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.

RESEARCH ARTICLE

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

Authors:

Nolwenn Guéguen[1,$], Yannick Sérès[1,$], Félix Cicéron[1,¶], Valérie Gros[1,#], Grégory Si Larbi[1], Denis Falconet[1], Etienne Deragon[1], Siraba D. Gueye[1], Damien Le Moigne[1], Marion Schilling[1], Mathilde Cussac[1], Dimitris Petroutsos[1,2], Hanhua Hu[3], Yangmin Gong[4,5], Morgane Michaud[1], Juliette Jouhet[1], Juliette Salvaing[1], Alberto Amato[1,*], Eric Maréchal[1,*].

Affiliations:

[1]Laboratoire de Physiologie Cellulaire et Végétale, Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement, Centre National de la Recherche Scientifique, Commissariat à l'Energie Atomique et aux Energies Alternatives, Université Grenoble Alpes; IRIG, CEA-Grenoble, 17 rue des Martyrs; 38000 Grenoble, France

[2]Department of Organismal Biology, Uppsala University, 75236, Uppsala, Sweden

[3]Key Laboratory of Algal Biology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, China

[4]Oil Crops Research Institute of the Chinese Academy of Agricultural Sciences, Wuhan 430062, China

[5]Key Laboratory of Biology and Genetic Improvement of Oil Crops, Ministry of Agriculture, Oil Crops Research Institute of Chinese Academy of Agricultural Sciences, Wuhan 430062, China

[¶]Present address: Living Systems Institute - Exeter University Stocker Road; EX4 4QD, Exeter, UK

[#]Present address: Institut des Biomolécules Max Mousseron; Pôle Chimie Balard Recherche, 1919 route de Mende, 34293 Montpellier

[$]Equal contribution

***Corresponding authors:**

Alberto Amato (alberto.amato@cea.fr) and Eric Maréchal (eric.marechal@cea.fr)

Short title: Multigenic MGD family in diatoms

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

Abstract.

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

Introduction.

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

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

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

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

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

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

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

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

UDP-Gal + DAG —(MGD)→ MGDG + UDP

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

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

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

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

Results.

Gene structure and molecular phylogeny of *P. tricornutum MGDα*, *MGDβ* and *MGDγ*

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

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

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

MGDα, MGDβ and MGDγ predicted targeting sequences and protein structures

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

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

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

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

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

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

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

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

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

***Phaeodactylum tricornutum* MGDα, MGDβ and MGDγ subcellular localization.**

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

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

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

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

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

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

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

Generation of MGDα, MGDβ and MGDγ edited lines.

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

For *MGDα*, we selected three pure KO lines with frameshift disruptions *i.e.* one KO obtained with sgRNA 125i (*mgdαi1*), and two independent KO lines obtained with sgRNA 125j (*mgdαj1* and *mgdαj3*) (Supplemental Figure 9). Two *mgdα* mutants conserved for at least one year highlighted homozygous profiles and were stable after repeated rounds of cultivation in liquid medium (Supplemental Figure 10).

For *MGDβ*, we selected one KO line with sgRNA 168b (*mgdβb1*), and three mutants obtained with sgRNA 168c (*mgdβc1*, *mgdβc2* and *mgdβc3*) (Supplemental Figure 9). The three lines *mgdβc1*, *mgdβc2* and *mgdβc3* derive from the same initially transformed cell. They harbor the same vector insertion, but different editions at the targeted locus. The *mgdβc2* mutant has a deletion of 6 nucleotides (DEL6), which does not generate a frameshift in the sequence (Supplemental Figure 9). This mutant lacks two amino acids between the predicted signal peptide and the first N-terminal conserved region. This deletion was not expected to impair the enzyme function and thus *mgdβc2* was kept as a possible control for *mgdβc1* and

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

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

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

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

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

monitored at low light intensity (20 µmol photons.m^{-2}.s^{-1}) (Ruban et al., 2004; Goss and Jakob, 2010). Under all tested conditions, no significant changes in photosynthetic efficiency or non-photochemical quenching (NPQ) were observed, except in the case of mgdgγi1. This line showed lower photosynthetic efficiency under 55 and 335 µmol photons m-2 s-1 and reduced recovery under low light intensity following high light intensity measurements. Additionally, mgdgγi1 exhibited slower relaxation of NPQ under the same low light conditions (Supplemental Figure 12C and F).

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

Cell morphology of MGD KO and overexpressing lines was examined by light and epifluorescence microscopy (Supplemental Figure 13). All mutant cells were fusiform with a normally sized plastid compared to the WT. No change in cell size was detected. Potential impact of MGD KO on plastid membranes was checked by STEM (Figure 4), with cells from *mgdαj3*, *mgdβc3* and *mgdγh2* strains. No modifications could be observed in plastids nor in any other parts of the cell in any MGD KO line. In MGDβ null mutants, blob-like structures were still present (Figure 4). Likewise, no change in cell ultrastructure could be noticed in overexpressing lines (Supplemental Figure 14).

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

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

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

in the WT) (Supplemental Figure 15B). In MGDβ KO strains, the proportion of PC was higher (14.10%) compared to both WT and silent *mgdβc*2 controls (12.7%) (Supplemental Figure 15B), consistently with an opposite trend in the MGDβ overexpressing strains (Supplemental Figure 16B). All MGDγ KO strains had slightly lower proportions of PG (5.1% *vs* 6.3% in the WT) and slightly higher proportions of DAG (1.23% *vs* 1.05% in the WT) (Supplemental Figure 15B). DAG tended to decrease in MGDγ overexpressing strains (Supplemental Figure 16B), in line with a more active utilization of this MGD substrate.

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

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

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

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

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

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

For the analysis of MGDβ KO lines, we used the silent *mgdβc2* strain as a negative control in addition to the WT. MGDG molecular species containing a 20:5 in position *sn-1* decreased (Figure 6), with the exception of 20:5-16:0 and 20:5-16:4. MGDG 16:1-16:0 and MGDG 16:1-16:1 decreased, while other MGDG species containing a C16 FA in positions *sn-1* and *sn-2* and two or more unsaturations on a C16 increased. The opposite trend was observed in the MGDβ overexpressing strains (Supplemental Figure 17). By comparison, DGDG molecular species were little affected (Figure 6). DGDG 14:0-16:1 tended to decrease and 16:1-16:1 strongly decreased, in line with similar changes in MGDG. The DGDG species containing 20:5 were almost not affected, with only a slight increase in 20:5-16:0 compared to the controls. By contrast with MGDα mutants, only a few changes were observed in other glycerolipids in MGDβ KO lines: for PC, DGTA and PE, 16:1-16:2, 16:1-16:3, and 20:5-16:2 species tended to increase in the mutants, while 20:5-18:1, 20:5-18:2 and 20:5-18:3 tended to decrease (Supplemental Figure 20). PE 20:5-20:5 species increased in MGDβ KO, while it decreased consistently in the MGDβ overexpressing strains (Supplemental Figure 18).

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

compared to the WT. For several molecular species, the *mgdγh1* and *mgdγh2* mutants behaved slightly differently compared to *mgdγi1* (Figure 7). Nevertheless, the common increase in MGDG 16:1-16:3, 16:2-16:3 and 16:3-16:3 species was coherent with an opposite trend in the overexpressors (Supplemental Figure 17). No change was observed in 20:5-containing MGDG species, except a decrease in 20:5-16:4 mirrored by an increase in the overexpressing lines. In DGDG and SQDG, 20:5-containing species tended to increase except for DGDG 20:5-16:4. PG molecular species distribution was globally unaffected by MGDγ mutations.

In other glycerolipid classes linked with the endomembrane system, many changes in species proportions were observed in MGDγ mutants (Supplemental Figure 21). In PC and 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 DGTA 16:1-18:3 and PC 16:1-18:4 decreased. There was a striking decrease in 20:5-20:4 and 20:5-20:5 in PC, DGTA and PE. By comparison with MGDα and MGDβ mutations, 14:0-containing species were not impacted in MGDγ mutants. In DAG, changes were limited to 16:0-16:1 and 20:5-16:0, which tended to increase and decrease, respectively.

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

We performed RT-qPCR analyses on MGD KO strains to check for possible compensations at the transcriptional level. In replete conditions, the WT, *MGDβ* and *MGDγ* were expressed at similar levels, while *MGDα* was ~4 times more expressed (Supplemental Figure 22). We observed a decrease in *MGDβ* gene expression in MGDβ KO lines (2.25-fold decrease on average), which was not observed in the silent mutant *mgdβc2*. In MGDγ mutants, an increase of *MGDα* and *MGDβ* expression (1.37- and 1.69-fold increase, respectively), suggests the possible activation of a compensation mechanism at the transcriptional level, when *MGDγ* is affected. No apparent transcriptional compensation could be detected when *MGDα* or *MGDβ* were affected (Supplemental Figure 22). This does not exclude some other compensation mechanisms at the post-transcriptional level.

Impact of *MGDα*, *MGDβ* and *MGDγ* mutations on the response of *P. tricornutum* to a nitrogen starvation

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

527 respectively. By contrast, MGDβ 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: *mgdαi1*, *mgdαj3*, *mgdβb1*, *mgdβc1*, *mgdγh2*, and *mgdγi1*. 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 ~9×10^6 cells.mL^{-1} at day 5, which is about
536 half of the concentration that would be expected in 10N10P condition. The growth of MGDα
537 and MGDβ KO lines were similar to that of the WT, whereas growth was affected in both
538 MGDγ 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 MGDβ 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 *mgdγi1* 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 ~60% of total lipid
555 content (Supplemental Figure 26B).

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

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

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

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

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

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

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 MGDγ 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 MGDγ overexpressing lines (Supplemental Figures 18 and 27).

597 Considering that MGDα is naturally absent under nitrogen shortage, the knockout of
598 either MGDβ or MGDγ could be seen as a double invalidation (*mgdα*/*mgdβ* or *mgdα*/*mgdγ*,
599 respectively), resulting in stronger phenotypes compared to those observed in replete medium.
600 MGDβ looked important to keep a high level of polyunsaturated galactolipids in the absence
601 of MGDα. MGDγ seems to play a specific role in response to nitrogen shortage, in particular
602 on galactolipids with a low unsaturation level. MGDγ KO had a stronger effect on non-
603 plastidial lipids than MGDβ KO. The increase in DGTA 16:0-16:1 and 16:1-16:1 in MGDγ
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 MGDγ products and TAG synthesis.

607 **MGDβ/MGDγ double mutations**

608 To better understand the role of MGDα, three MGDβ/MGDγ 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 *mgdβ*/*mgdγ* KO proved non-lethal. The double KO mutants showed no variation in the
612 MGDG levels compared to the WT, demonstrating that MGDα 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 MGDβ and MGDγ 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 MGDα. 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 MGDα 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

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

Discussion.

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

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

P. tricornutum MGDα, MGDβ and MGDγ are active MGDG synthases and have distinct substrate specificities.

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

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

P. tricornutum MGDα, MGDβ and MGDγ localize to distinct subcellular compartments inside and outside of the plastid

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

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

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

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

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

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

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

731 the MGDβ-eGFP fluorescence signal was more diffuse, either suggesting that the VN itself
732 loses the blob structure and become more scattered, or that MGDβ 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 MGDβ 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 MGDβ 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 MGDα and MGDβ KO lines. By contrast, in MGDγ mutants,
751 an increase in *MGDα* and *MGDβ* 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 MGDα in thylakoid membrane synthesis**

757 Expression analysis in *P. tricornutum* WT strain showed that *MGDα* was the most
758 expressed isoform under our optimal laboratory conditions, with relative expressions ~4 times
759 higher than that of *MGDβ* and *MGDγ*. In a recent study, *MGDγ* 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 *MGDα* expression. MGDα basal role under optimal conditions is consistent with

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

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

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

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

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

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

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

Under optimal conditions, MGDβ mutation strongly decreased the accumulation of 20:5-containing MGDG species, except MGDG 20:5-16:4. Interestingly, MGDG 14:0-16:1, 16:1-16:0 and 16:1-16:1 species also decreased. This firstly suggests that under normal conditions, MGDβ mostly uses 20:5-16:0 and 20:5-16:1 DAG, and that the corresponding MGDG products are rapidly desaturated, leading to the accumulation of 20:5-containing MGDG species with a higher unsaturation level of the C16 at position *sn-2* (Figure 9 and Table 1). Given the localization of this isoform in the blob/PPM, the 20:5-containing MGDG it produces need to be rapidly transported to the thylakoid membranes where C16-desaturases are located. The role of MGDβ in the production of 20:5-containing MGDG species is also consistent with its subcellular localization, since 20:5 FA is produced outside or at the periphery of the plastid following elongation and desaturation steps (Figure 9). For the massive incorporation of this FA into the plastid, 20:5 and/or 20:5-containing precursor need to cross the four limiting membranes, via the PPM where MGDβ 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 MGDβ, but that most of them are not directly
829 available for desaturation. The heterologous expression of MGDβ 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 MGDβ overexpressing
833 lines, while the MGDG 16:2-16:3 and 16:3-16:3 decreased. By this mean, MGDβ can produce
834 more saturated forms of MGDG, remaining away from stromal C16-desaturases, thus
835 compensating for the loss of MGDγ.

836 Thus, on the one hand, MGDβ 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, MGDβ 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 MGDβ 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 MGDβ KO
844 lines, while DGDG 14:0-16:1 tended to decrease. This would be consistent with the role of
845 MGDβ 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 MGDα and MGDγ. 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 MGDα-associated DGD pointed above.

852 Finally, under nitrogen deprivation, at low MGDα level and in the absence of MGDβ,
853 MGDγ could compensate for the lack of MGDG, and for the provision of substrates for
854 DGDG species.

855 **MGDγ 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, MGDγ 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.

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

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

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

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

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

MGD isoforms and extraplastidial lipid homeostasis

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

MGDγ mutation had a stronger impact on non-plastidial lipid composition. In particular, 16:1 and 16:0 containing PC and DGTA increased both under optimal and nitrogen deprived conditions. MGDγ being located at the periphery and outside of the plastid, it is possible that it sits near to PC and DGTA synthesis sites. Therefore, it is coherent that the DAG species not used by MGDγ to produce MGDG are used as substrate for the synthesis of PC and DGTA (Figure 9). Following MGDγ KO, 16:0 amounts available for elongation into 18:0 by Δ0-elongases (Dolch et al., 2017a), the committing step for further desaturations up to 18:4 and 20:5 in phospholipids and DGTA increased. The amounts of PC, PE and DGTA containing an 18:1 or an 18:2 increased. It seems however that the rates of 18:2-to-18:3 and 18:3-to-18:4 desaturations and 18:4-to-20:4 elongation were slower than the 16:0 export rate.

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

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

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

In both cases, the emergence of 'novel' MGD isoforms (type B MGDs in angiosperm and MGDβ/γ in diatoms) strengthens the integration of the organelle within the cell and 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 (Petroutsos 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 MGDγ and to a lesser extent, MGDβ , 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 MGDα level is drastically reduced, MGDγ compensates for the loss
976 of MGDβ 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 (MGDβ) supposed to derive from the red alga endosymbiont plasma membrane, and the EpM
983 and ER (MGDγ). 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

Materials and Methods.

Primary sequence analyses.

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

Protein structure predictions.

The software Phyre^2 (Protein Homology/analogY Recognition Engine V 2.0) (Kelley et al., 2015), and AlphaFold (Jumper et al., 2021; Jumper and Hassabis, 2022) run from ChimeraX (Mirdita et al., 2022) were used to model *P. tricornutum* MGD sequences, which were then viewed on PyMol (Janson et al., 2017). AtMGD1 structure (4X1T- wwPDB consortium, 2019) (Rocha et al., 2016) was used to build Phyre^2 PtMGDs models. For each *P. tricornutum* MGD, the peptide sequences used as input were those deduced after cDNA sequencing and signal peptide prediction, as represented in Figure 3E. These correspond to the following sequences: for *MGDα*, the sequence retrieved from Phatr3_J14125.t1 starting from the second methionine (MCKL to TRTS, 460 amino acid); for *MGDβ,* the whole sequence from Phatr3_J54168.t1 (MVWS to LNNK, 559 amino acid); for *MGDγ*, the whole sequence from Phatr3_J9619.t1 (MATG to EPSR, 618 amino acids).

MGD phylogenetic analysis.

All MGD protein sequences used for phylogenetic analysis were retrieved from the National Centre for Biotechnology Information (NCBI) (Supplemental Table 1) after a protein BLAST search using MGDα (XM_002181649.1), MGDβ (XM_002186319) and MGDγ

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

Cultivation of *Phaeodactylum tricornutum* cells

Phaeodactylum tricornutum (Pt1) Strain 8.6 CCAP 1055/1 (CCMP2561- Culture Collection of Marine Phytoplankton, now known as NCMA: National Center for Marine Algae and Microbiota) was used to generate overexpressing and KO lines. Cells were maintained and grown in a modified ESAW (Enriched Seawater, Artificial Water) medium (NaCl 362.7 mM; Na_2SO_4 25 mM; KCl 8.03 mM; $NaHCO_3$ 2.067 mM; KBr 0.725 mM; H_3BO_3 0.372 mM; NaF 0.0657 mM; $MgCl_2$ 47.18 mM; $CaCl_2$ 9.134 mM; $SrCl_2$ 0.082 mM; Na_2-glycerophosphate 21.8 µM; Na_2SiO_3 105.6 µM; disodium ethylenediaminetetraacetate dehydrate (Na_2EDTA) 14.86 µM; $Fe(NH_4)_2(SO_4)_2$ 5.97 µM; $FeCl_3$ 0.592 µM; $MnSO_4$ 2.42 µM; $ZnSO_4$ 0.254 µM; $CoSO_4$ 0.0569 µM; Na_2MoO_4 0.52 µM; H_3BO_3 61.46 µM; Na_2SeO_3 10 nM; biotin (vitamin H) 8.18 nM; cobalamin (vitamin B_{12}) 2.94 nM; thiamine (vitamin B_1) 0.594 µM) (Falciatore et al., 2000) using ten times enriched nitrogen and phosphate sources

(10N10P containing 5.49 mM $NaNO_3$ and 0.224 mM NaH_3PO_4) (Abida et al., 2015). Culture in liquid media was performed at 20°C, on a 12:12 light (40 µmol photons.m^{-2}.sec^{-1}) dark cycle, under gentle agitation, in an INFORS HT Multitron Pro incubator. Cell concentration was kept between 0.5×10^6 and 10×10^6 cell.mL^{-1}. For nitrogen starvation, *P. tricornutum* cells were cultivated in triplicates in 50 mL of ESAW 10N10P in 250-mL flasks until a concentration of 3-4×10^6 cell.mL^{-1}. After centrifugation at 1,500 g for 10 minutes and removal of supernatant, cells were washed with 50 mL of ESAW 0N10P medium, and transferred into to 250-mL flaks containing 50 mL of ESAW 0N10P, at a concentration of 2.5-3×10^6 cell·mL^{-1}. Cultures on solid medium were performed with ESAW 10N10P complemented with Agar (1% m/v), at 20°C under continuous light (50 µmol photons.m^{-2}.s^{-1}), in a MLR-352-PE Climate Chamber. To guarantee axenic conditions, all media were supplemented with carbenicillin disodium salt (0.237 µM). For culture of mutant lines transformed with CRISPR-Cas9 vectors and of overexpressing lines, zeocin (Invitrogen; 0.07 µM) was added to the media. For each analysis, three independent lines were cultivated. Growth was evaluated by cell counting using a TECAN infinite M1000Pro plate reader and determined following the equation $y = 0.01834\ x + 0.03758$ with y the absorbance at 730 nm and x the number of cells in million per milliliter (Conte et al., 2018).

DNA and RNA extractions.

Genomic DNA extraction from *P. tricornutum* was performed on a cell pellet corresponding to 100×10^6 cells, suspended in 500 µL of lysis buffer (250 mM Tris pH 8.2, 100 mM EDTA, 2% SDS v/w, 350 mM NaCl) and transferred to a 1.5-mL tube. The sample was incubated 15 seconds at 60°C under agitation (450 rpm) using a thermobloc. 500 µL of cold phenol:chloroform:isoamyl alcohol (25:24:1) were added to the sample and homogenized gently. The sample was centrifuged at 13,000×g at 4 °C for 10 minutes. 300 µL of the upper aqueous phase was collected. A second extraction was performed by addition of one volume of chloroform:isoamyl alcohol (24:1), and 200 µL of the upper aqueous phase was collected. DNA was precipitated with 30 µL of sodium acetate 3 M pH 5 and three volumes of ethanol. After a 20-minute incubation at -20°C and a 10-minute centrifugation at 13,000×g at 4°C the precipitated DNA was washed with 700 µL of ethanol 70% v/v, centrifuged at 13,000×g at 4°C for 5 minutes and air dried at room temperature. DNA was resuspended in DNAse-free water. Concentration and purity were evaluated using a NanoDrop2000 (ThermoFisher Scientific). For RNA extraction from *P. tricornutum* cells, 1.5 mL of TRI Reagent (Sigma-Aldrich) were added to a frozen cell pellet corresponding to

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

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

This method is detailed in supplemental material.

Heterologous expression of MGD isoforms in yeast

This method is detailed in supplemental material.

Construction of MGD overexpressing lines.

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

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

Construction of CRISPR-Cas9 single and double knockout lines.

Single guide RNAs (sgRNA) were designed using the PhytoCrispex online tool (Rastogi et al., 2016), choosing NGG as Protospacer Adjacent Motif (PAM) sequence, and allowing CRISPR targets to start with any nucleotide. sgRNA were selected based on their proximity to START codons and active site residues. sgRNA forward and reverse sequences preceded with the short nucleotide sequences TCGA and AAAC respectively are shown in Supplemental Table 3. The sgRNA sequences were inserted in the pKSdiaCas9_sgRNA-zeo vector, kindly provided by Cécile Giustini (LPCV, Grenoble), derived from pKS diaCas9_sgRNA vector (Addgene). *P. tricornutum* cells were transformed with corresponding vectors by particle bombardment as described previously (Falciatore et al., 1999; Dolch et al., 2017b). For MGDβ/MGDγ double KO, *mgdβc1* single KO was chosen as genetic background. The sgRNA 619h (Supplemental Table 3) was cloned between the two BsaI sites of the pYS31 episome, derived from PtPuc3_diaCas9_sgRNA episome (Addgene) with a blasticidin resistance cassette instead of a zeocin. The PtPuc3_diaCas9_sgRNA was digested by SalI and StuI in order to excise the zeocin resistance gene. The Blasticidin S deaminase (BSD) coding sequence, conferring the resistance to blasticidin, was amplified using the primers olYS40 (5'-GAAAAATTAACCAAGATGGCCAAGCCTTTGTCTCA-3') and olYS41 (5'CCCAGATCTCCGAGGTCAGCCCTCCCACACATAAC-3'). 50 ng of the vector were used for Gibson assembly with a 2-fold molar excess of the BSD fragment. The episome was transformed by bacterial conjugation (Karas et al., 2015). Colony PCR (Phire Plant Direct PCR Master Mix, F160, ThermoFisher Scientific) were conducted to analyse the

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

Bacterial conjugation

This method is detailed in supplemental material.

Glycerolipid analyses.

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

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

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

Epifluorescence microscopy.

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

Laser scanning confocal microscopy.

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

Transmission electron microscopy.

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

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

Accession numbers.

Species and associated MGD protein sequences used for phylogeny (NCBI accession numbers): *Vitrella brassicaformis* (CEM14063.1, CEM34604.1); *Gregarina niphandrodes* (XP_011130274.5); *Pseudo-nitzschia multistriata* (VEU37191.1, VEU39505.1); *Fragilariopsis cylindrus* (OEU16730.1, OEU06991.1); *Fistulifera solaris* (GAX23913.1, GAX10983.1, GAX09439.1); *Phaeodactylum tricornutum* (XP_002186355, XP_002176800, XP_002181685), *Thalassiosira pseudonana* (XP_002295865, XP_002293576.1, XP_002294242); *Ectocarpus siliculosus* (CBJ28381.1, CBJ28372.1, CBN79326.1); *Aureococcus anophagefferens* (XP_009033523.1, XP_009038839.1, XP_009035780.1); Microchloropsis/Nannochloropsis gaditana (XP_005855249.1); *Coccomyxa subellipsoidea* (XP_005651388.1); *Chlorella sorokiniana* (PRW45643.1); *Ostreococcus tauri* (OUS42062.1); *Chlamydomonas reinhardtii* (PNW74102.1); *Monoraphidium neglectum* (XP_013903204.1); *Raphidocelis subcapitata* (GBF88428.1); *Arabidopsis thaliana* (NP_194906.1, NP_565352.1, NP_568394.2); *Coffea arabica* (XP_027069738.1, XP_027084801.1); *Spinacia oleracea* (XP_021867203.1, XP_021852153.1); *Nymphaea colorata* (XP_031495579.1, XP_031480803.1); *Amborella trichopoda* (XP_006852865.1, XP_006845407.1); *Brachypodium distachyon* (XP_010238179.1, XP_003573675.2, XP_003570331.1); *Oryza sativa* (XP_015611851.1, XP_015649135.1, XP_015627712.1); *Physcomitrium patens* (XP_024403181.1, XP_024404076.1); *Selaginella moellendorffii* (XP_024533920.1); *Cyanidioschyzon merolae* (XP_005536420.1); *Cyanidiococcus yangmingshanensis* (KAF6002292.1); *Gracilariopsis chorda* (PXF49423.1, PXF42956.1, PXF46603.1); *Blastochloris viridis* (WP_055036643.1). *Phaeodactylum tricornutum* MGD gene accession numbers: *MGDα* (PHATRDRAFT_14125, NCBI; Phatr3_J14125, Ensembl); *MGDβ* (PHATR_43938, NCBI; Phatr3_J54168, Ensembl); *MGDγ* (PHATRDRAFT_9619, NCBI; Phatr3_J9619, Ensembl).

Supplementary Data

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

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

Supplementary Figure S3. Multiple sequence alignment of *Arabidopsis thaliana* AtMGD1, AtMGD2 and AtMGD3 and *Phaeodactylum tricornutum* MGDα, MGDβ and MGDγ proteins

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

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

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

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

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

Supplementary Figure S9. Selected knock-out MGDα, MGDβ and MGDγ mutants, generated by CRISPR-Cas9 editing.

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

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

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

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

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

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

Supplementary Figure S16. Quantitative analysis of fatty acid and glycerolipid content in 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 MGDα mutations on the molecular species constituting
1313 PC, PE, DGTA and DAG in *P. tricornutum*.

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

1316 Supplementary Figure S21. Impact of MGDγ 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 *MGDβ*/*MGDγ* double knock-out mutants, generated by
1330 CRISPR-Cas9 editing.

1331 Supplementary Figure S29. Growth curves of *MGDβ*/*MGDγ* double mutants compared to the
1332 WT.

1333 Supplementary Figure S30. Quantitative analysis of glycerolipid content in *MGDβ*/*MGDγ*
1334 double mutant overexpressing lines

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

Supplementary Figure S32. Impact of MGDβ/MGDγ double mutations on PC, DGTA, PE and TAG molecular species.

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

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

Supplementary Table S3. Single guide RNA list.

Supplementary Table S4. Primers list.

Supplementary Methods

Funding

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., D.P., G.S.L., Y.S. and E.M were supported by the French National Research Agency (GRAL Labex ANR-10-LABEX-04, EUR CBS ANR-17-EURE-0003, ANR AlpAlga ANR-20-CE02-0020, ANR DIM ANR-21-CE02-0021, PEPR Algadvance A-22-PEBB-0002) and IDEX Université Grenoble-Alpes (Glyco@Alps Cross-Disciplinary Program; Grant ANR-15-IDEX-02). H.H. and Y.G. were supported by the National Natural Science Foundation of China (41976119, 91751117). A.A., E.M., H.H. and Y.G. were supported by a CEA-CAS bilateral program. Imaging was performed at the μLife Platform of IRIG. Lipid analyzes were performed at the LIPANG platform supported by the Rhône-Alpes Region, the FEDER funds, and ANR (GRAL Labex ANR-10-LABEX-04, EUR CBS ANR-17-EURE-0003).

Acknowledgements.

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

Author contributions.

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

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

Tables

Figure Legends

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 and eGFP signal 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 intraplastidial localization was obtained when MGDα-eGFP is expressed. With MGDβ-eGFP, the eGFP signal corresponds to the blob-like structure,

indicating a localization of MGDβ at the level of the periplasitidal compartment, and possibly the PPM. MGDγ-eGFP is detected mainly in the ER, and at lower level in the EpM. Scale bar: 5 µm. 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 MGDβ-eGFP fluorescence. Observation of blob-like structure localization by confocal microscopy (A) and transmission electron microscopy (B). *P. tricornutum* transgenic lines overexpressing MGDβ-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.

F**igure 4: *P. tricornutum* cell ultrastructure in MGD KO lines determined by scanning transmission electron microscopy** A WT and KO *mgdαj3* and *mgdγh2* lines were cultured in parallel. In a separate experiment, a WT and a *mgdβc3* 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, SQDG and PG molecular species. Lipid molecular species in MGDG, DGDG, SQDG and PG were analyzed 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, SQDG 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 ± 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 MGDγ mutations on MGDG, DGDG, SQDG and PG molecular species. Lipid molecular species were analyzed 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 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 ± 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 MGDα, MGDβ and MGDγ 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). MGDα is localized at the level of thylakoids, and functional studies reported here are consistent with a utilization of DAG available inside the plastid (3α), to form MGDG species rapidly desaturated (16:1 → 16:4) by plastid desaturases (4α). Based on obtained phenotype of KO and overexpressing lines, MGDG produced by MGDα is likely used by a specific DGD enzyme (5α) to form DGDG. MGDβ 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 (3β). MGDG produced can be desaturated (16:1 → 16:4) by plastid desaturases (4β). Based on obtained phenotype of KO and overexpressing lines, MGDG produced by MGDβ is likely used by a specific DGD enzyme (5β) to form DGDG. MGDγ 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 (3γ), producing MGDG at a location far from plastid desaturases. Based on obtained phenotype of KO and overexpressing lines, MGDG produced by MGDγ is also likely associated to a specific DGD enzyme (5γ). In nitrogen shortage, MGDα protein level decreases drastically, and MGDG production relies on MGDβ and MGDγ. 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.

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

References.

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

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.

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.

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.

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.

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

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.

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.

Billey, E., Magneschi, L., Leterme, S., Bedhomme, M., Andres-Robin, A., Poulet, L., Michaud, M., Finazzi, G., Dumas, R., Crouzy, S., Laueffer, 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.

Botte, C., Jeanneau, C., Snajdrova, L., Bastien, O., Imberty, 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.

Boudiere, L., Michaud, M., Petroutsos, 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.

Bowler, C., Allen, A.E., Badger, J.H., Grimwood, J., Jabbari, K., Kuo, A., Maheswari, U., Martens, C., Maumus, F., Otillar, 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., Montsant, 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.

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.

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.

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.

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.

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.

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.

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.

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.

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

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

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., Mulot, 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.
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.
Dolch, L.J., and Marechal, E. (2015). Inventory of fatty acid desaturases in the pennate diatom Phaeodactylum tricornutum. Marine drugs 13, 1317-1339.
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.
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.
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.
Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32, 1792-1797.
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.
Falciatore, A., Alcalà, M.R., Croot, P., and Bowler, C. (2000). Perception of Environmental Signals by a Marine Diatom. Science 288, 2363.
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.
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.
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.
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.
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.
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.
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.
Goss, R., and Jakob, T. (2010). Regulation and function of xanthophyll cycle-dependent photoprotection in algae. Photosynth Res 106, 103-122.

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.
Grosche, C., Hempel, F., Bolte, K., Zauner, S., and Maier, U.G. (2014). The periplastidal compartment: a naturally minimized eukaryotic cytoplasm. Curr Opin Microbiol 22, 88-93.
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.
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.
Gschloessl, B., Guermeur, Y., and Cock, J.M. (2008). HECTAR: a method to predict subcellular targeting in heterokonts. BMC Bioinformatics 9, 393.
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.
Guiry, M.D. (2012). How Many Species of Algae Are There? J Phycol 48, 1057-1063.
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.
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.
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.
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.
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.
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.
Jensen, P.E., and Leister, D. (2014). Chloroplast evolution, structure and functions. F1000Prime Rep 6, 40.
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.
Jumper, J., and Hassabis, D. (2022). Protein structure predictions to atomic accuracy with AlphaFold. Nat Methods 19, 11-12.
Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, 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.
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.
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.
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.
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.
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.
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.
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.
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.
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.
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.
Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol 35, 1547-1549.
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.
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.
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.
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.
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.
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.
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.
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.
Lupette, J., Tardif, M., Brugiere, 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.
Lupette, J., Jaussaud, A., Seddiki, K., Morabito, C., Brugiere, S., Schaller, H., Kuntz, M., Putaux, J.L., Jouneau, P.H., Rebellle, 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.
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.
Mann, D.G. (1999). The species concept in diatoms. Phycologia 38, 437-495.
Mann, D.G., and Droop, S.J.M. (1996). Biodiversity, biogeography and conservation of diatoms. . Hydrobiologia, 19–32.
Marechal, E. (2024). How Did Thylakoids Emerge in Cyanobacteria, and How Were the Primary Chloroplast and Chromatophore Acquired? Methods Mol Biol 2776, 3-20.
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.
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-+.
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.
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.
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.
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.
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.

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.

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.

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.

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.

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.

Serôdio, 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.

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.

Siaut, M., Heijde, M., Mangogna, M., Montsant, A., Coesel, S., Allen, A., Manfredonia, A., Falciatore, A., and Bowler, C. (2007). Molecular toolbox for studying diatom biology in Phaeodactylum tricornutum. Gene 406, 23-35.

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 tricornutum. Plant J 106, 1247-1259.

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

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 tricornutum. Protist 166, 506-521.

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.

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., Luypaert, 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.

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

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

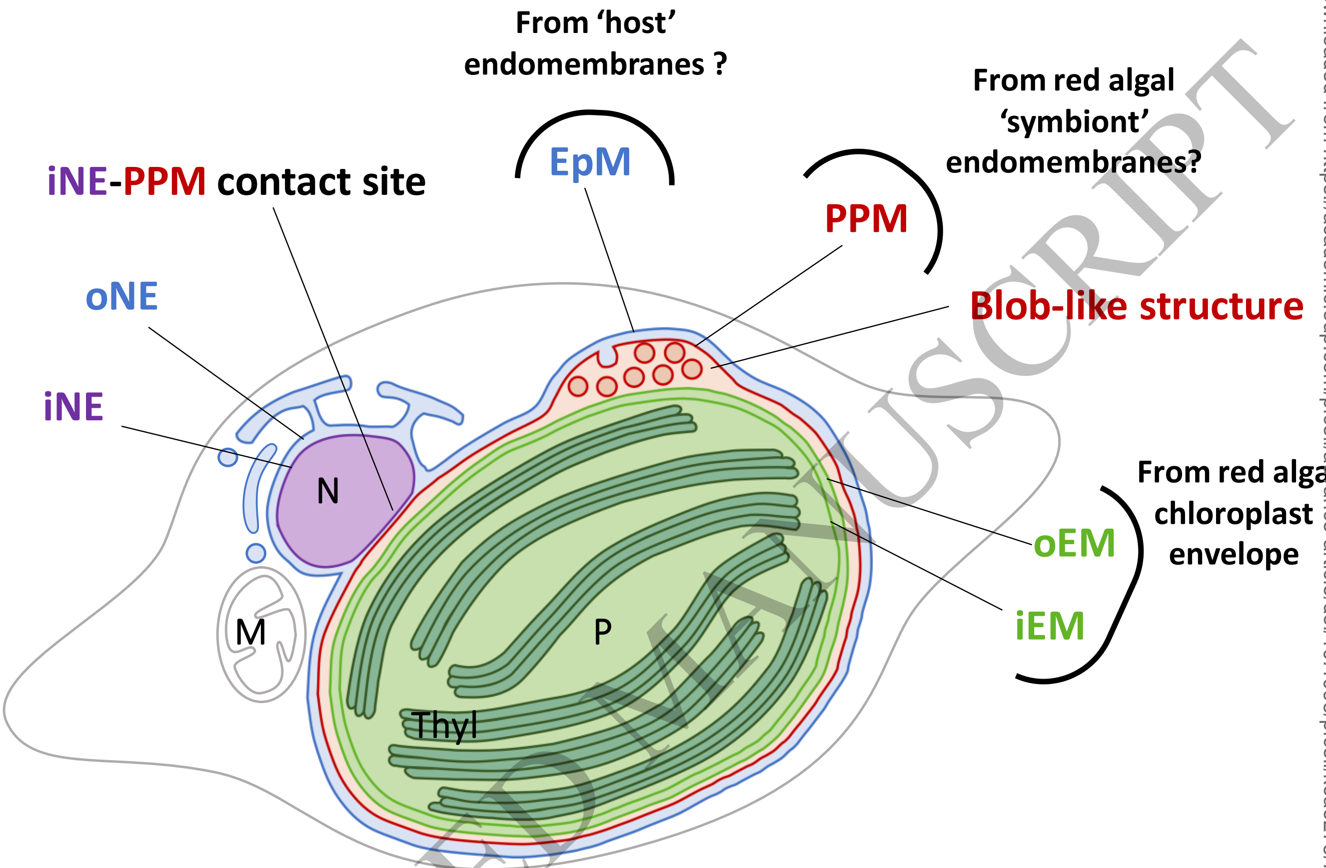

Figure 1: Chimeric origin of the secondary plastid in diatom. The scheme shows a fusiform cell of *Phaeodactylum tricornutum*. The plastid is limited by four membranes. The epiplastidial membrane (EpM), shown in blue, is continuous with the outer nuclear envelope (oNE). The periplastidial membrane (PPM), 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.

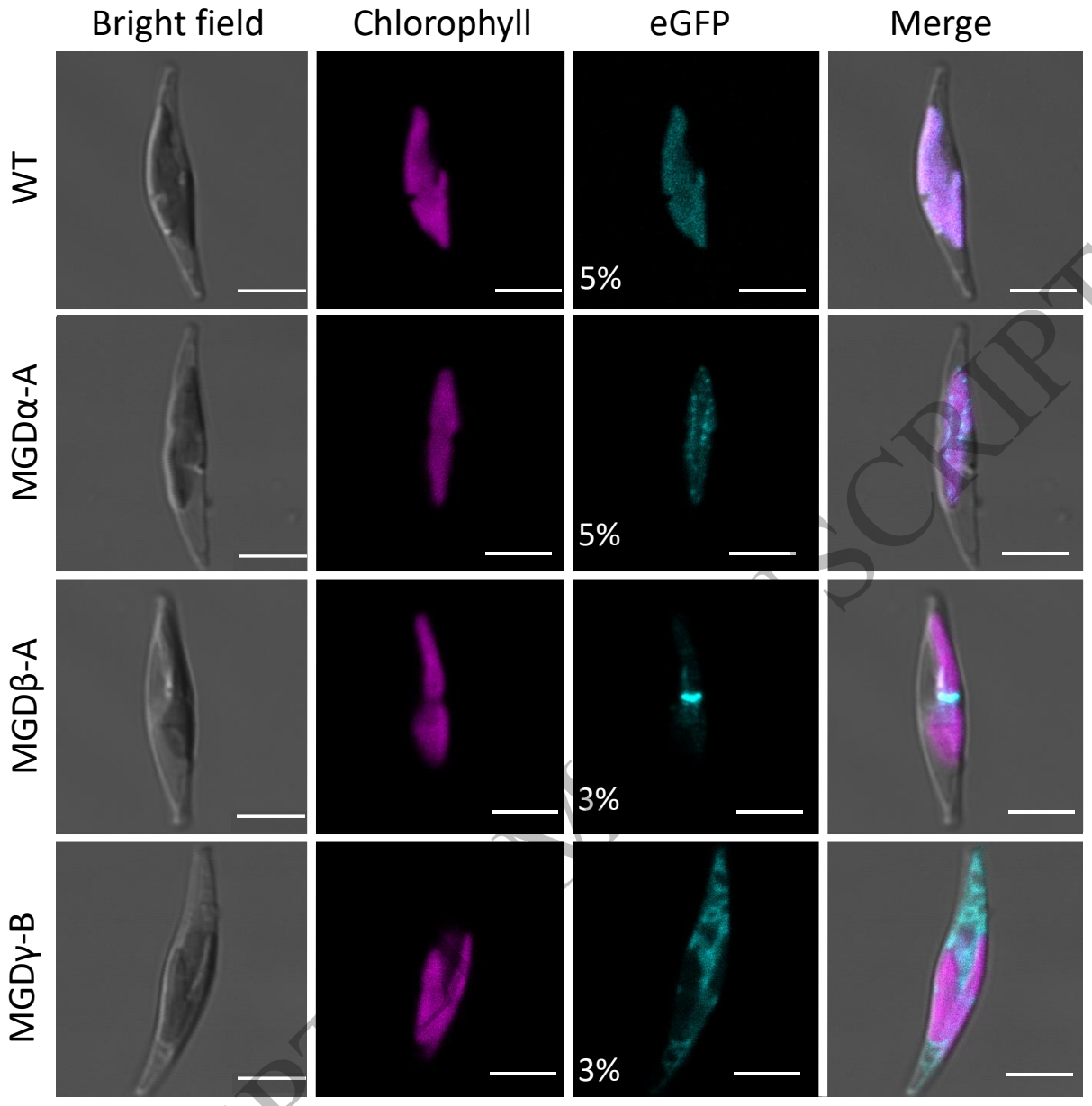

Figure 2: Subcellular localization of MGD isoforms fused to eGFP in *P. tricornutum*. Detection of MGD-eGFP expressed in transgenic *P. tricornutum* lines with wild type as a negative control. For each cell, bright field, chlorophyll fluorescence (magenta) and eGFP signal (light blue) are shown, with 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 intraplastidial localization was obtained when MGDα-eGFP is expressed. With MGDβ-eGFP, the eGFP signal corresponds to the blob-like structure, indicating a localization of MGDβ at the level of the periplasitidal compartment, and possibly the PPM. MGDγ-eGFP is detected mainly in the ER, and at lower level in the EpM. Scale bar: 5 µm. Similar independent observations are provided in Supplemental Figure 7. MGD, monogalactosyldiacylglycerol synthase; WT, wild type.

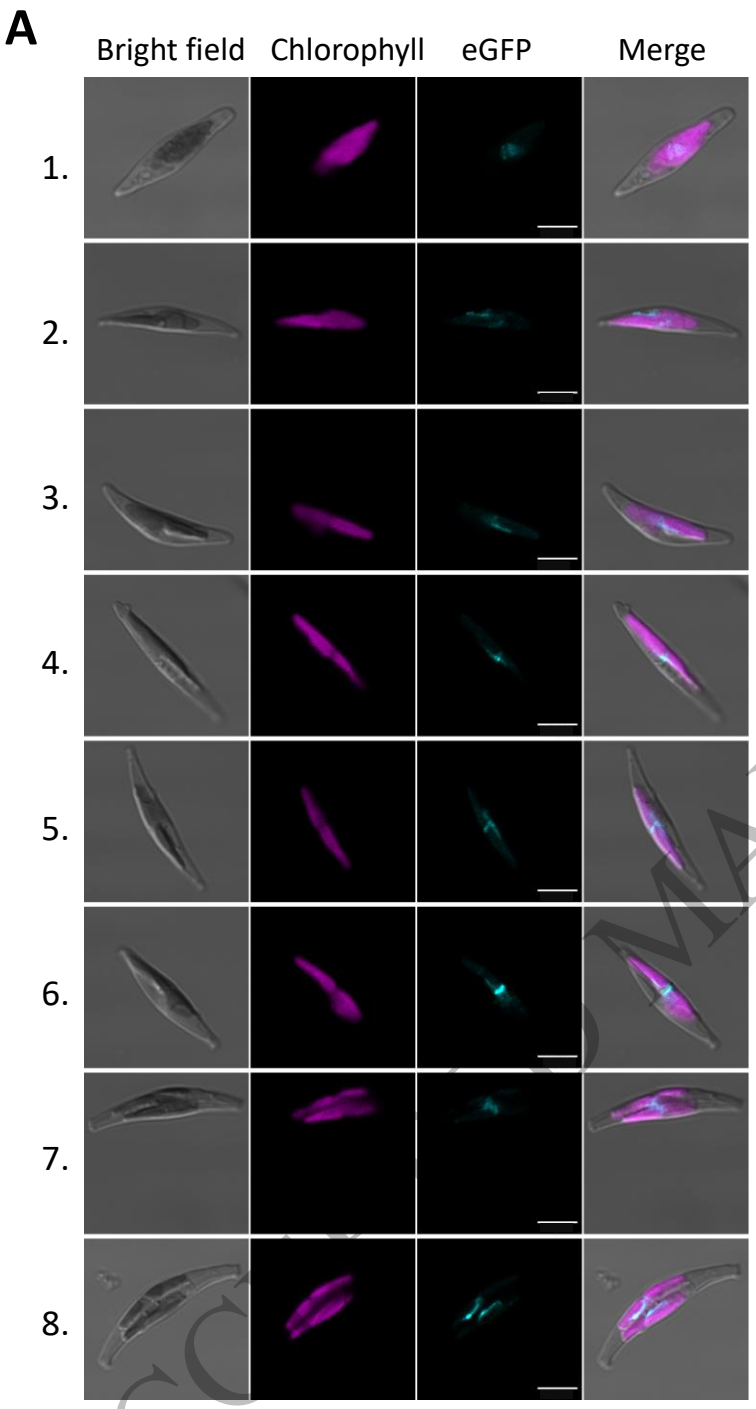

Figure 3: Subcellular localization of the blob-like structure, associated to MGDβ-eGFP fluorescence. Observation of blob-like structure localization by confocal microscopy (A) and transmission electron microscopy (B). *P. tricornutum* transgenic lines overexpressing MGDβ-eGFP were used to monitor the 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.

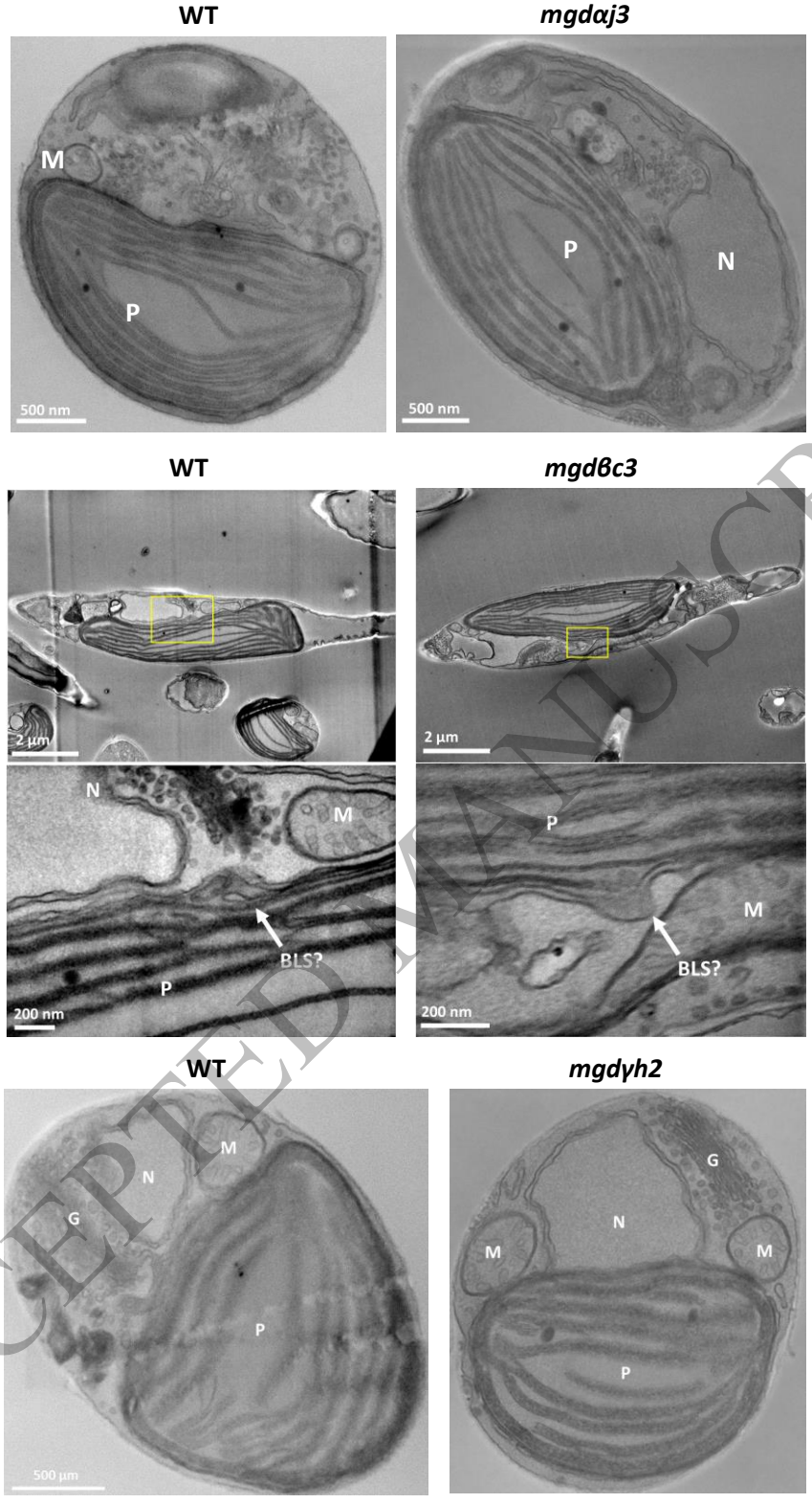

Figure 4: *P. tricornutum* cell ultrastructure in MGD KO lines determined by scanning transmission electron microscopy A WT and KO *mgdαj3* and *mgdγh2* lines were cultured in parallel. In a separate experiment, a WT and a *mgdβc3* mutant were cultured in parallel. Cell ultrastructure is shown in each 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.

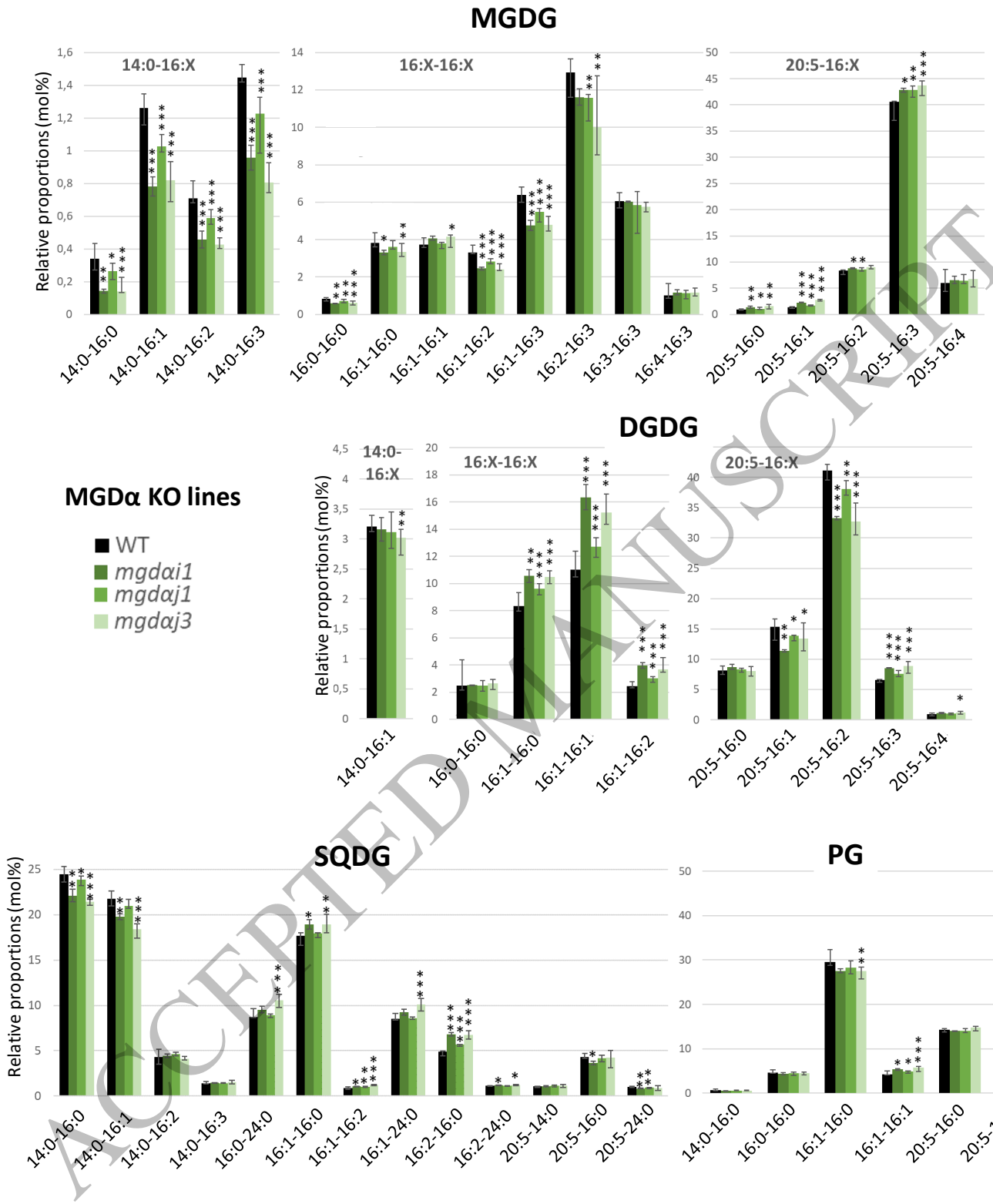

Figure 5: Impact of MGDa mutations on MGDG, DGDG, SQG and PG molecular species. Lipid molecular species in MGDG, DGDG, SQDG and PG were analysed by LC-MS/MS. Each result is the median of six biological replicates ± min and max values. (*), *P-value* < 5.10^{-2}; (**), *P-value* < 1.10^{-2}; (***), *P-value*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

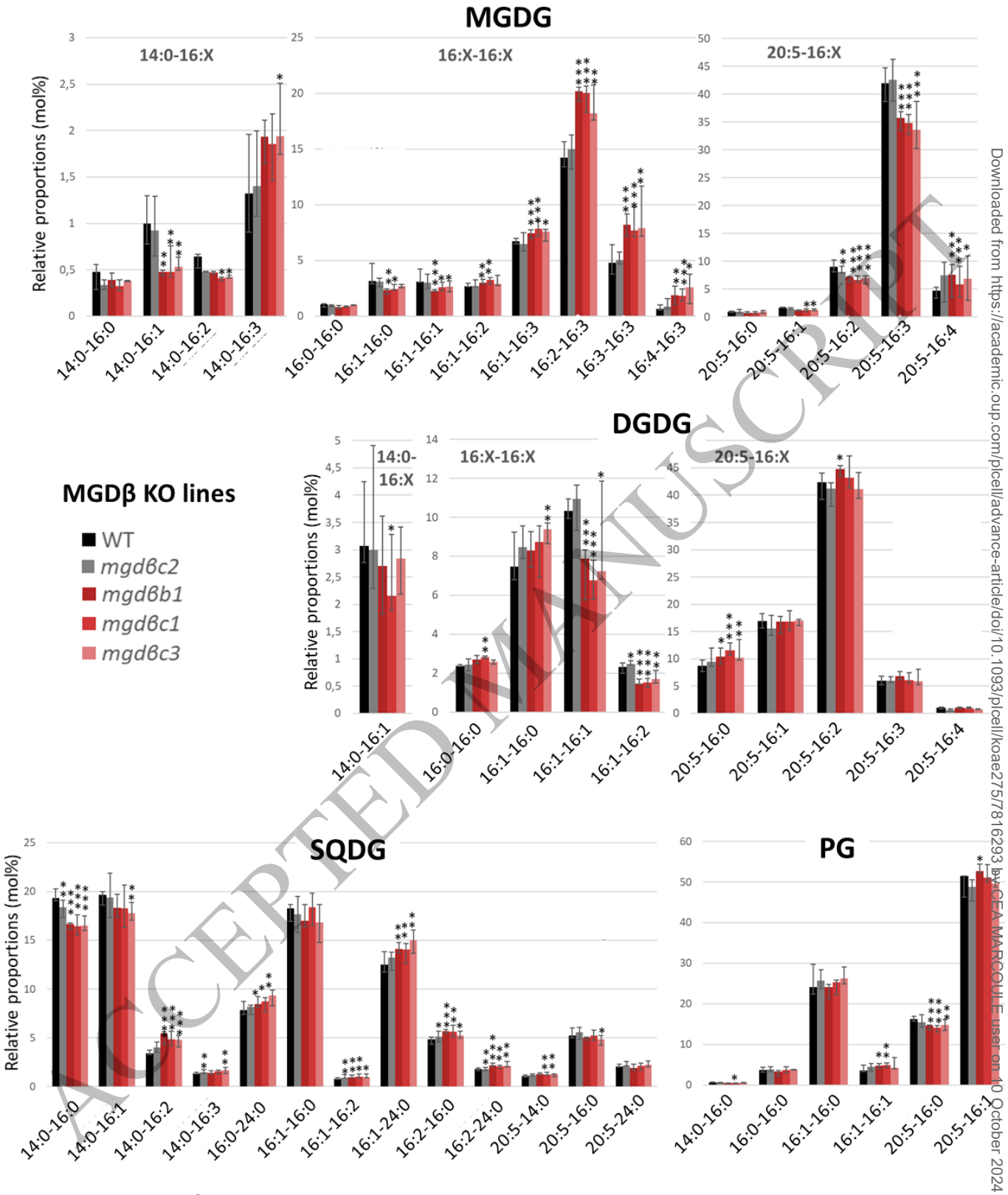

Figure 6: Impact of MGDβ mutations on MGDG, DGDG, SQG and PG molecular species. Lipid molecular species were analyzed by LC-MS/MS. The *mgdβc2* mutant contains a silent mutation and was used as a control. Each result is the median of six biological replicates ± min and max values. (*), *P-value* < 5.10^{-2}; (**), *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.

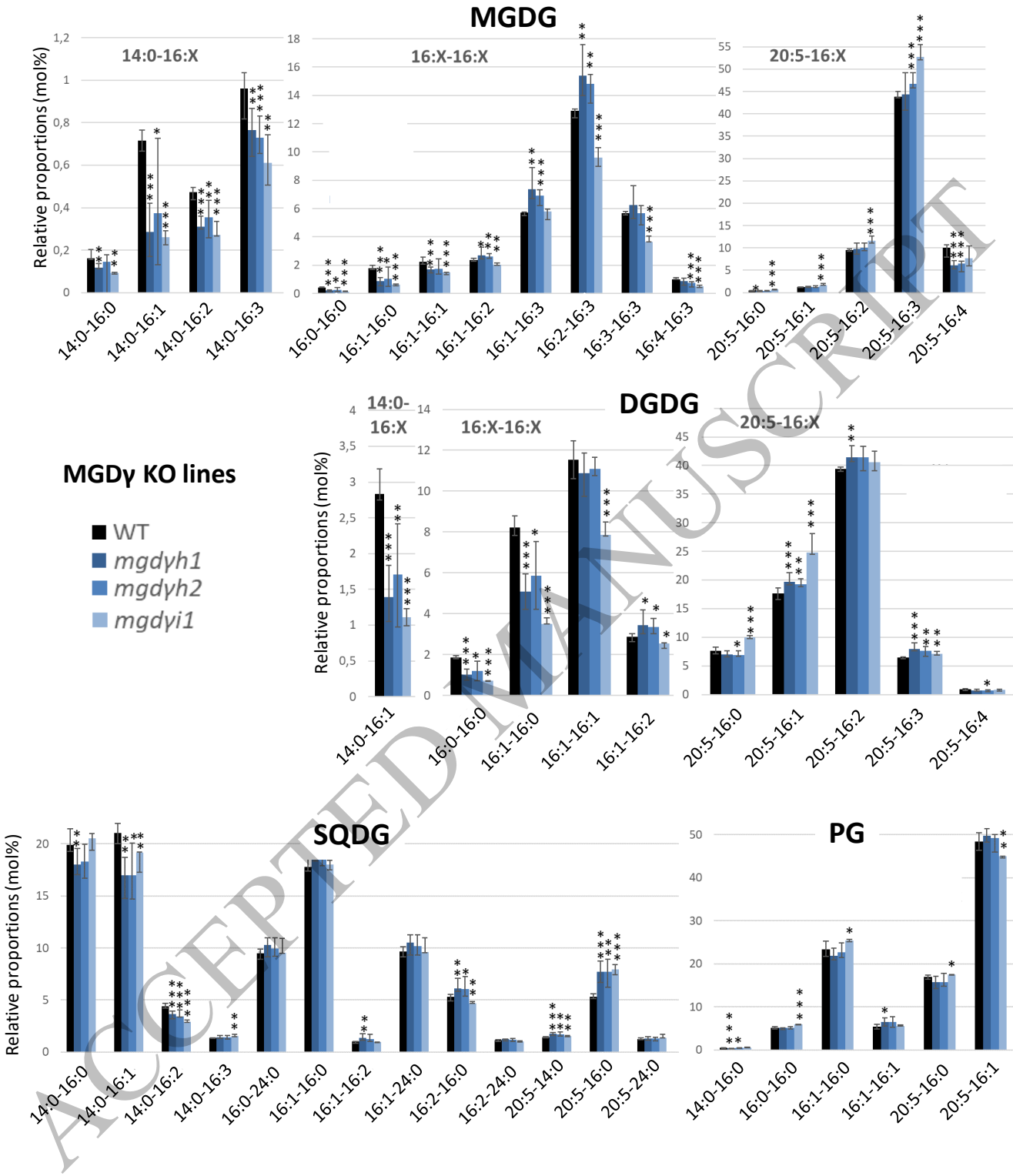

Figure 7: Impact of MGDγ mutations on MGDG, DGDG, SQG and PG molecular species. Lipid molecular species were analysed by LC-MS/MS. Each result is the median of six biological replicates ± min and max values. (*), *P-value* < 5.10^{-2}; (**), *P-value* < 1.10^{-2}; (***), *P-value* < 1.10^{-3}, based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

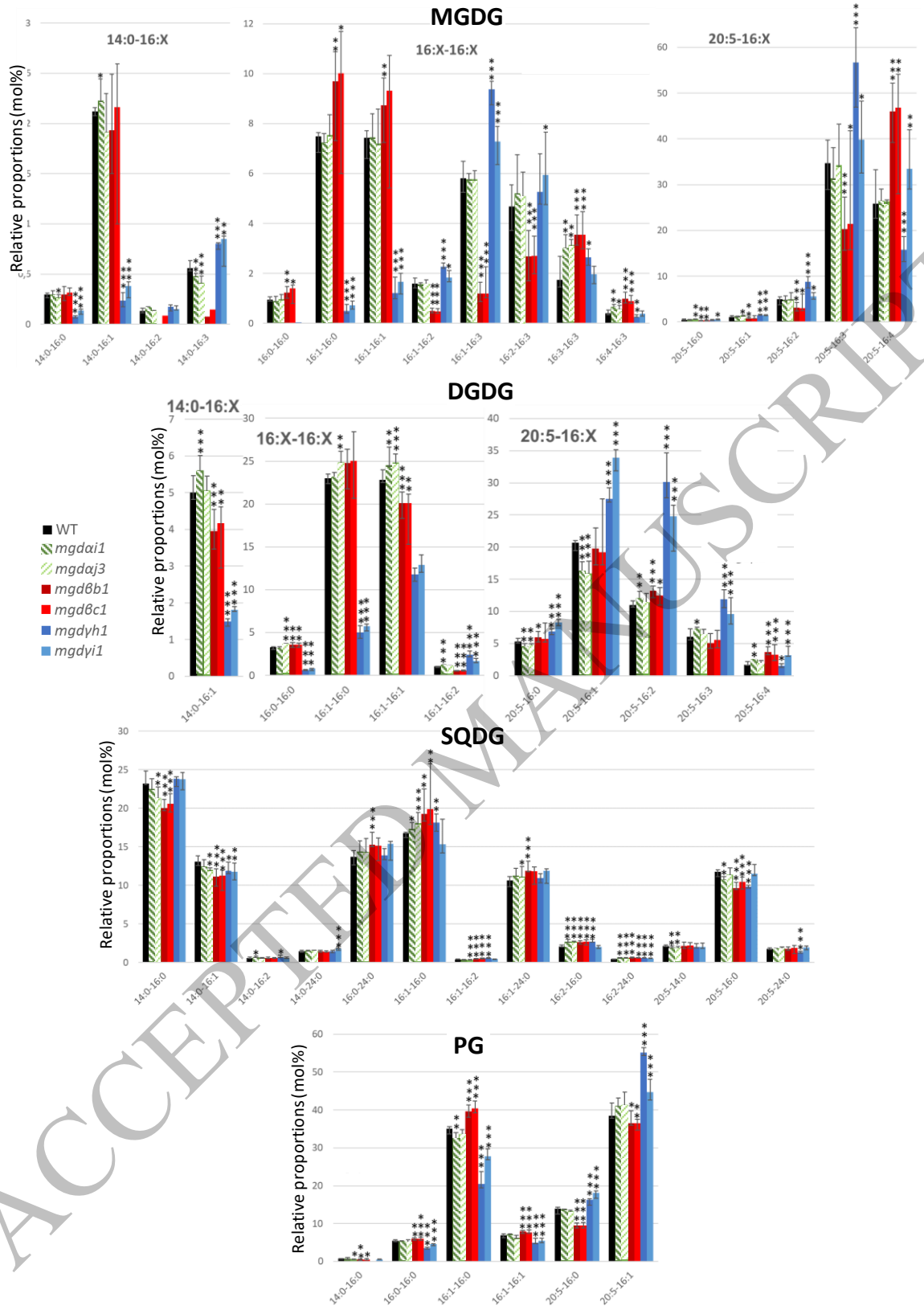

Figure 8: Impact of MGD mutations on the molecular species constituting MGDG, DGDG, SQDG and PG in nitrogen-deprived *P. tricornutum* cells. Lipid molecular species in MGDG, DGDG, SQDG and PG were analyzed by LC-MS/MS. Each result is the median of six biological replicates ± min and max values. (*), *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.

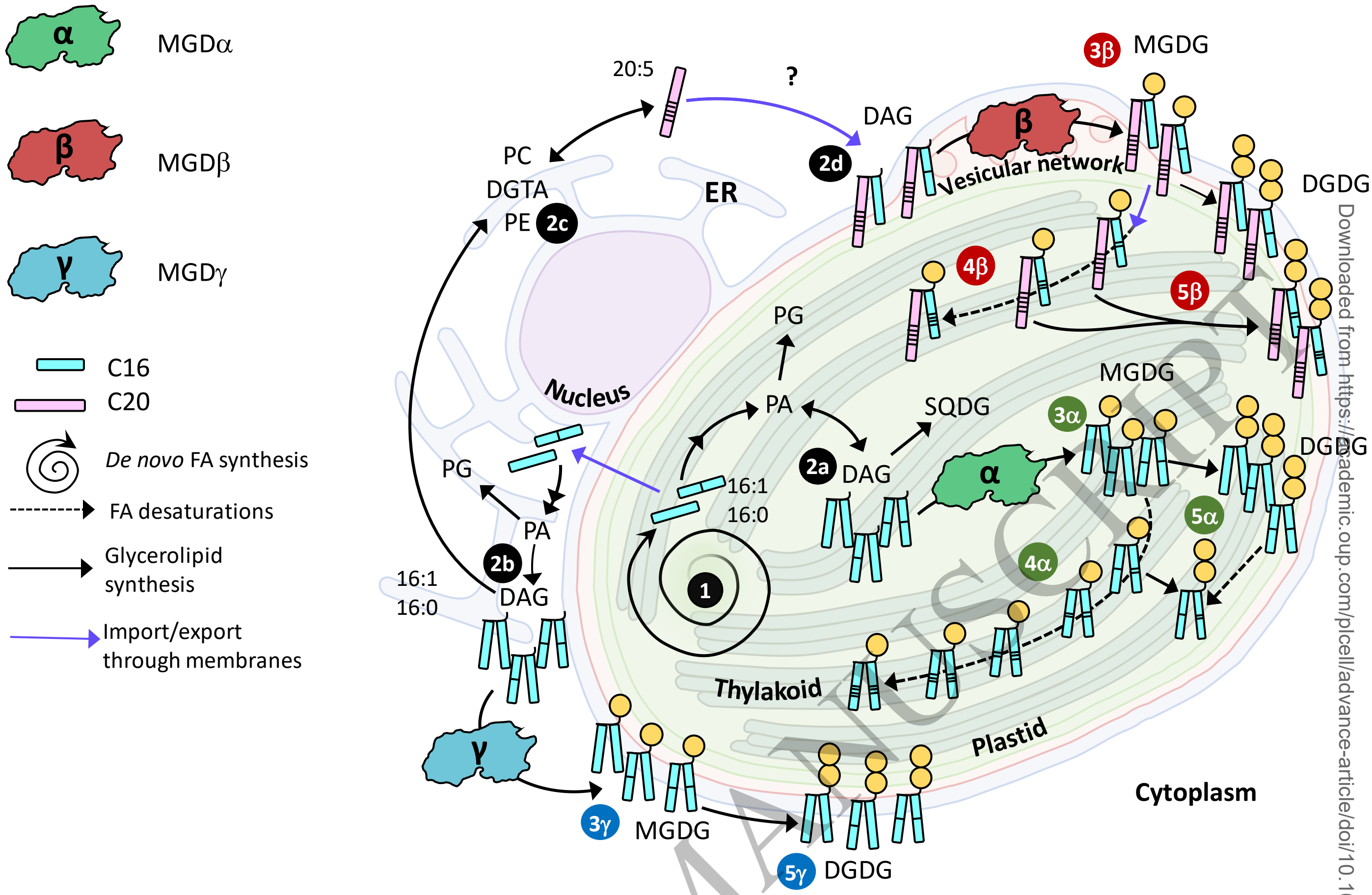

Figure 9. Function of MGDα, MGDβ and MGDγ in P. tricornutum cell. The model summarizes all compartmentalization, structural and functional data obtained in this study. FA are synthesized in the stroma as 16:0- and 16:1-ACP (1), where they can be used for the synthesis of DAG (2a). Alternatively, FA are 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). MGDα is localized at the level of thylakoids, and functional studies reported here are consistent with a utilization of DAG available inside the plastid (3α), to form MGDG species rapidly desaturated (16:1 □ 16:4) by plastid desaturases (4α). Based on obtained phenotype of KO and overexpressing lines, MGDG produced by MGDα is likely used by a specific DGD enzyme (5α) to form DGDG. MGDβ 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 (3β). MGDG produced can be desaturated (16:1 □ 16:4) by plastid desaturases (4β). Based on obtained phenotype of KO and overexpressing lines, MGDG produced by MGDβ is likely used by a specific DGD enzyme (5β) to form DGDG. MGDγ 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 (3γ), producing MGDG at a location far from plastid desaturases. Based on obtained phenotype of KO and overexpressing lines, MGDG produced by MGDγ is also likely associated to a specific DGD enzyme (5γ). In nitrogen shortage, MGDα protein level decreases drastically, and MGDG production relies on MGDβ and MGDγ. 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.

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., Petroutsos, 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., Laueffer, 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., Imberty, 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., Petroutsos, 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., Otillar, 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., Montsant, 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., Falciatore, A., Juillerat, A., Beurdeley, M., Voytas, D.F., Cavarec, L., and Duchateau, P. (2014). Genome engineering empowers the diatom Phaeodactylum tricornutum 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 Falciatore, A. (2009). Gene silencing in the marine diatom Phaeodactylum tricornutum. 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., Mulot, 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 tricornutum. 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 periplastidal 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., Kohl, 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., Knyaz, 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., Brugiere, 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., Brugiere, S., Schaller, H., Kuntz, M., Putaux, J.L., Jouneau, P.H., Rebellle, 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

Serôdio, 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., Falciatore, 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 tricornutum. 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 tricornutum. 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., Luypaert, 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

Zaslavskaia, L.A., Lippmeier, J.C., Kroth, P.G., Grossman, A.R., and Apt, K.E. (2000). Transformation of the diatom Phaeodactylum tricornutum (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 tricornutum by electroporation. Mar Genomics 16, 63-66.
Google Scholar: Author Only Title Only Author and Title