblegum acyl-CoA synthetase of *Microchloropsis gaditana*. *Plant Physiol* 185, 815-835.

Google Scholar: [Author Only](#) [Author and Title](#)

Botte, C., Jeanneau, C., Snajdrova, L., Bastien, O., Imbert, A., Breton, C., and Marechal, E. (2005). Molecular modeling and site-directed mutagenesis of plant chloroplast monogalactosyldiacylglycerol synthase reveal critical residues for activity. *J Biol Chem* 280, 34691-34701.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Boudiere, L., Michaud, M., Petroustos, D., Rebeille, F., Falconet, D., Bastien, O., Roy, S., Finazzi, G., Rolland, N., Jouhet, J., Block, M.A., and Marechal, E. (2014). Glycerolipids in photosynthesis: composition, synthesis and trafficking. *Biochim Biophys Acta* 1837, 470-480.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Bowler, C., Allen, A.E., Badger, J.H., Grimwood, J., Jabbari, K., Kuo, A., Maheswari, U., Martens, C., Maumus, F., Otilar, R.P., Rayko, E., Salamov, A., Vandepoele, K., Beszteri, B., Gruber, A., Heijde, M., Katinka, M., Mock, T., Valentin, K., Verret, F., Berges, J.A., Brownlee, C., Cadoret, J.P., Chiovitti, A., Choi, C.J., Coesel, S., De Martino, A., Detter, J.C., Durkin, C., Falciatore, A., Fournet, J., Haruta, M., Huysman, M.J., Jenkins, B.D., Jiroutova, K., Jorgensen, R.E., Joubert, Y., Kaplan, A., Kroger, N., Kroth, P.G., La Roche, J., Lindquist, E., Lommer, M., Martin-Jezequel, V., Lopez, P.J., Lucas, S., Mangogna, M., McGinnis, K., Medlin, L.K., Monsant, A., Oudot-Le Secq, M.P., Napoli, C., Obornik, M., Parker, M.S., Petit, J.L., Porcel, B.M., Poulsen, N., Robison, M., Rychlewski, L., Rynearson, T.A., Schmutz, J., Shapiro, H., Siaut, M., Stanley, M., Sussman, M.R., Taylor, A.R., Vardi, A., von Dassow, P., Vyverman, W., Willis, A., Wyrwicz, L.S., Rokhsar, D.S., Weissenbach, J., Armbrust, E.V., Green, B.R., Van de Peer, Y., and Grigoriev, I.V. (2008). The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature* 456, 239-244.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Brinkman, E.K., Chen, T., Amendola, M., and van Steensel, B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* 42, e168.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Browse, J., Warwick, N., Somerville, C.R., and Slack, C.R. (1986). Fluxes through the prokaryotic and eukaryotic pathways of lipid synthesis in the 16-3 Plant *Arabidopsis thaliana*. *Biochemical Journal* 235, 25-31.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Bullmann, L., Haarmann, R., Mirus, O., Bredemeier, R., Hempel, F., Maier, U.G., and Schleiff, E. (2010). Filling the gap, evolutionarily conserved Omp85 in plastids of chromalveolates. *J Biol Chem* 285, 6848-6856.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Buseman, C.M., Tamura, P., Sparks, A.A., Baughman, E.J., Maatta, S., Zhao, J., Roth, M.R., Esch, S.W., Shah, J., Williams, T.D., and Welti, R. (2006). Wounding stimulates the accumulation of glycerolipids containing oxophytodienoic acid and dinor-oxophytodienoic acid in *Arabidopsis* leaves. *Plant Physiol* 142, 28-39.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Cavalier-Smith, T. (2018). Kingdom Chromista and its eight phyla: a new synthesis emphasising periplastid protein targeting, cytoskeletal and periplastid evolution, and ancient divergences. *Protoplasma* 255, 297-357.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Conte, M., Lupette, J., Seddiki, K., Mei, C., Dolch, L.J., Gros, V., Barette, C., Rebeille, F., Jouhet, J., and Marechal, E. (2018). Screening for biologically annotated drugs that trigger triacylglycerol accumulation in the diatom *Phaeodactylum*. *Plant Physiol*.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Corteggiani Carpinelli, E., Telatin, A., Vitulo, N., Forcato, C., D'Angelo, M., Schiavon, R., Vezzi, A., Giacometti, G.M., Morosinotto, T., and Valle, G. (2014). Chromosome scale genome assembly and transcriptome profiling of *Nannochloropsis gaditana* in nitrogen depletion. *Molecular plant* 7, 323-335.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Criscuolo, A., and Gribaldo, S. (2010). BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC evolutionary biology* 10, 210.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Daboussi, F., Leduc, S., Marechal, A., Dubois, G., Guyot, V., Perez-Michaut, C., Amato, A., Falcatore, A., Juillerat, A., Beurdeley, M., Voytas, D.F., Cavarec, L., and Duchateau, P. (2014). Genome engineering empowers the diatom *Phaeodactylum tricorutum* for biotechnology. *Nat Commun* 5, 3831.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

De Riso, V., Raniello, R., Maumus, F., Rogato, A., Bowler, C., and Falcatore, A. (2009). Gene silencing in the marine diatom *Phaeodactylum tricorutum*. *Nucleic Acids Res* 37, e96.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahe, F., Logares, R., Lara, E., Berney, C., Le Bescot, N., Probert, I., Carmichael, M., Poulain, J., Romac, S., Colin, S., Aury, J.M., Bittner, L., Chaffron, S., Dunthorn, M., Engelen, S., Flegontova, O., Guidi, L., Horak, A., Jaillon, O., Lima-Mendez, G., Lukes, J., Malviya, S., Morard, R., Mulo, M., Scalco, E., Siano, R., Vincent, F., Zingone, A., Dimier, C., Picheral, M., Searson, S., Kandels-Lewis, S., Tara Oceans, C., Acinas, S.G., Bork, P., Bowler, C., Gorsky, G., Grimsley, N., Hingamp, P., Iudicone, D., Not, F., Ogata, H., Pesant, S., Raes, J., Sieracki, M.E., Speich, S., Stemmann, L., Sunagawa, S., Weissenbach, J., Wincker, P., and Karsenti, E. (2015). Ocean plankton. Eukaryotic plankton diversity in the sunlit ocean. *Science* 348, 1261605.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Deme, B., Cataye, C., Block, M.A., Marechal, E., and Jouhet, J. (2014). Contribution of galactoglycerolipids to the 3-dimensional architecture of thylakoids. *FASEB J* 28, 3373-3383.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Dolch, L.J., and Marechal, E. (2015). Inventory of fatty acid desaturases in the pennate diatom *Phaeodactylum tricorutum*. *Marine drugs* 13, 1317-1339.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Dolch, L.J., Rak, C., Perin, G., Tourcier, G., Broughton, R., Leterrier, M., Morosinotto, T., Tellier, F., Faure, J.D., Falconet, D., Jouhet, J., Sayanova, O., Beaudoin, F., and Marechal, E. (2017a). A Palmitic Acid Elongase Affects Eicosapentaenoic Acid and Plastidial Monogalactosyldiacylglycerol Levels in *Nannochloropsis*. *Plant Physiol* 173, 742-759.

Google Scholar: [Author Only](#) [Author and Title](#)

Dolch, L.J., Lupette, J., Tourcier, G., Bedhomme, M., Collin, S., Magneschi, L., Conte, M., Seddiki, K., Richard, C., Corre, E., Fourage, L., Laeuffer, F., Richards, R., Reith, M., Rebeille, F., Jouhet, J., McGinn, P., and Marechal, E. (2017b). Nitric Oxide Mediates Nitrite-Sensing and Acclimation and Triggers a Remodeling of Lipids. *Plant Physiol* 175, 1407-1423.

Google Scholar: [Author Only](#) [Author and Title](#)

Dubots, E., Audry, M., Yamaryo, Y., Bastien, O., Ohta, H., Breton, C., Marechal, E., and Block, M.A. (2010). Activation of the chloroplast monogalactosyldiacylglycerol synthase MGD1 by phosphatidic acid and phosphatidylglycerol. *J Biol Chem* 285, 6003-6011.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32, 1792-

1797.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Emanuelsson, O., Brunak, S., von Heijne, G., and Nielsen, H. (2007). Locating proteins in the cell using TargetP, SignalP and related tools. *Nature protocols* 2, 953-971.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Falciatore, A., Alcalà, M.R., Croot, P., and Bowler, C. (2000). Perception of Environmental Signals by a Marine Diatom. *Science* 288, 2363.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Falciatore, A., Casotti, R., Leblanc, C., Abrescia, C., and Bowler, C. (1999). Transformation of Nonselectable Reporter Genes in Marine Diatoms. *Mar Biotechnol* (NY) 1, 239-251.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Fawley, M.W., Jameson, I., and Fawley, K.P. (2015). The phylogeny of the genus *Nannochloropsis* (Monodopsidaceae, Eustigmatophyceae), with descriptions of *N. australis* sp. nov. and *Microchloropsis* gen. nov. *Phycologia* 54, 545-552.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Flori, S., Jouneau, P.H., Finazzi, G., Marechal, E., and Falconet, D. (2016). Ultrastructure of the Periplastidial Compartment of the Diatom *Phaeodactylum tricornutum*. *Protist* 167, 254-267.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Flori, S., Jouneau, P.H., Gallet, B., Estrozi, L.F., Moriscot, C., Schoehn, G., Finazzi, G., and Falconet, D. (2018). Imaging Plastids in 2D and 3D: Confocal and Electron Microscopy. *Methods Mol Biol* 1829, 113-122.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Folch, J., Lees, M., and Sloane Stanley, G.H. (1957). A simple method for the isolation and purification of total lipides from animal tissues. *J Biol Chem* 226, 497-509.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Garab, G., Yaguzhinsky, L.S., Dlouhy, O., Nesterov, S.V., Spunda, V., and Gasanoff, E.S. (2022). Structural and functional roles of non-bilayer lipid phases of chloroplast thylakoid membranes and mitochondrial inner membranes. *Prog Lipid Res* 86, 101163.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Gascuel, O. (1997). BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Mol Biol Evol* 14, 685-695.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Goss, R., and Jakob, T. (2010). Regulation and function of xanthophyll cycle-dependent photoprotection in algae. *Photosynth Res* 106, 103-122.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Gould, S.B., Maier, U.G., and Martin, W.F. (2015). Protein Import and the Origin of Red Complex Plastids. *Curr Biol* 25, R515-R521.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Grosche, C., Hempel, F., Bolte, K., Zauner, S., and Maier, U.G. (2014). The periplastidial compartment: a naturally minimized eukaryotic cytoplasm. *Curr Opin Microbiol* 22, 88-93.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Gruber, A., Rocap, G., Kroth, P.G., Armbrust, E.V., and Mock, T. (2015). Plastid proteome prediction for diatoms and other algae with secondary plastids of the red lineage. *Plant J* 81, 519-528.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Gruber, A., Vugrinec, S., Hempel, F., Gould, S.B., Maier, U.G., and Kroth, P.G. (2007). Protein targeting into complex diatom plastids: functional characterisation of a specific targeting motif. *Plant Mol Biol* 64, 519-530.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Gschloessl, B., Guermeur, Y., and Cock, J.M. (2008). HECTAR: a method to predict subcellular targeting in heterokonts. *BMC Bioinformatics* 9, 393.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Gu, X., Cao, L., Wu, X., Li, Y., Hu, Q., and Han, D. (2021). A Lipid Bodies-Associated Galactosyl Hydrolase Is Involved in Triacylglycerol Biosynthesis and Galactolipid Turnover in the Unicellular Green Alga *Chlamydomonas reinhardtii*. *Plants* (Basel) 10.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Guiry, M.D. (2012). How Many Species of Algae Are There? *J Phycol* 48, 1057-1063.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Hempel, F., Bullmann, L., Lau, J., Zauner, S., and Maier, U.G. (2009). ERAD-derived preprotein transport across the second

outermost plastid membrane of diatoms. *Mol Biol Evol* 26, 1781-1790.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Hori, K., Nobusawa, T., Watanabe, T., Madoka, Y., Suzuki, H., Shibata, D., Shimojima, M., and Ohta, H. (2016). Tangled evolutionary processes with commonality and diversity in plastidial glycolipid synthesis in photosynthetic organisms. *Biochim Biophys Acta* 1861, 1294-1308.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Horton, P., Park, K.J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., and Nakai, K. (2007). WoLF PSORT: protein localization predictor. *Nucleic Acids Res* 35, W585-587.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Huang, T., Pan, Y., Marechal, E., and Hu, H. (2024). Proteomes reveal the lipid metabolic network in the complex plastid of *Phaeodactylum tricornutum*. *Plant J* 117, 385-403.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Iwai, M., Yamada-Oshima, Y., Asami, K., Kanamori, T., Yuasa, H., Shimojima, M., and Ohta, H. (2021). Recycling of the major thylakoid lipid MGDG and its role in lipid homeostasis in *Chlamydomonas reinhardtii*. *Plant Physiol* 187, 1341-1356.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Janson, G., Zhang, C., Prado, M.G., and Paiardini, A. (2017). PyMod 2.0: improvements in protein sequence-structure analysis and homology modeling within PyMOL. *Bioinformatics* 33, 444-446.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Jensen, P.E., and Leister, D. (2014). Chloroplast evolution, structure and functions. *F1000Prime Rep* 6, 40.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Jouhet, J., Lupette, J., Clerc, O., Magneschi, L., Bedhomme, M., Collin, S., Roy, S., Maréchal, E., and Rébeillé, F. (2017). LC-MS/MS versus TLC plus GC methods: Consistency of glycerolipid and fatty acid profiles in microalgae and higher plant cells and effect of a nitrogen starvation. *PLOS ONE* 12, e0182423.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Jumper, J., and Hassabis, D. (2022). Protein structure predictions to atomic accuracy with AlphaFold. *Nat Methods* 19, 11-12.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohli, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., and Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583-589.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Karas, B.J., Diner, R.E., Lefebvre, S.C., McQuaid, J., Phillips, A.P., Noddings, C.M., Brunson, J.K., Valas, R.E., Deerinck, T.J., Jablanovic, J., Gillard, J.T., Beeri, K., Ellisman, M.H., Glass, J.I., Hutchison, C.A., 3rd, Smith, H.O., Venter, J.C., Allen, A.E., Dupont, C.L., and Weyman, P.D. (2015). Designer diatom episomes delivered by bacterial conjugation. *Nat Commun* 6, 6925.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols* 10, 845-858.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Kilian, O., and Kroth, P.G. (2005). Identification and characterization of a new conserved motif within the presequence of proteins targeted into complex diatom plastids. *Plant J* 41, 175-183.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Kobayashi, K., Nakamura, Y., and Ohta, H. (2009a). Type A and type B monogalactosyldiacylglycerol synthases are spatially and functionally separated in the plastids of higher plants. *Plant Physiology and Biochemistry* 47, 518-525.

Google Scholar: [Author Only](#) [Author and Title](#)

Kobayashi, K., Kondo, M., Fukuda, H., Nishimura, M., and Ohta, H. (2007a). Galactolipid biosynthesis is essential for proper chloroplast biogenesis and embryogenesis. *Photosynthesis Research* 91, 216-216.

Google Scholar: [Author Only](#) [Author and Title](#)

Kobayashi, K., Kondo, M., Fukuda, H., Nishimura, M., and Ohta, H. (2007b). Galactolipid synthesis in chloroplast inner envelope is essential for proper thylakoid biogenesis, photosynthesis, and embryogenesis. *Proc Natl Acad Sci U S A* 104, 17216-17221.

Google Scholar: [Author Only](#) [Author and Title](#)

Kobayashi, K., Awai, K., Nakamura, M., Nagatani, A., Masuda, T., and Ohta, H. (2009b). Type-B monogalactosyldiacylglycerol synthases are involved in phosphate starvation-induced lipid remodeling, and are crucial for low-phosphate adaptation. *Plant Journal* 57, 322-331.

Google Scholar: [Author Only Author and Title](#)

Kooistra, W.H., De Stefano, M., Mann, D.G., and Medlin, L.K. (2003). The phylogeny of the diatoms. *Prog Mol Subcell Biol* 33, 59-97.

Google Scholar: [Author Only Title Only Author and Title](#)

Kooistra, W.H., Gersonde, R., Medlin, L.K., and Mann, D.G. (2007). The Origin and Evolution of the Diatoms: Their Adaptation to a Planktonic Existence. In *Evolution of Primary Producers in the Sea* P.G. Falkowski and A.H. Knoll, eds (Cambridge, U.S.A: Academic Press), pp. 207-249.

Google Scholar: [Author Only Title Only Author and Title](#)

Kroth, P.G., Bones, A.M., Daboussi, F., Ferrante, M.I., Jaubert, M., Kolot, M., Nymark, M., Rio Bartulos, C., Ritter, A., Russo, M.T., Serif, M., Winge, P., and Falciatore, A. (2018). Genome editing in diatoms: achievements and goals. *Plant Cell Rep* 37, 1401-1408.

Google Scholar: [Author Only Title Only Author and Title](#)

Kumar, S., Stecher, G., Li, M., Nknyaz, C., and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol* 35, 1547-1549.

Google Scholar: [Author Only Title Only Author and Title](#)

Lang, M., Apt, K.E., and Kroth, P.G. (1998). Protein transport into "complex" diatom plastids utilizes two different targeting signals. *J Biol Chem* 273, 30973-30978.

Google Scholar: [Author Only Title Only Author and Title](#)

Le, S.Q., Lartillot, N., and Gascuel, O. (2008). Phylogenetic mixture models for proteins. *Philos Trans R Soc Lond B Biol Sci* 363, 3965-3976.

Google Scholar: [Author Only Title Only Author and Title](#)

Lemoine, F., Correia, D., Lefort, V., Doppelt-Azeroual, O., Mareuil, F., Cohen-Boulakia, S., and Gascuel, O. (2019). NGPhylogeny.fr: new generation phylogenetic services for non-specialists. *Nucleic Acids Res* 47, W260-W265.

Google Scholar: [Author Only Title Only Author and Title](#)

Letunic, I., and Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 49, W293-W296.

Google Scholar: [Author Only Title Only Author and Title](#)

Leyland, B., Zarka, A., Didi-Cohen, S., Boussiba, S., and Khozin-Goldberg, I. (2020). High Resolution Proteome of Lipid Droplets Isolated from the Pennate Diatom *Phaeodactylum tricornutum* (Bacillariophyceae) Strain pt4 provides mechanistic insights into complex intracellular coordination during nitrogen deprivation. *J Phycol.*

Google Scholar: [Author Only Title Only Author and Title](#)

Li-Beisson, Y., Shorrosh, B., Beisson, F., Andersson, M.X., Arondel, V., Bates, P.D., Baud, S., Bird, D., Debono, A., Durrett, T.P., Franke, R.B., Graham, I.A., Katayama, K., Kelly, A.A., Larson, T., Markham, J.E., Miquel, M., Molina, I., Nishida, I., Rowland, O., Samuels, L., Schmid, K.M., Wada, H., Welti, R., Xu, C., Zallot, R., and Ohlrogge, J. (2010). Acyl-lipid metabolism. In *The Arabidopsis Book*, pp. e0133.

Google Scholar: [Author Only Title Only Author and Title](#)

Li, X., Moellering, E.R., Liu, B., Johnny, C., Fedewa, M., Sears, B.B., Kuo, M.H., and Benning, C. (2012). A galactoglycerolipid lipase is required for triacylglycerol accumulation and survival following nitrogen deprivation in *Chlamydomonas reinhardtii*. *Plant Cell* 24, 4670-4686.

Google Scholar: [Author Only Title Only Author and Title](#)

Liu, X.J., Hempel, F., Stork, S., Bolte, K., Moog, D., Heimerl, T., Maier, U.G., and Zauner, S. (2016). Addressing various compartments of the diatom model organism via sub-cellular marker proteins. *Algal Res* 20, 249-257.

Google Scholar: [Author Only Title Only Author and Title](#)

Lupette, J., Tardif, M., Brugiére, S., Coute, Y., Salvaing, J., and Marechal, E. (2022). Quantitative proteomic analyses reveal the impact of nitrogen starvation on the proteome of the model diatom *Phaeodactylum tricornutum*. *Proteomics* 22, e2200155.

Google Scholar: [Author Only Title Only Author and Title](#)

Lupette, J., Jaussaud, A., Seddiki, K., Morabito, C., Brugiére, S., Schaller, H., Kuntz, M., Putaux, J.L., Juneau, P.H., Rebelle, F., Falconet, D., Coute, Y., Jouhet, J., Tardif, M., Salvaing, J., and Marechal, E. (2019). The architecture of lipid droplets in the diatom *Phaeodactylum tricornutum*. *Algal Res* 38.

Google Scholar: [Author Only Title Only Author and Title](#)

Makshakova, O., Breton, C., and Perez, S. (2020). Unraveling the complex enzymatic machinery making a key galactolipid in chloroplast membrane: a multiscale computer simulation. *Scientific reports* 10, 13514.

Google Scholar: [Author Only Title Only Author and Title](#)

Mann, D.G. (1999). The species concept in diatoms. *Phycologia* 38, 437-495.

Google Scholar: [Author Only Title Only Author and Title](#)

- Mann, D.G., and Droop, S.J.M. (1996). Biodiversity, biogeography and conservation of diatoms. *Hydrobiologia*, 19–32.
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)
- Marechal, E. (2024). How Did Thylakoids Emerge in Cyanobacteria, and How Were the Primary Chloroplast and Chromatophore Acquired? *Methods Mol Biol* 2776, 3-20.
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)
- Martens, C., Vandepoele, K., and Van de Peer, Y. (2008). Whole-genome analysis reveals molecular innovations and evolutionary transitions in chromalveolate species. *Proc Natl Acad Sci U S A* 105, 3427-3432.
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)
- Mirdita, M., Schutze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nature Methods* 19, 679-+.
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)
- Moog, D., Stork, S., Zauner, S., and Maier, U.G. (2011). In silico and in vivo investigations of proteins of a minimized eukaryotic cytoplasm. *Genome Biol Evol* 3, 375-382.
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)
- Murakami, R., and Hashimoto, H. (2009). Unusual nuclear division in *Nannochloropsis oculata* (Eustigmatophyceae, Heterokonta) which may ensure faithful transmission of secondary plastids. *Protist* 160, 41-49.
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)
- Nitenberg, M., Makshakova, O., Rocha, J., Perez, S., Marechal, E., Block, M.A., Girard-Egrot, A., and Breton, C. (2020). Mechanism of activation of plant monogalactosyldiacylglycerol synthase 1 (MGD1) by phosphatidylglycerol. *Glycobiology* 30, 396-406.
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)
- Parks, M.B., Nakov, T., Ruck, E.C., Wickett, N.J., and Alverson, A.J. (2018). Phylogenomics reveals an extensive history of genome duplication in diatoms (Bacillariophyta). *Am J Bot* 105, 330-347.
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)
- Petroutsos, D., Amiar, S., Abida, H., Dolch, L.J., Bastien, O., Rebeille, F., Jouhet, J., Falconet, D., Block, M.A., McFadden, G.I., Bowler, C., Botte, C., and Marechal, E. (2014). Evolution of galactoglycerolipid biosynthetic pathways--from cyanobacteria to primary plastids and from primary to secondary plastids. *Prog Lipid Res* 54, 68-85.
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)
- Rastogi, A., Murik, O., Bowler, C., and Tirichine, L. (2016). PhytoCRISP-Ex: a web-based and stand-alone application to find specific target sequences for CRISPR/CAS editing. *BMC Bioinformatics* 17, 261.
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)
- Reyes-Prieto, A., Weber, A.P., and Bhattacharya, D. (2007). The origin and establishment of the plastid in algae and plants. *Annu Rev Genet* 41, 147-168.
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)
- Rocha, J., Sarkis, J., Thomas, A., Pitou, L., Radzimanowski, J., Audry, M., Chazalet, V., de Sanctis, D., Palcic, M.M., Block, M.A., Girard-Egrot, A., Marechal, E., and Breton, C. (2016). Structural insights and membrane binding properties of MGD1, the major galactolipid synthase in plants. *Plant J* 85, 622-633.
Google Scholar: [Author Only](#) [Author and Title](#)
- Ruban, A., Lavaud, J., Rousseau, B., Guglielmi, G., Horton, P., and Etienne, A.L. (2004). The super-excess energy dissipation in diatom algae: comparative analysis with higher plants. *Photosynth Res* 82, 165-175.
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)
- Sato, N., and Awai, K. (2017). "Prokaryotic Pathway" Is Not Prokaryotic: Noncyanobacterial Origin of the Chloroplast Lipid Biosynthetic Pathway Revealed by Comprehensive Phylogenomic Analysis. *Genome Biol Evol* 9, 3162-3178.
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)
- Seródio, J., and Lavaud, J. (2020). Diatoms and Their Ecological Importance. In *Life Below Water*, W. Leal Filho, A.M. Azul, L. Brandli, A Lange Salvia, and T. Wall, eds (Cham: Springer International Publishing), pp. 1-9.
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)
- Shang, S., Liu, R., Luo, L., Li, X., Zhang, S., Zhang, Y., Zheng, P., Chen, Z., and Wang, B. (2022). Functional Characterization of the Monogalactosyldiacylglycerol Synthase Gene ptMGD2 in the Diatom *Phaeodactylum tricornutum*. *Front. Mar. Sci.* 9, e874448.
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)
- Siaut, M., Heijde, M., Mangogna, M., Montsant, A., Coesel, S., Allen, A., Manfredonia, A., Falcatore, A., and Bowler, C. (2007). Molecular toolbox for studying diatom biology in *Phaeodactylum tricornutum*. *Gene* 406, 23-35.
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)
- Smith, R., Jouhet, J., Gandini, C., Nekrasov, V., Marechal, E., Napier, J.A., and Sayanova, O. (2021). Plastidial acyl carrier protein

Delta9-desaturase modulates eicosapentaenoic acid biosynthesis and triacylglycerol accumulation in *Phaeodactylum tricoratum*. *Plant J* 106, 1247-1259.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Sorhannus, U. (2007). A nuclear-encoded small-subunit ribosomal RNA timescale for diatom evolution. *Mar Micropaleontol* 65, 1-12.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Tanaka, A., De Martino, A., Amato, A., Montsant, A., Mathieu, B., Rostaing, P., Tirichine, L., and Bowler, C. (2015). Ultrastructure and Membrane Traffic During Cell Division in the Marine Pennate Diatom *Phaeodactylum tricoratum*. *Protist* 166, 506-521.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Vieler, A., Wu, G., Tsai, C.H., Bullard, B., Cornish, A.J., Harvey, C., Reca, I.B., Thornburg, C., Achawanantakun, R., Buehl, C.J., Campbell, M.S., Cavalier, D., Childs, K.L., Clark, T.J., Deshpande, R., Erickson, E., Armenia Ferguson, A., Handee, W., Kong, Q., Li, X., Liu, B., Lundback, S., Peng, C., Roston, R.L., Sanjaya, Simpson, J.P., Terbush, A., Warakanont, J., Zauner, S., Farre, E.M., Hegg, E.L., Jiang, N., Kuo, M.H., Lu, Y., Niyogi, K.K., Ohlrogge, J., Osteryoung, K.W., Shachar-Hill, Y., Sears, B.B., Sun, Y., Takahashi, H., Yandell, M., Shiu, S.H., and Benning, C. (2012). Genome, functional gene annotation, and nuclear transformation of the heterokont oleaginous alga *Nannochloropsis oceanica* CCMP1779. *PLoS Genet* 8, e1003064.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Yates, A.D., Allen, J., Amode, R.M., Azov, A.G., Barba, M., Becerra, A., Bhai, J., Campbell, L.I., Martinez, M.C., Chakiachvili, M., Chougule, K., Christensen, M., Contreras-Moreira, B., Cuzick, A., Fioretto, L.D., Davis, P., De Silva, N.H., Diamantakis, S., Dyer, S., Elser, J., Filippi, C.V., Gall, A., Grigoriadis, D., Guijarro-Clarke, C., Gupta, P., Hammond-Kosack, K.E., Howe, K.L., Jaiswal, P., Kaikala, V., Kumar, V., Kumari, S., Langridge, N., Le, T., Luybaert, M., Maslen, G.L., Maurel, T., Moore, B., Muffato, M., Mushtaq, A., Naamati, G., Naithani, S., Olson, A., Parker, A., Paulini, M., Pedro, H., Perry, E., Preece, J., Quinton-Tulloch, M., Rodgers, F., Rosello, M., Ruffier, M., Seager, J., Sitnik, V., Szpak, M., Tate, J., Tello-Ruiz, M.K., Trevanion, S.J., Urban, M., Ware, D., Wei, S., Williams, G., Winterbottom, A., Zarowiecki, M., Finn, R.D., and Flicek, P. (2022). Ensembl Genomes 2022: an expanding genome resource for non-vertebrates. *Nucleic Acids Research* 50, D996-D1003.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Zaslavskaja, L.A., Lippmeier, J.C., Kroth, P.G., Grossman, A.R., and Apt, K.E. (2000). Transformation of the diatom *Phaeodactylum tricoratum* (Bacillariophyceae) with a variety of selectable marker and reporter genes. *J Phycol* 36, 379-386.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Zhang, C., and Hu, H. (2014). High-efficiency nuclear transformation of the diatom *Phaeodactylum tricoratum* by electroporation. *Mar Genomics* 16, 63-66.

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

ACCEPTED MANUSCRIPT