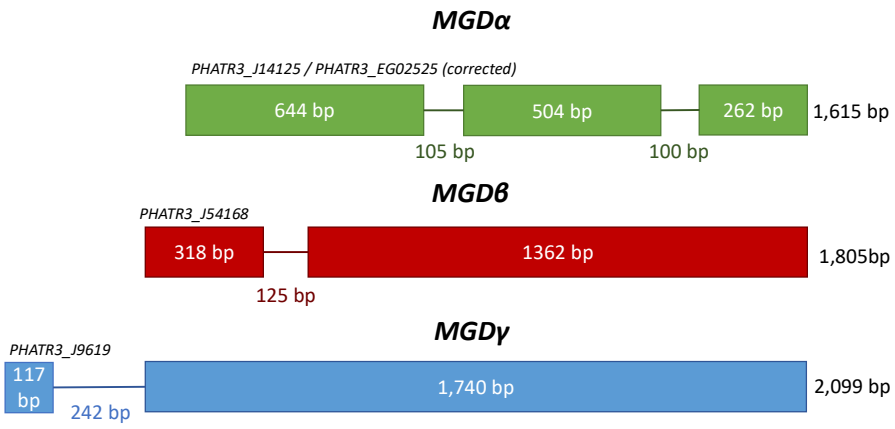
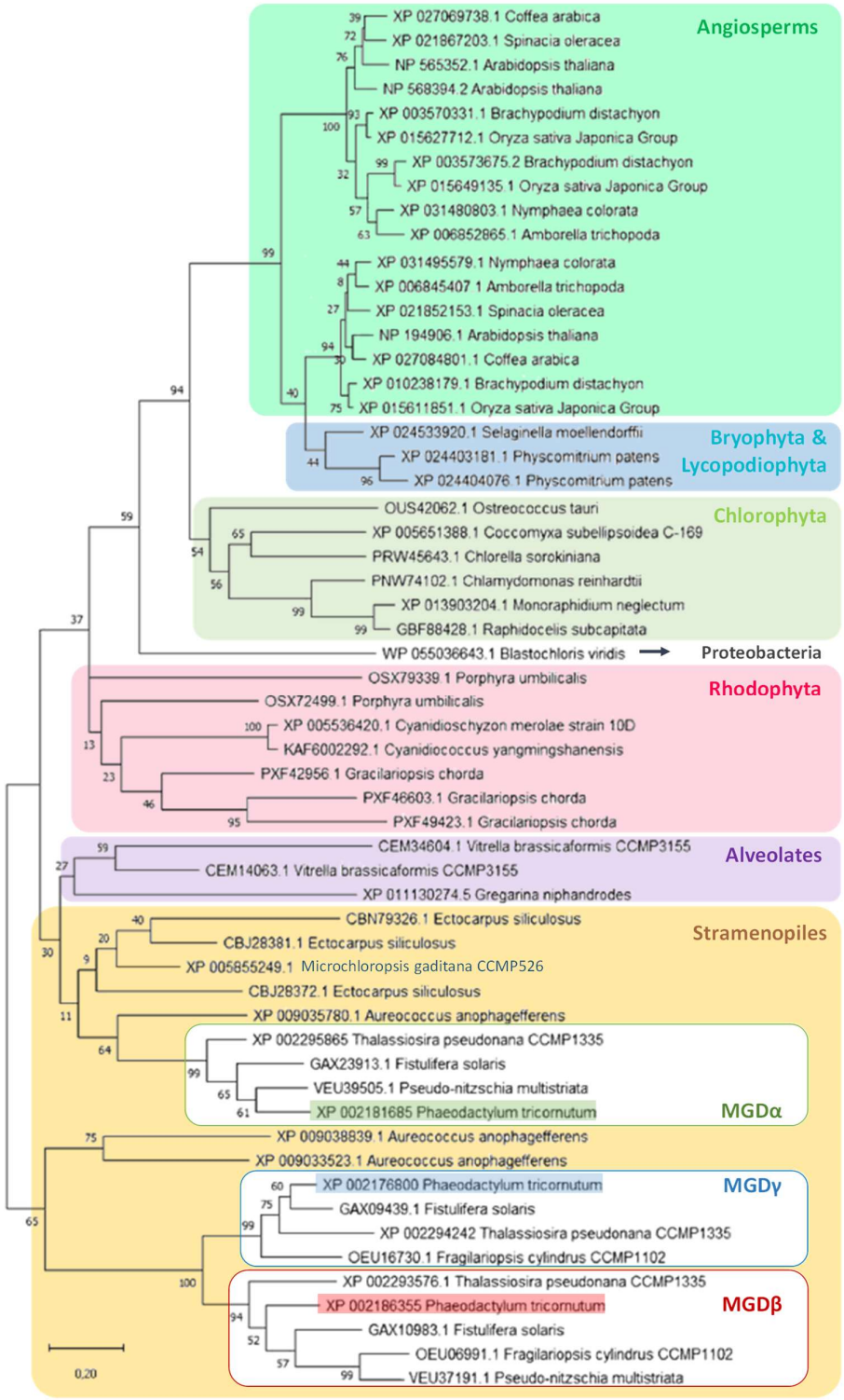
Monogalactosyldiacylglycerol synthase isoforms play diverse roles inside and outside the diatom plastid

Nolwenn Guéguen, Yannick Sérès, Félix Cicéron, Valérie Gros, Grégory Si Larbi, Denis Falconet, Etienne Deragon, Siraba D. Gueye, Damien Le Moigne, Marion Schilling, Mathilde Cussac, Dimitris Petroutsos, Hanhua Hu, Yangmin Gong, Morgane Michaud, Juliette Jouhet, Juliette Salvaing, Alberto Amato, Eric Maréchal.

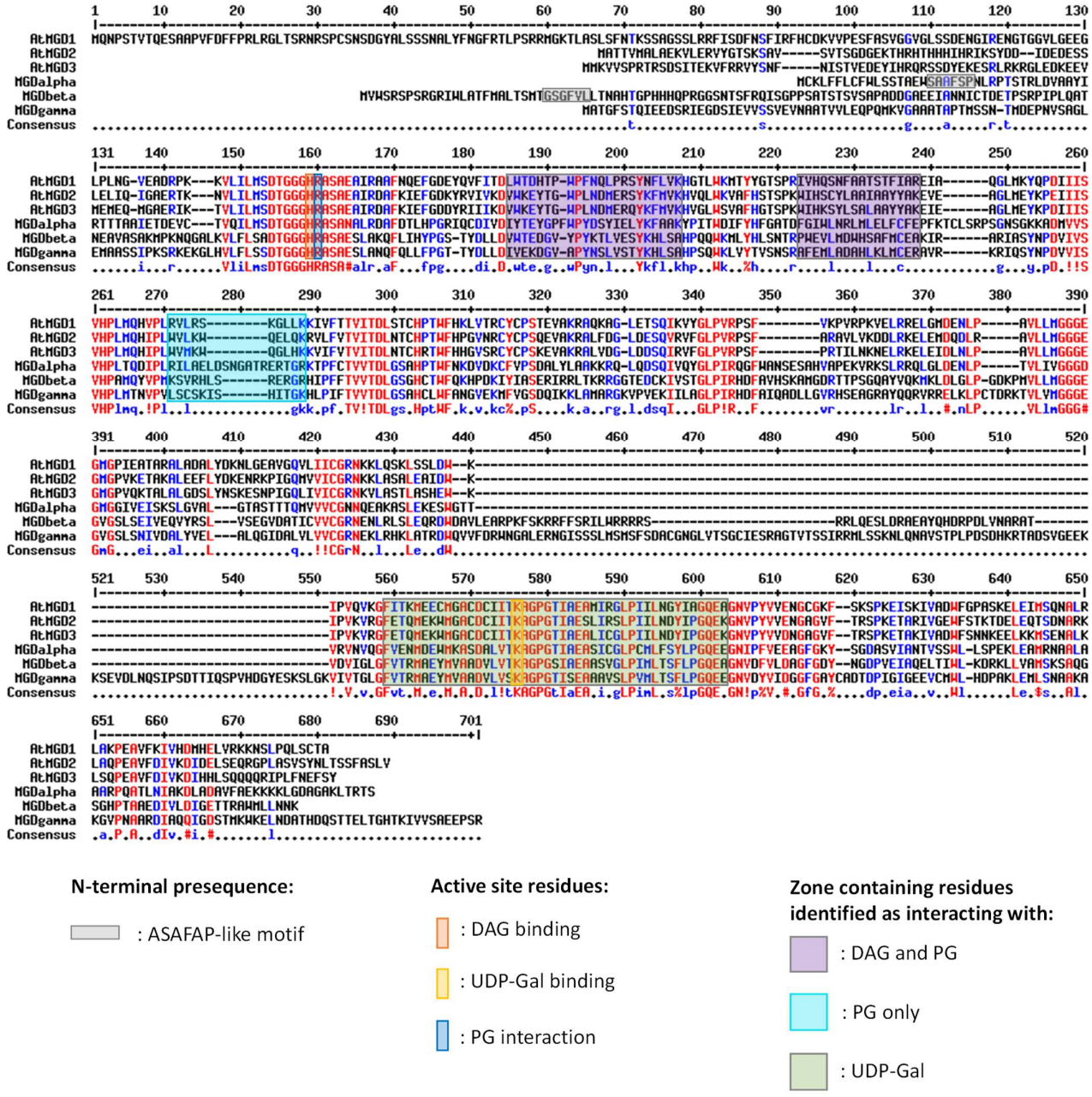
Supplementary Material



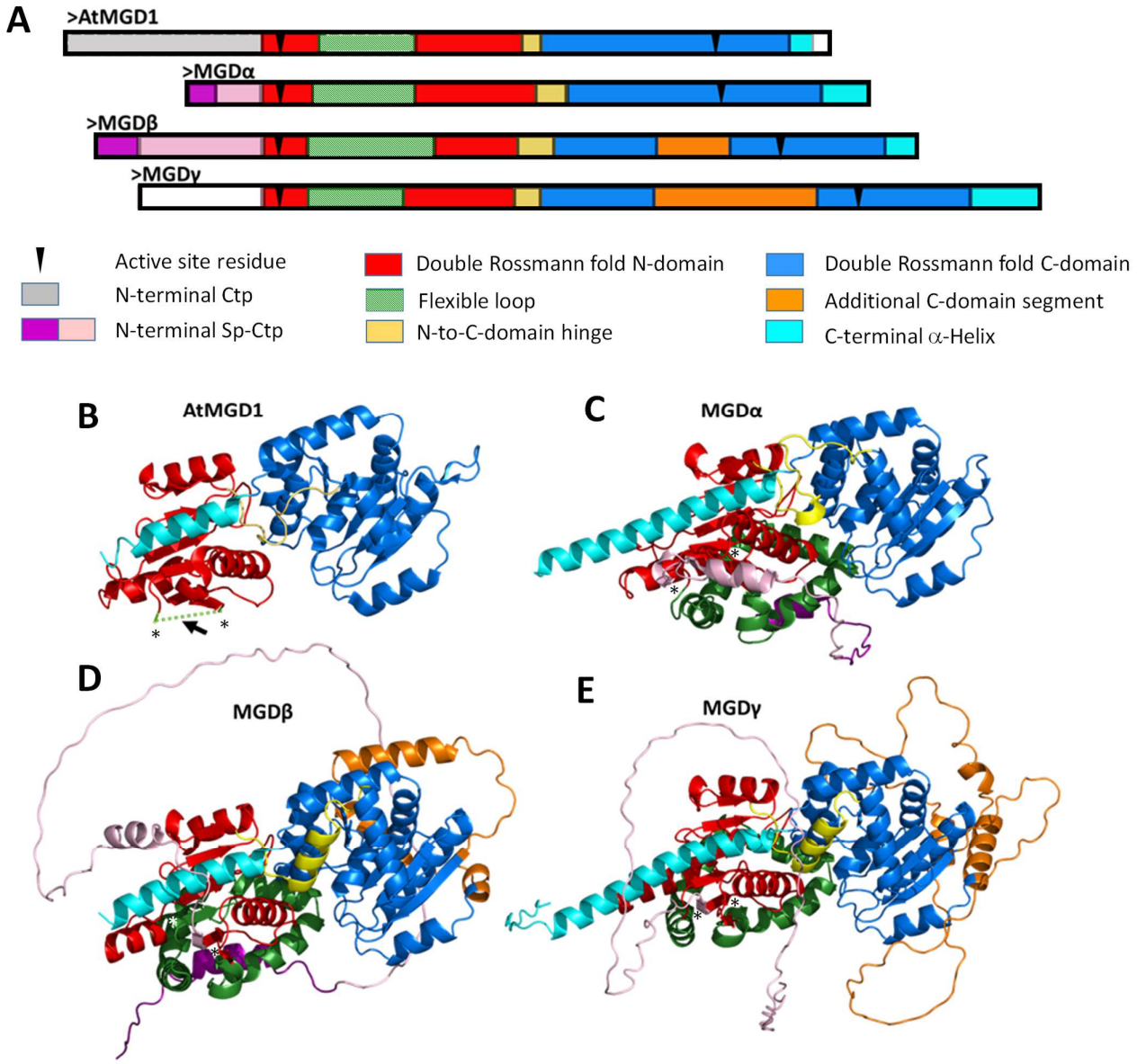
Supplementary Figure S1: Schematic view of *MGD* genes in *P. tricornutum*. Different sequences were available for each gene according to NCBI and EnsemblProtists databases (Supplementary Table 1). Manual curation of start codons combined cDNA sequencing in Pt1 8.6 strain and analysis of N-terminal addressing sequence in translated proteins. Based on EnsemblProtists, MGDα corresponds to two predicted CDS, Phatr3_J14125 and Phatr3_EG052525, containing 1,828 and 1,930 bp, respectively. Sequencing of MGDα cDNA and manual curation confirmed the position of introns and corrected the position of the start codon, determining a sequence length of 1,615 bp (see also Supplementary Table 1). MGDβ is a 1,805-bp gene according to the Phatr3_J54168 sequence from EnsemblProtists, harboring two exons. This structure was validated by cDNA sequencing and manual curation. MGDγ is the longest of the three genes. The Phatr3_J9619 sequence from EnsemblProtists has a length 2,099 bp and contains two exons. Sequencing of MGDγ cDNA and manual curation confirmed the splicing of the first intron of 242 bp. Lines indicate introns and boxes indicate exons.



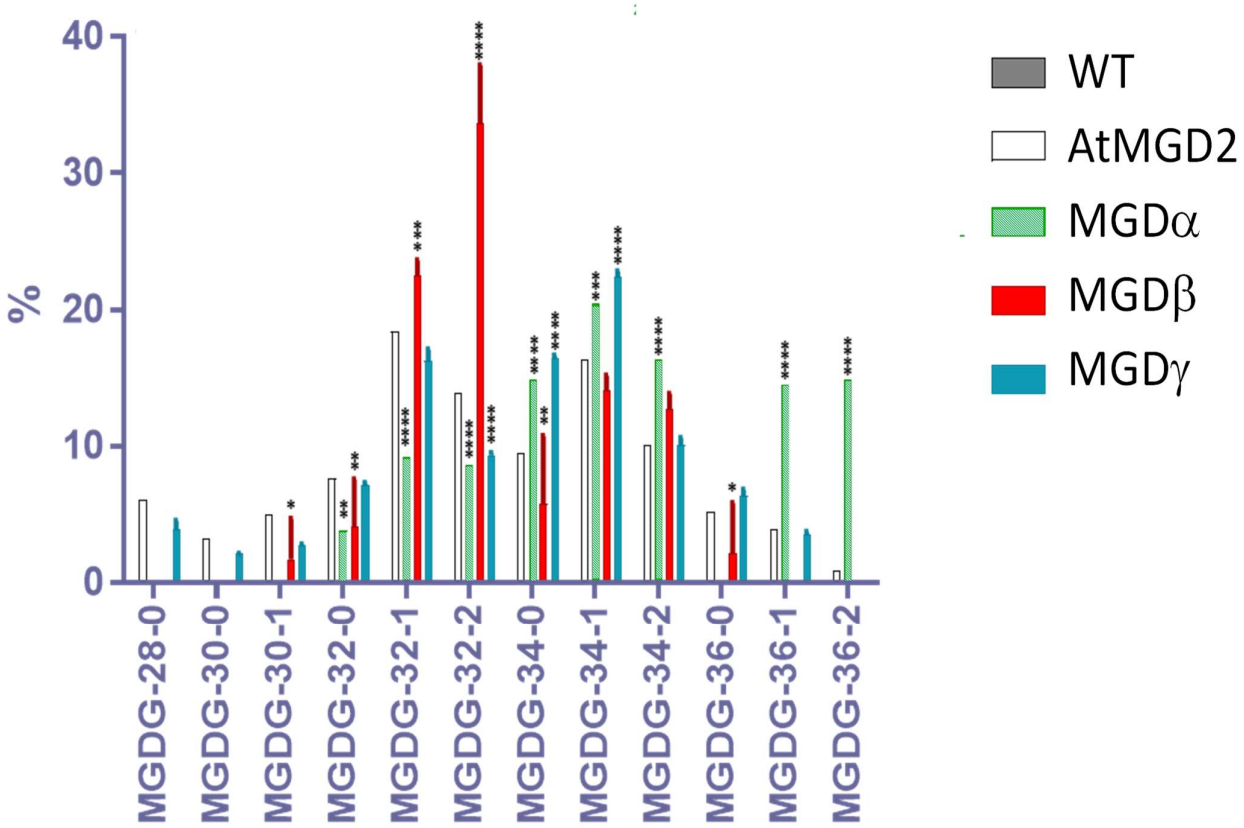
Supplementary Figure S2: Phylogenetic analysis of MGDs from plastid-containing eukaryotes. MGD sequences were selected to cover the biodiversity of plastid-containing eukaryotes from green algae (*Chlamydomonas*, *Chlorella*, *Coccomixa*, *Monoraphidium*, *Ostreococcus* and *Raphidocelis* spp.), red algae (*Cyanidioschyzon*, *Cyanidiococcus*, *Gracillariopsis* and *Porphyra* spp.), mosses (*Physcomitrium*), lycopods (*Selaginella*), angiosperms (*Amborella*, *Arabidopsis*, *Brachypodium*, *Coffea*, *Nimphea*, *Oryza* and *Spinacia* spp.), stramenopiles (*Aureococcus*, *Ectocarpus*, *Microchloropsis*, - including diatoms, *i.e. Fistulifera*, *Fragilariopsis*, *Phaeodactylum*, *Pseudo-nitzschia* and *Thalassiosira* spp.), to alveolates (*Vitrella* and *Gregarina* spp.). *Fistulifera solaris* being a diatom with an allodiploid genome structure, only one of each homeologous gene was used in the phylogeny reconstruction. The tree was inferred by Maximum Likelihood, with 5,000 bootstrap pseudoreplicates, using the MEGAX software. The percentage of trees in which the associated taxa clustered together is shown next to the corresponding branches. Branch lengths are proportional to the number of substitutions per site (scale bar = 0.20). Main phyla are highlighted. Clusters corresponding to diatom MGDα, MGDβ and MGDγ are framed in green, red and blue, respectively. This tree highlights the conservation of a MGD multigenic family in diatoms, and the absence of evolutionary relationships with the multigenic family previously characterized in angiosperms.



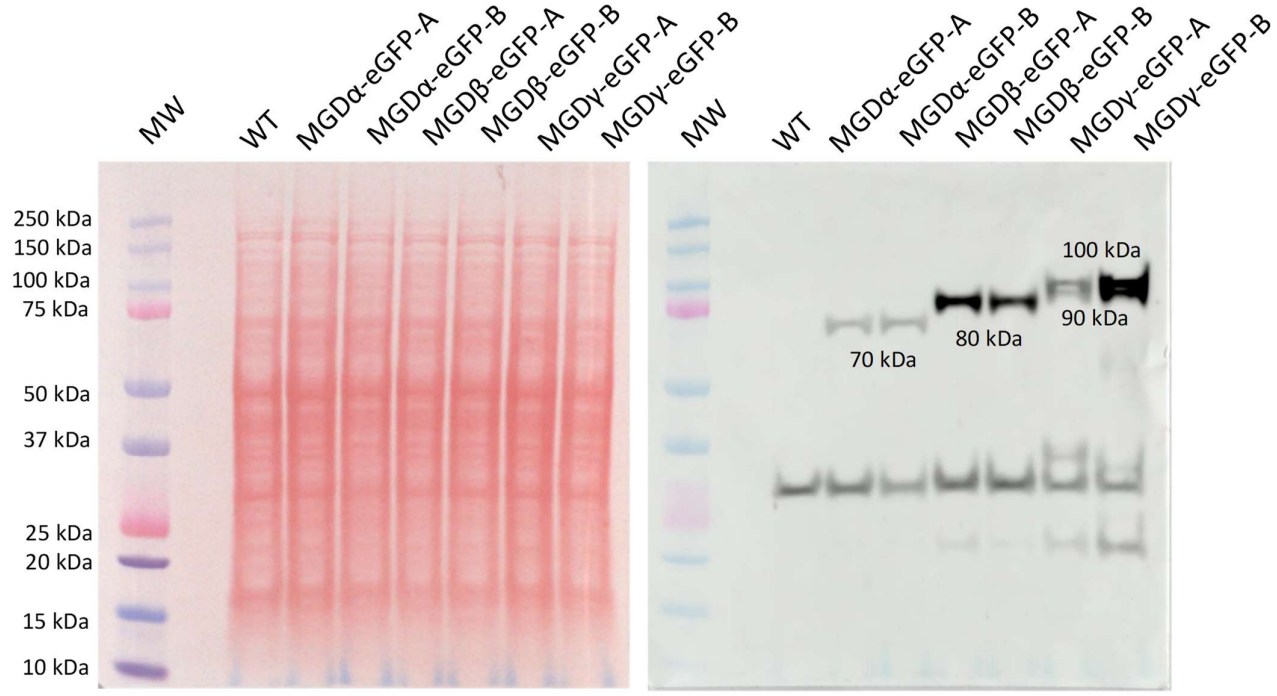
Supplementary Figure S3: Multiple sequence alignment of *Arabidopsis thaliana* AtMGD1, AtMGD2 and AtMGD3 and *Phaeodactylum tricornutum* MGDα, MGDβ and MGDγ proteins. Comparison of the position of amino acids known to be involved in the interaction of AtMGD1 with its substrates diacylglycerol (DAG) and UDP-Galactose (UDP-Gal), as well as the activator phosphatidylglycerol (PG), with the amino acid sequences of *P. tricornutum* MGDs. Residue colour differentiates high (red) and medium (blue) conservation. The catalytic residues binding to DAG and UDP-Gal are a histidine (framed in orange) in position 159 and a lysine (framed in yellow) in position 576 of the alignment, respectively. The residue interacting with PG and allowing a potential PG-His catalytic dyad is an arginine (framed in dark blue) in position 160 of the alignment. Zones containing a high number of residues in AtMGD1 interacting with DAG, PG and UDP-Gal are indicated with purple squares when residues interact with both DAG and PG, with light blue squares when interacting with PG only, and with green squares when interacting with UDP-Gal.



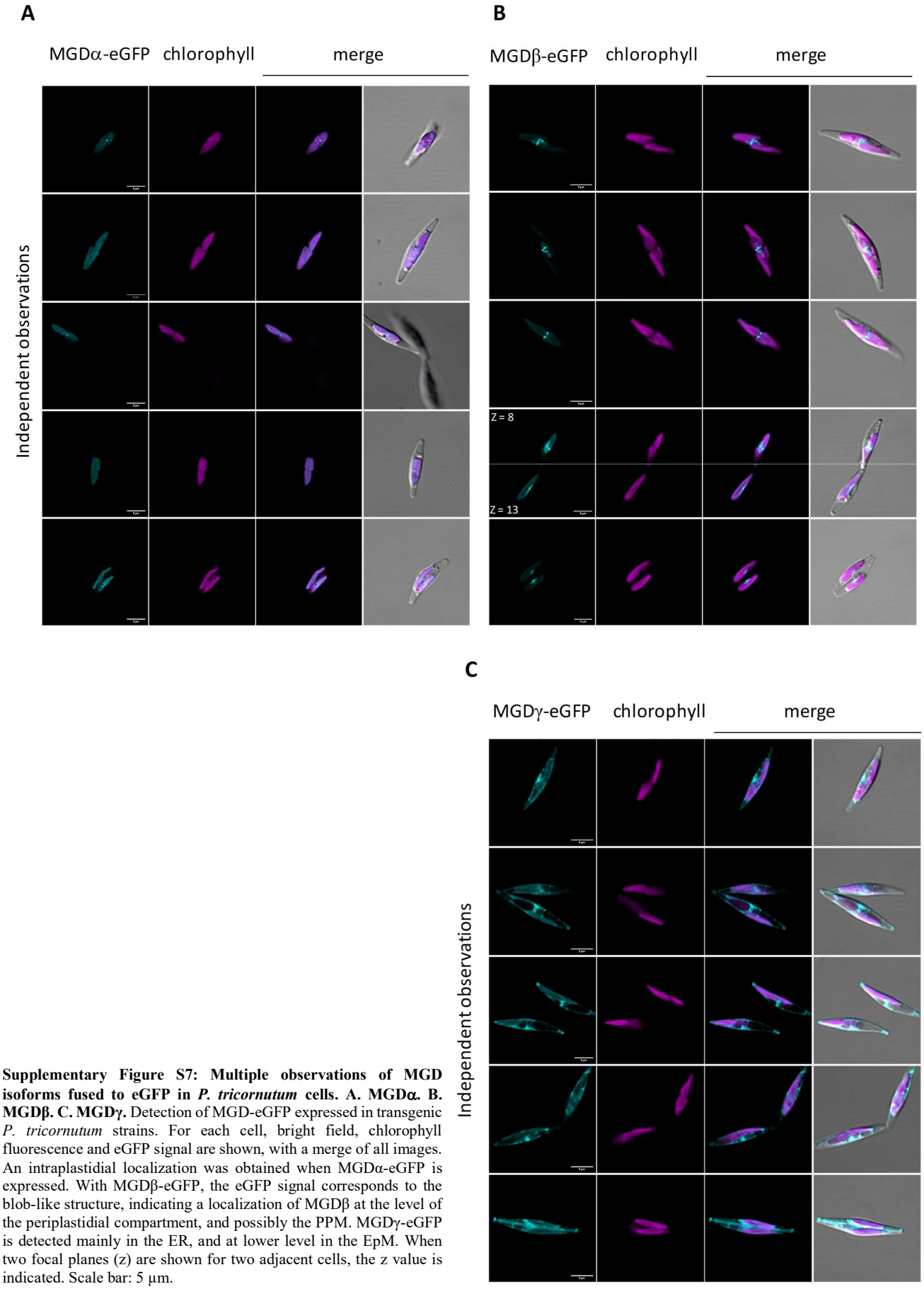
Supplementary Figure S4: Protein models of *P. tricornutum* MGDs. (A) Linear representation of conserved structural domains. (B) Protein structure of AtMGD1 from *A. thaliana*. The AtMGD1 protein model (4X1T) (Rocha et al., 2016) was retrieved from the Protein Data Bank (wwPDB consortium, 2019). (C-E) Protein models of MGDα, MGDβ and MGDγ from *P. tricornutum*. Models were obtained with corrected MGD protein sequences, using AlphaFold. Structures are viewed with PyMOL software in the same orientation. The identified protein segments are shown as follows. In the N-terminal end, AtMGD1 Ctp is shown in grey; putative MGDα and MGDβ Sp-Ctp in purple; other non-conserved N-terminal sequences in pink. In the double Rossmann fold N-domain, conserved sequences are shown in red, the AtMGD1 loop with unresolved structure in green (with a pattern with stripes in A, and with stars at the start and end positions of the loop in C, D and E). The hinge connecting the N- and C-domains of the double Rossmann fold is shown in yellow. The double Rossmann fold C-domain is shown in blue; and additional segments in MGDβ and MGDγ in orange. The C-terminal alpha helix that folds in the direction of the N-domain is shown in light blue. Black triangles represent the localization of conserved amino acids involved in the enzymatic reaction. AlphaFold attributes a high accuracy index (90 – 100) for all modeled portions showing homology with AtMGD1 and a low one (50 – 70) for the non-conserved N-terminal ends, as well as for the second half of the C-terminal closing helix for MGDβ and MGDγ.



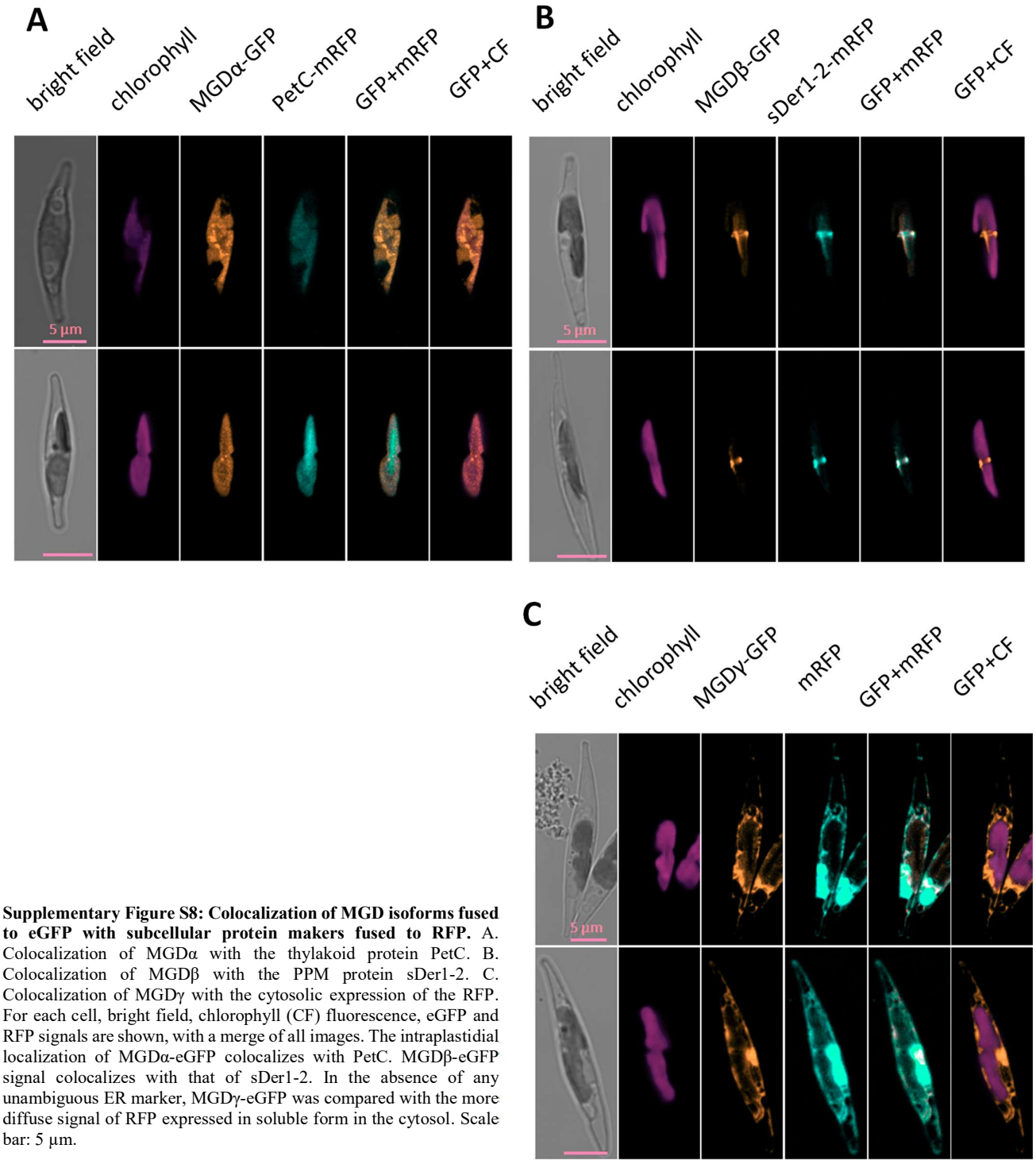
Supplementary Figure S5. Heterologous expression of *Phaeodactylum tricornutum* MGD isoforms in yeast. LC-MS/MS lipid quantification of *Saccharomyces cerevisiae* strain BY4741 expressing AtMGD2 (used as positive control), MGDα, MGDβ, and MGDγ. For *P. tricornutum* MGD isoforms, three yeast expressing lines were analyzed per gene with a total of six replicates. For MGDα, only one line showed a synthesis of MGDG, although the two other lines expressed the transgene. An ANOVA analysis was performed and the adjusted P values referring to the results of Tukey's multiple comparisons test with a cutoff of 95% confidence intervals are reported on the figure (* < 0.05, ** <0.01, *** < 0.005, **** < 0.0001). Graphical representations and One-way ANOVA followed by Dunnett's multiple comparisons test was performed using GraphPad Prism. MGD, monogalactosyldiacylglycerol synthase; MGDG, monogalactosyldiacylglycerol.



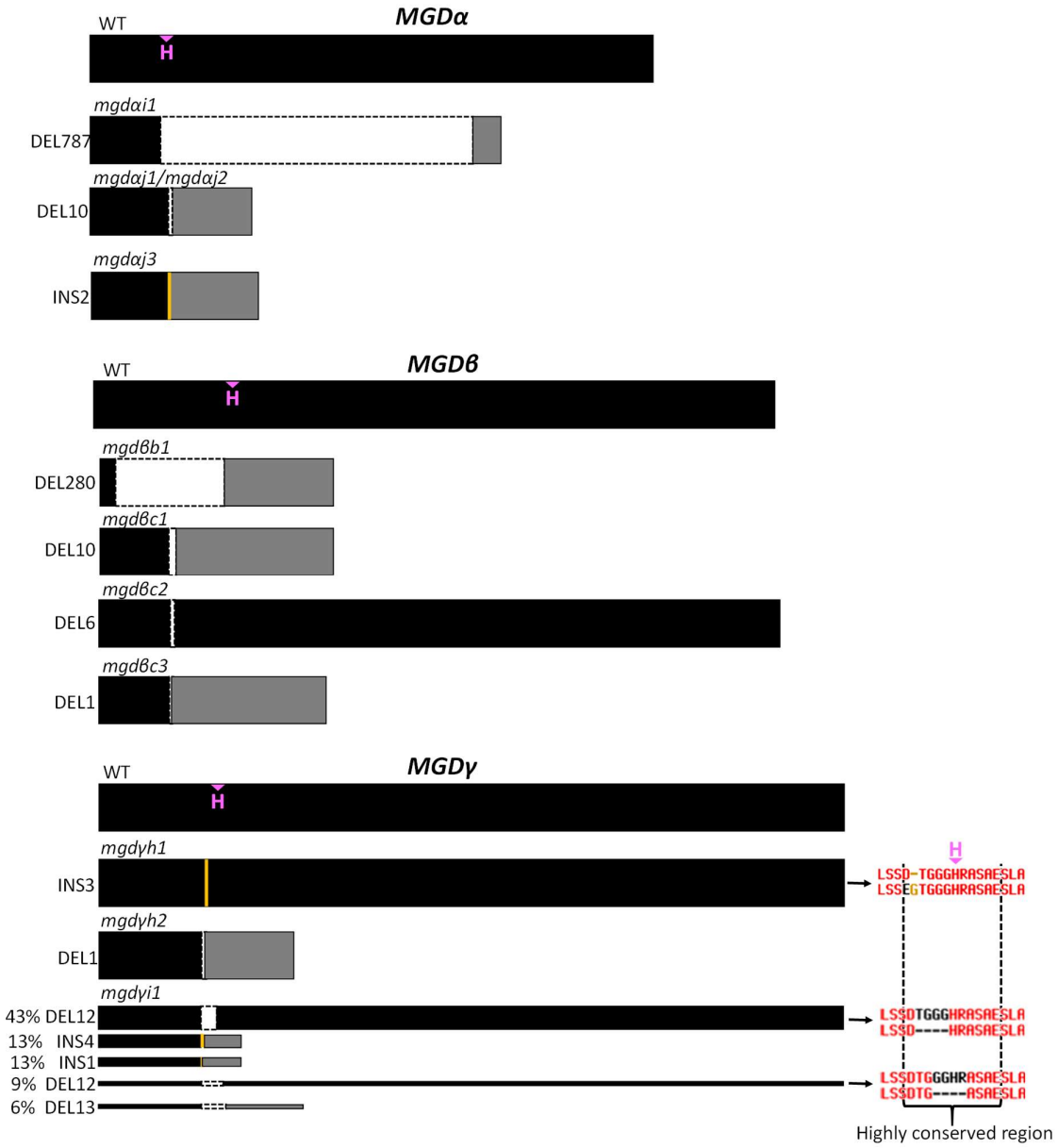
Supplementary Figure S6: Visualization of MGD-eGFP proteins overexpression by immunoblot. Ponceau staining (left) and western blot analyses (right) of total protein extracts from *MGD-eGFP* overexpressing *P. tricornutum* lines using wild type (WT) as a control. Quantity of total protein extract per well: 10 µg. Membrane revelation was performed with an anti-GFP-HR antibody at a 1/5000 dilution.



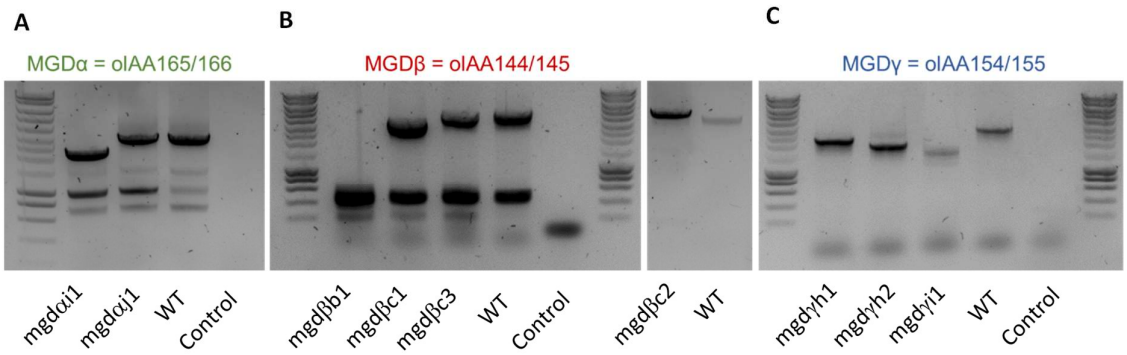
Supplementary Figure S7: Multiple observations of MGD isoforms fused to eGFP in *P. tricornutum* cells. A. MGDα. B. MGDβ. C. MGDγ. Detection of MGD-eGFP expressed in transgenic *P. tricornutum* strains. For each cell, bright field, chlorophyll fluorescence and eGFP signal are shown, with a merge of all images. 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 periplastidial compartment, and possibly the PPM. MGDγ-eGFP is detected mainly in the ER, and at lower level in the EpM. When two focal planes (z) are shown for two adjacent cells, the z value is indicated. Scale bar: 5 µm.



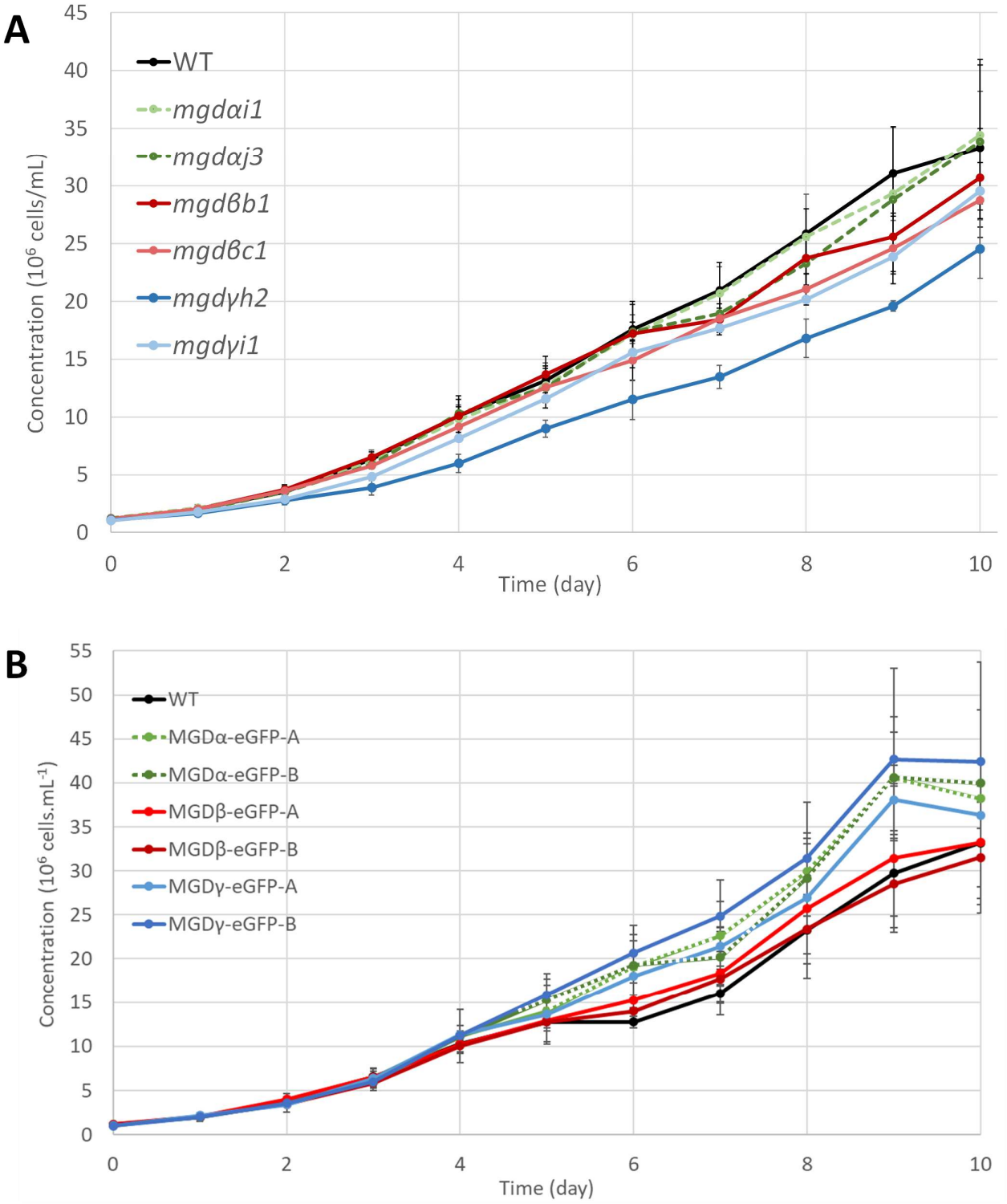
Supplementary Figure S8: Colocalization of MGD isoforms fused to eGFP with subcellular protein makers fused to RFP. A. Colocalization of MGDα with the thylakoid protein PetC. B. Colocalization of MGDβ with the PPM protein sDer1-2. C. Colocalization of MGDγ with the cytosolic expression of the RFP. For each cell, bright field, chlorophyll (CF) fluorescence, eGFP and RFP signals are shown, with a merge of all images. The intraplastidial localization of MGDα-eGFP colocalizes with PetC. MGDβ-eGFP signal colocalizes with that of sDer1-2. In the absence of any unambiguous ER marker, MGDγ-eGFP was compared with the more diffuse signal of RFP expressed in soluble form in the cytosol. Scale bar: 5 µm.



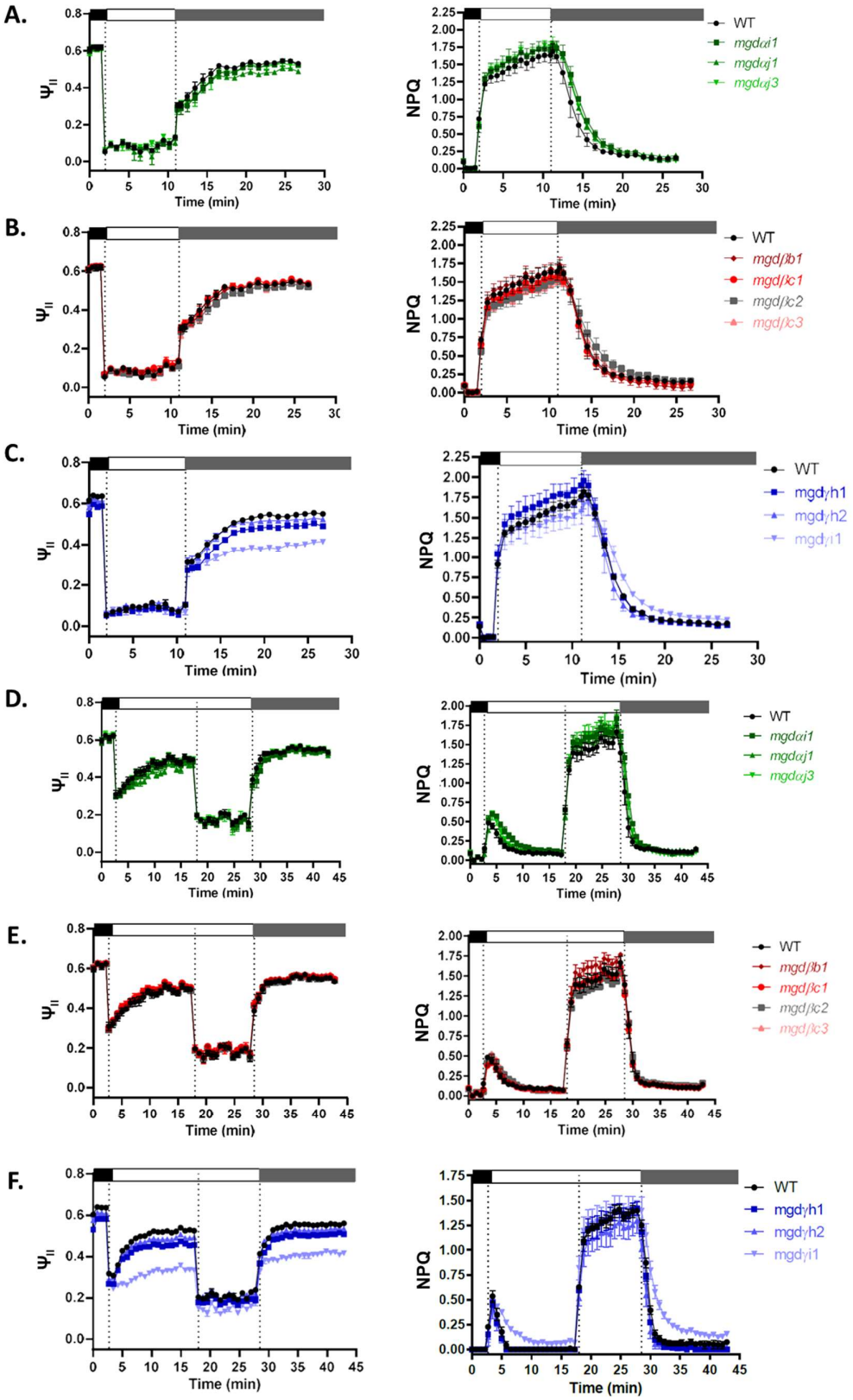
Supplementary Figure S9: Selected knock-out MGDα, MGDβ and MGDγ mutants, generated by CRISPR-Cas9 editing. Schematic representation of the impact of the indels introduced on the protein sequences in each mutant lines. The relative position of the catalytic histidine is indicated in pink in the wild type protein sequences. The nature of the indels is indicated next to the mutant protein sequence in base pair unit. The colors of the blocks indicate if the sequence is in frame (black) or not (grey) with the wild type protein sequence. White blocks correspond to portions of the sequence that are absent due to a deletion. Yellow marks indicate the position of an insertion. In the case of *mgdγi1* mixed mutant, no WT sequence remained in the mixed-mutated colony; the five most abundant indels were represented, with a height corresponding to their proportion. The mg*dγh1* and *mgdγi1* strains contain indels that do not induce a frameshift. For these, alignments of the wild type and mutant sequences show the impact of the indels on a highly conserved region of the protein. INS, insertion; DEL, deletion.



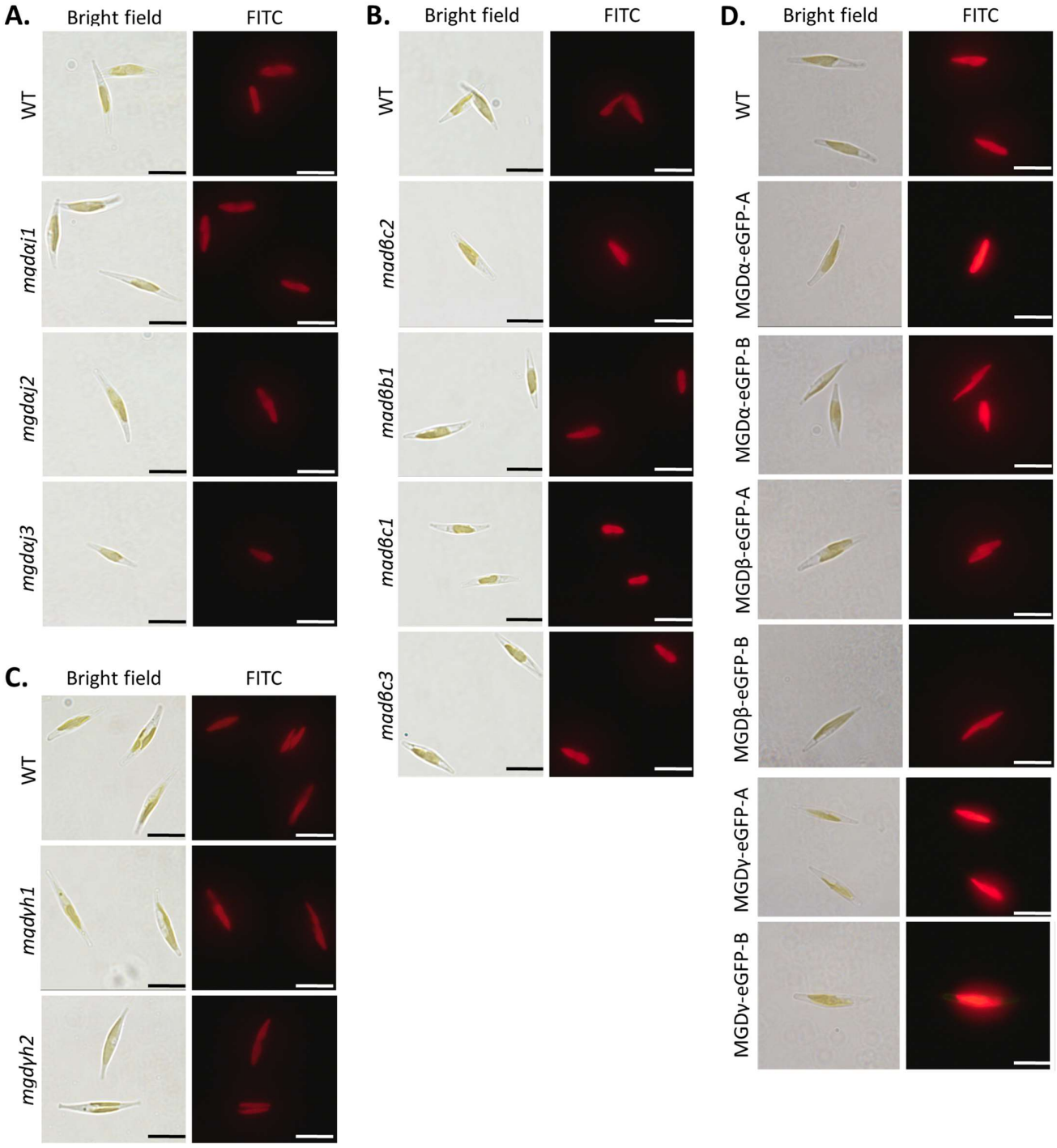
Supplementary Figure S10: Analysis of loss of heterozygosity and stability of MGD KO lines. In different biological systems, including *P. tricornutum*, the genomic regions where mutations have been induced by CRISPR-Cas9, can potentially experience a loss of heterozygosity (LoH) (Russo et al., 2018). LoH may involve large genomic regions and even entire chromatids. The disappearance of *P. tricornutum* single nucleotide polymorphism (SNP) in inter-allelic regions, allows a detection of LoH at each locus corresponding to mutated MGDs. All MGD mutants were analysed after one year of repeated rounds of cell cultivation in liquid medium. Amplicons were obtained with primer pairs designed ~700-800 bp up- and downstream each gene. Unspecific small size bands, present in both WT and mutant lines, were not considered. **A. MGDα KO lines.** All MGDα KO lines were stable and homozygous. In the *mgdαi1* diploid mutant, a single band with the initial 787 bp deletion is visible (A lane 2). As an illustration of LoH analysis, the 3,104 bp region amplified around the MGDα locus (-748 bp +741 bp) highlights three main SNPs in the WT, *i.e.* a TC repetition in the 5' region 232 bp upstream the gene and two C/T double peaks in exon 2. The mutant *mgdαi1* strain retained only one of the allelic variants identified in the WT, corroborating LoH. The 10-bp deletion in *mgdαi1* is faintly visible after migration on the gel (A, lane 3) and was confirmed as homozygous. **B. MGDβ KO lines.** All *mgdβ* mutants highlighted homozygous profiles and were stable at the time of analysis. The homozygous 280-bp deletion in *mgdβb1* was clearly detected (B, lane 2), whereas smaller deletions were less visible after gel migration. **C. MGDγ KO lines.** All *mgdγ* mutant lines have evolved after one year of liquid medium cultivation. The cultivated line corresponding to *mgdγh1* (C, lane 2) initially characterised by a 3-bp insertion and *mgdγh2* (C, lane 3), with a 1-bp deletion, have stabilized with larger deletions. The *mgdγh1* strain is homozygous, whereas *mgdγh2* line appears heterogeneous. The *mgdγhi1* line, which was initially heterogeneous, evolved as a homozygous line, with a MGDγ deletion larger than 1 kb (C, lane 4). This study confirms that all mutated strains used in this study could be mostly homozygous and stable over time (*mgdα*, *mgdβ*) or evolve with larger deletions after multiple rounds of cultivation (*mgdγ*).



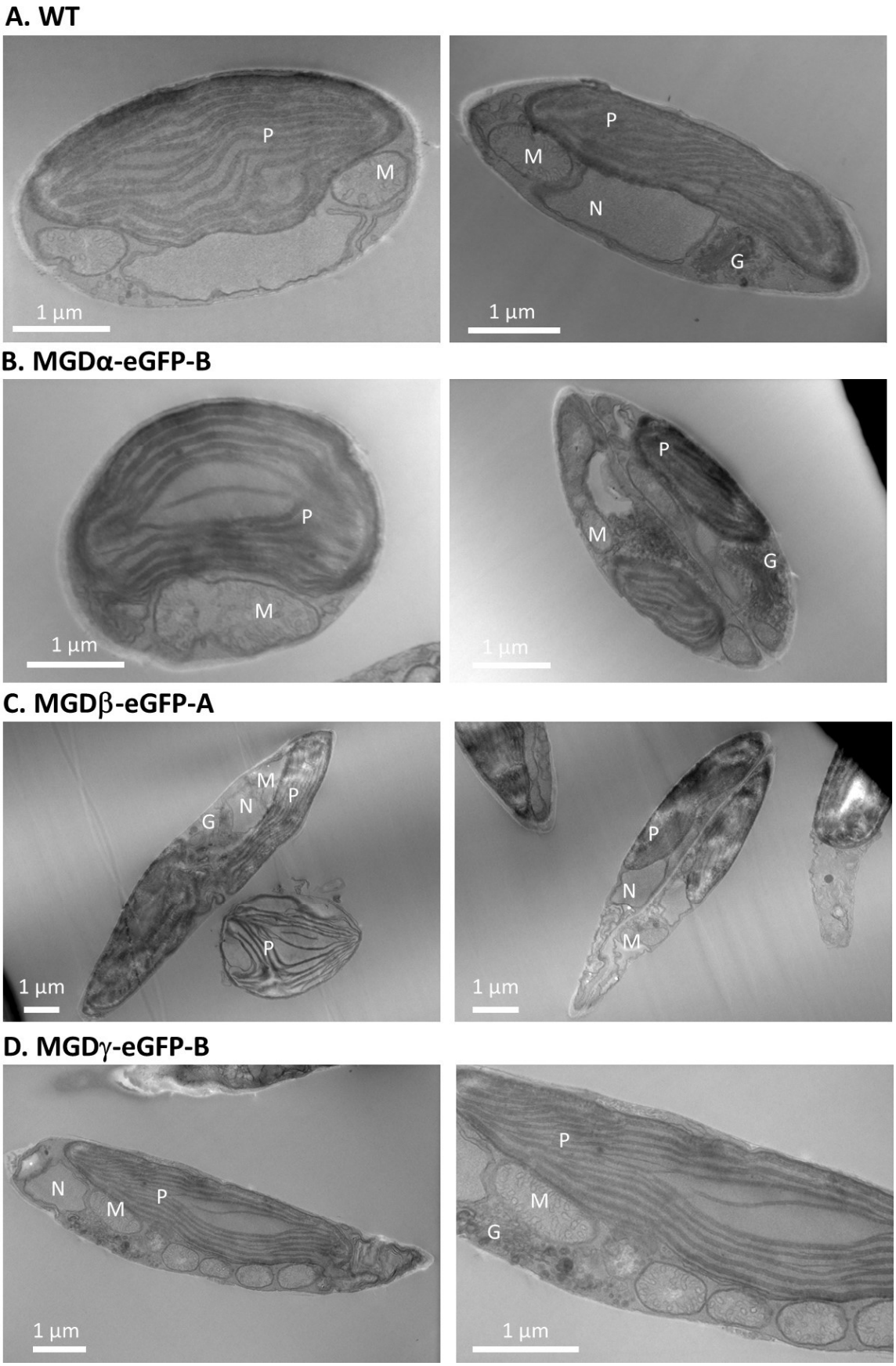
Supplementary Figure S11: Growth curves of MGD mutant and overexpressing strains compared to the WT. A. MGD KO lines. B. MGD-GFP overexpressing lines. Mutant and WT lines were cultivated in parallel at 20 °C in 100-mL ESAW 10N10P medium. Cell concentration (10^6 cells.mL^{-1}) was measured every day as indicated, using a TECAN Infinite M1000 PRO. Data are the average of three independent biological replicates ± SD.



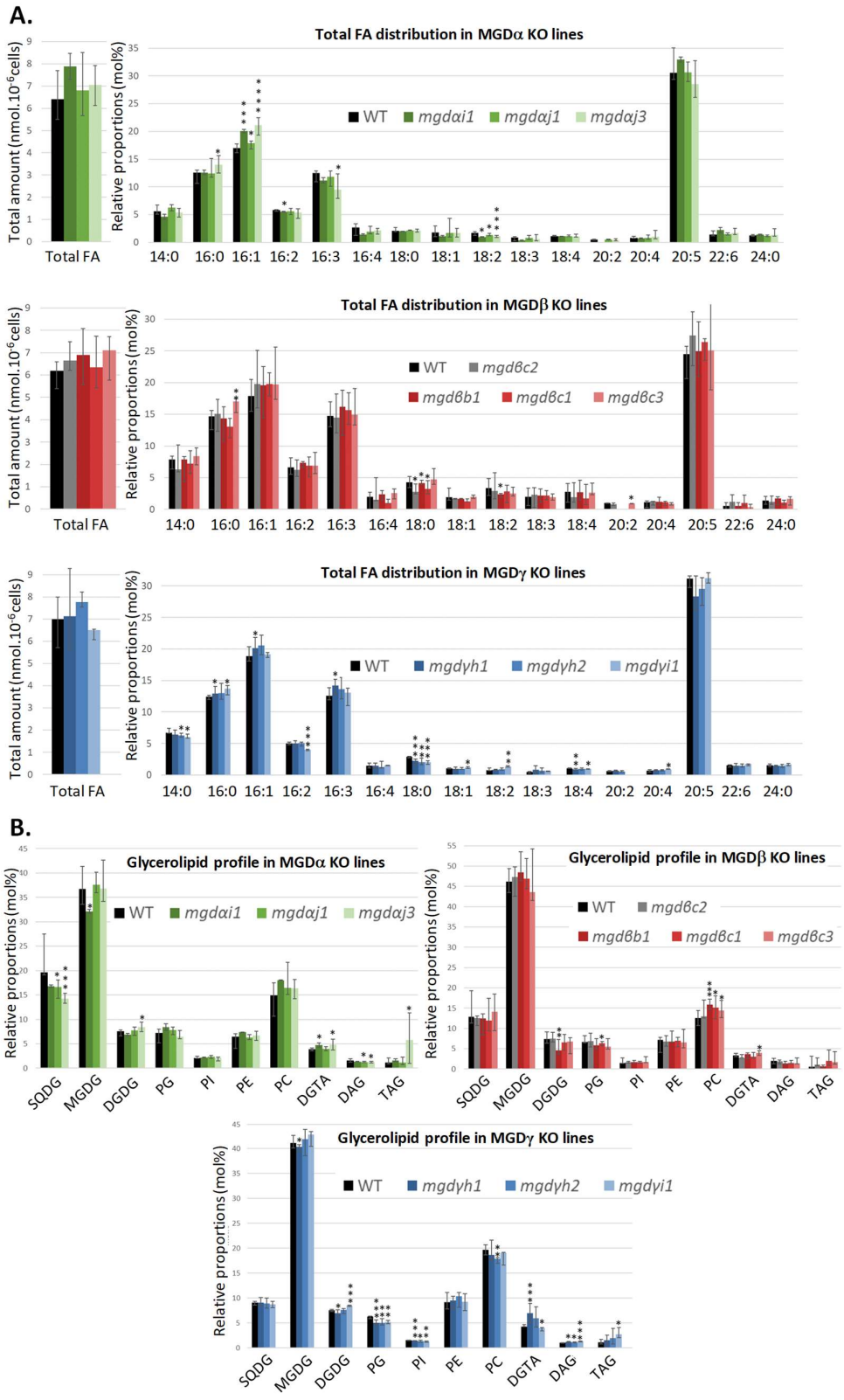
Supplementary Figure S12: Effect of MGDs mutations on photosynthetic properties under high and moderate light stresses. Cells were acclimated to darkness for 15 min prior to the onset of the measurements (first dashed line). For each *MGDα*, *MGDβ* and *MGDγ* mutant strains, chlorophyll fluorescence was recorded in the dark (shown as a black bar on top of the graph) and at 700 µmol photons m^{-2} s^{-1} (A-C, second dashed line) or at 55 and then 335 µmol photons m^{-2} s^{-1} (D-F, second and third dashed lines). In all cases the light stress was followed by relaxation at low light intensity (20 µmol photons m^{-2} s^{-1}; shown as a grey bar on top of the graph). Data were obtained with three independent replicates. Photosynthetic yield (ΨII) values were calculated as (*Fm*'-*F*)/*Fm*'; NPQ (non-photochemical quenching) values were calculated as (*F*m – *F*m')/*F*m. Data correspond to means ± SD.



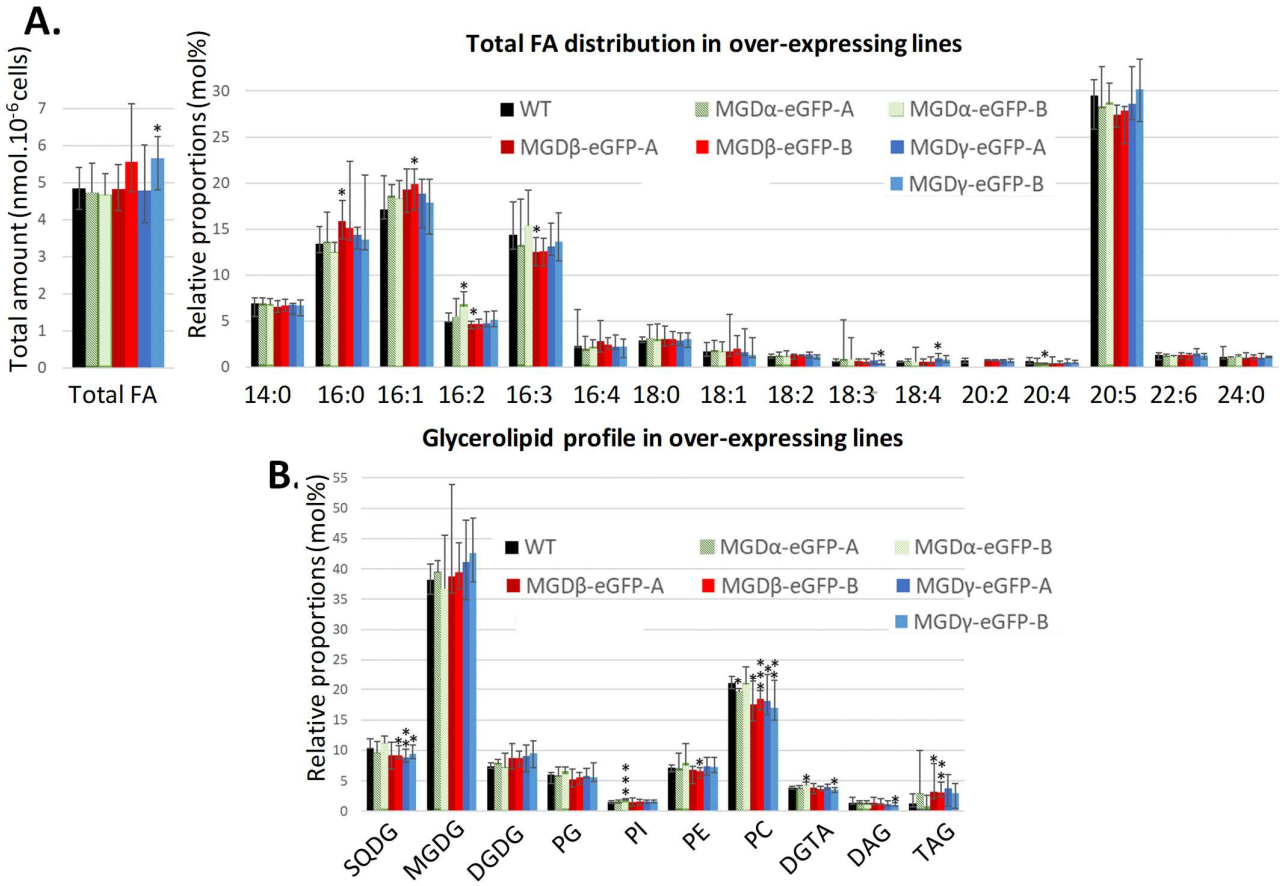
Supplementary Figure S13: Cell morphology of KO and overexpressing lines. Cells were observed with an epifluorescence microscope. Mutant strains for MGDα (A), MGDβ (B) and MGDγ (C) were cultured independently, in parallel with a wild type (WT), while overexpressing strains MGDα-eGFP-A and -B, MGDβ-eGFP-A and -B, and MGDγ-eGFP-A and -B were all cultured in parallel, with a WT (D). Bright field and chlorophyll autofluorescence images were captured with an oil-immersed objective 100x. Chlorophyll fluorescence was observed with an FITC filter (BP485/20, FT510, LP515). Scale bar: 10 µm.



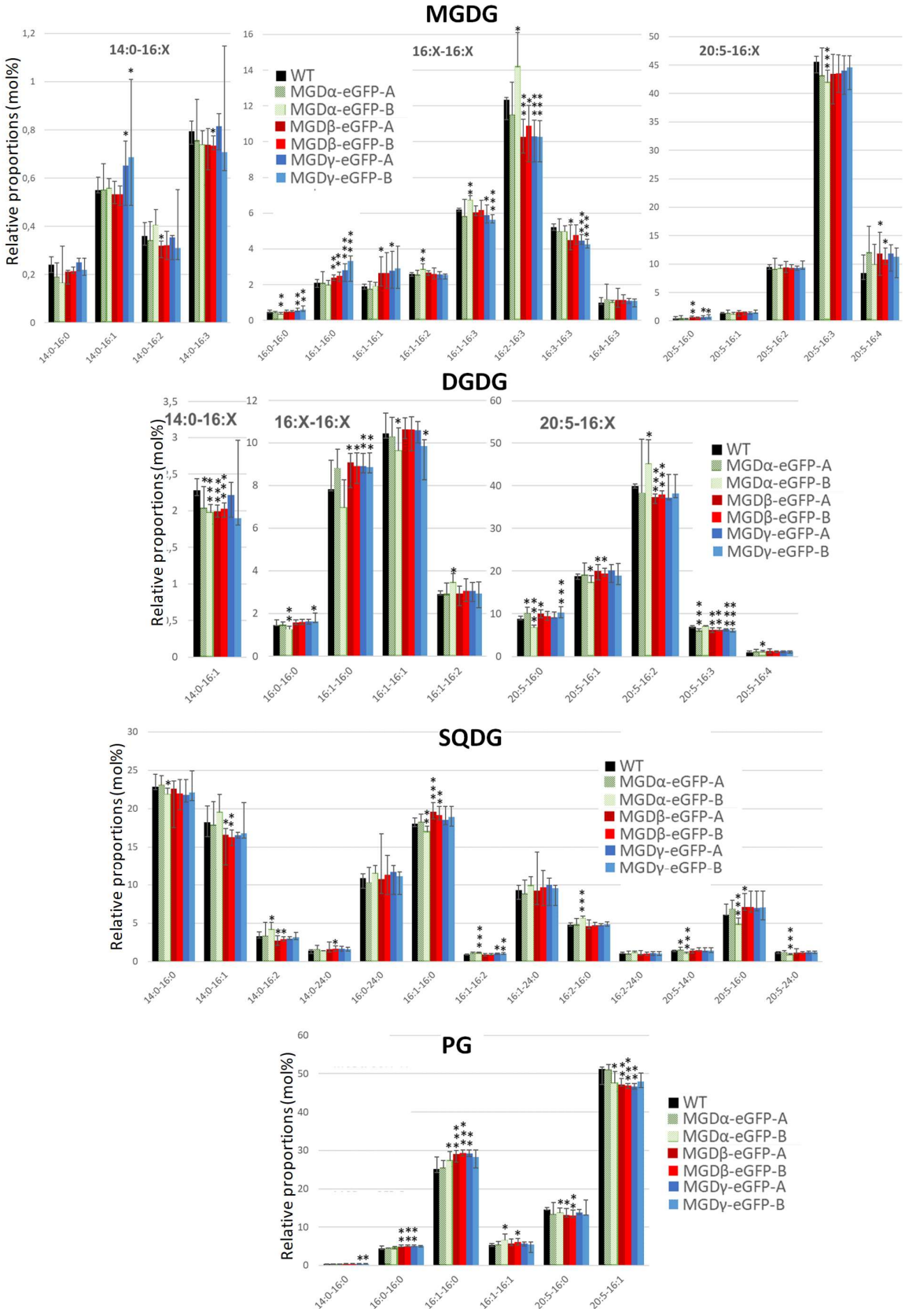
Supplementary Figure S14: *P. tricornutum* cell ultrastructure in MGD overexpressing lines. WT (A) and MGDα (B), MGDβ (C) or MGD1γ (D) overexpressing lines were observed by scanning transmission electron microscopy. Cell ultrastructure is shown in each mutant. In D, the sample on the right has been magnified from the sample on the left. No impact could be observed at the level of membrane compartments, including plastids. G, Golgi, M, mitochondria; N, nucleus, P, plastid



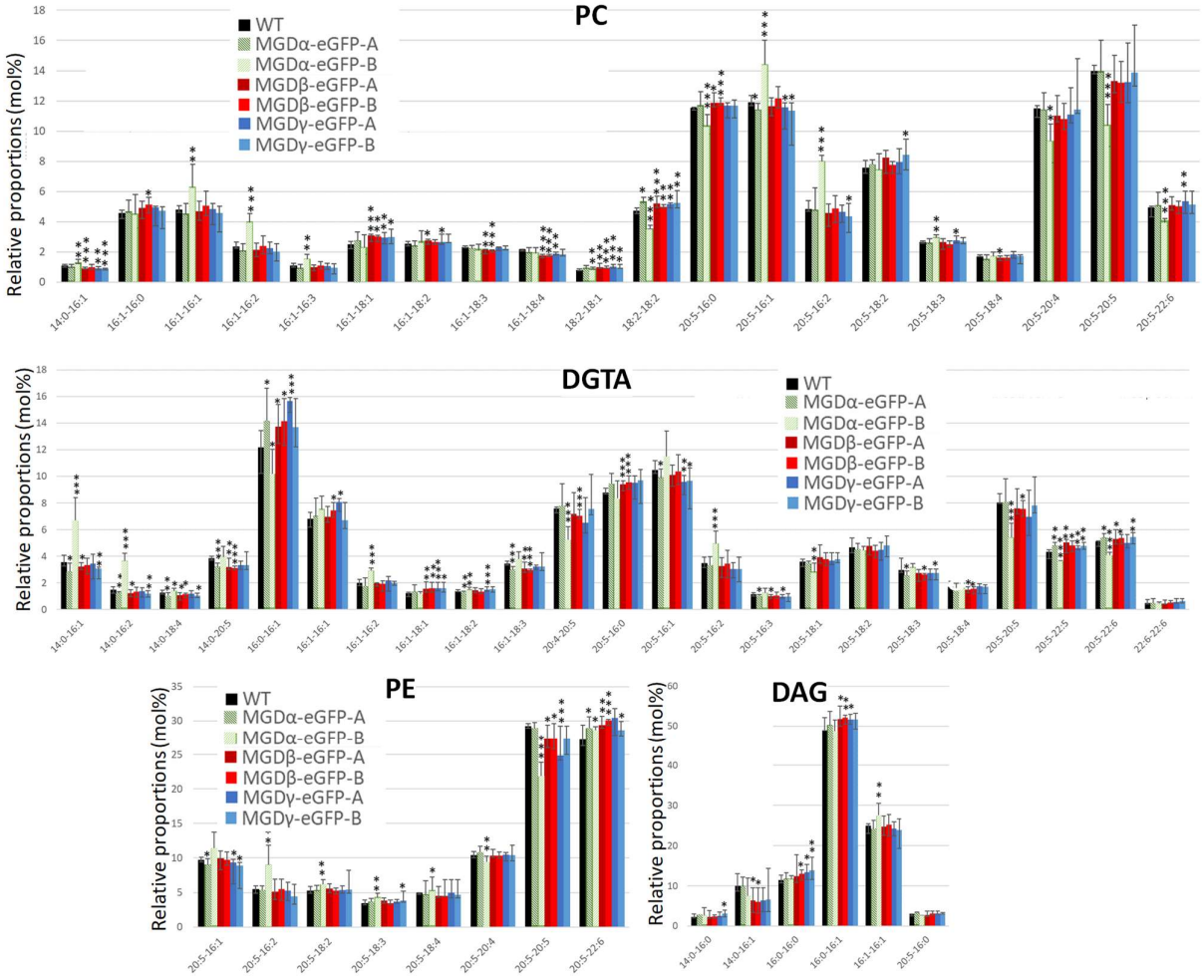
Supplementary Figure S15: Quantitative analysis of fatty acid and glycerolipid content in MGD KO mutants. Lipids from *P. tricornutum* WT and KO lines grown in 10N10P medium were extracted, and separated as described in Methods. A, total FA content (given in nmol.10^{-6} cells) and global FA profile (given in molar percentage) in total glycerolipid extracts. B, molar profile of the different glycerolipid classes in total glycerolipid extracts. The *mgdβc2* mutant contains a silent mutation and was used as a control for MGDβ knockout mutants. 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.



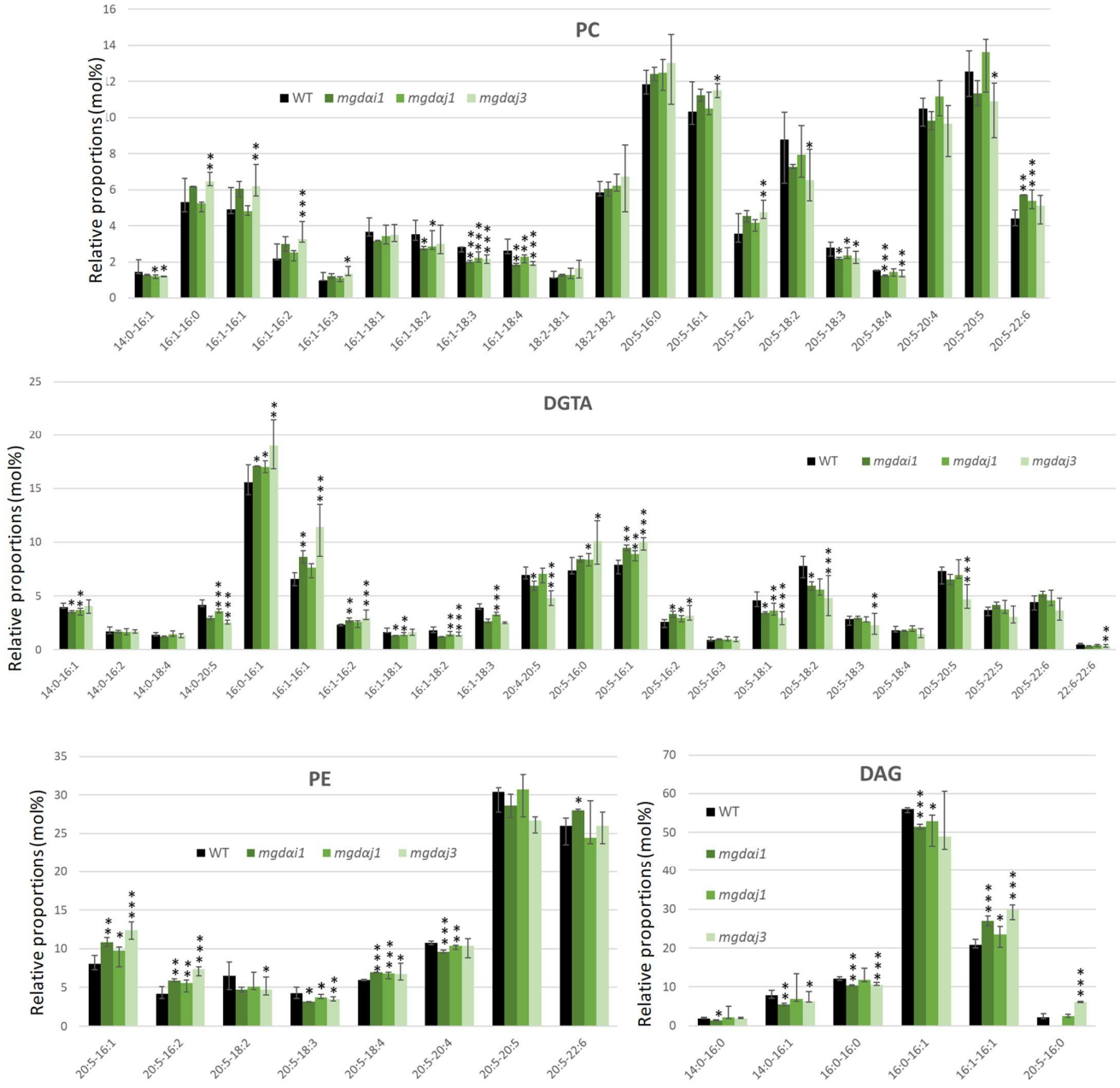
Supplementary Figure S16: Quantitative analysis of fatty acid and glycerolipid content in MGD overexpressing lines. Lipids from *P. tricornutum* WT and MGD overexpressing lines grown in 10N10P medium were extracted, and separated as described in Methods. **A**. Total FA content (given in nmol.10^{-6} cells) and global FA profile (given in molar percentage) in total glycerolipid extracts, analysed by GC-FID. **B**. Profile of the different glycerolipid classes in total glycerolipid extracts, analysed as described in Methods. 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.



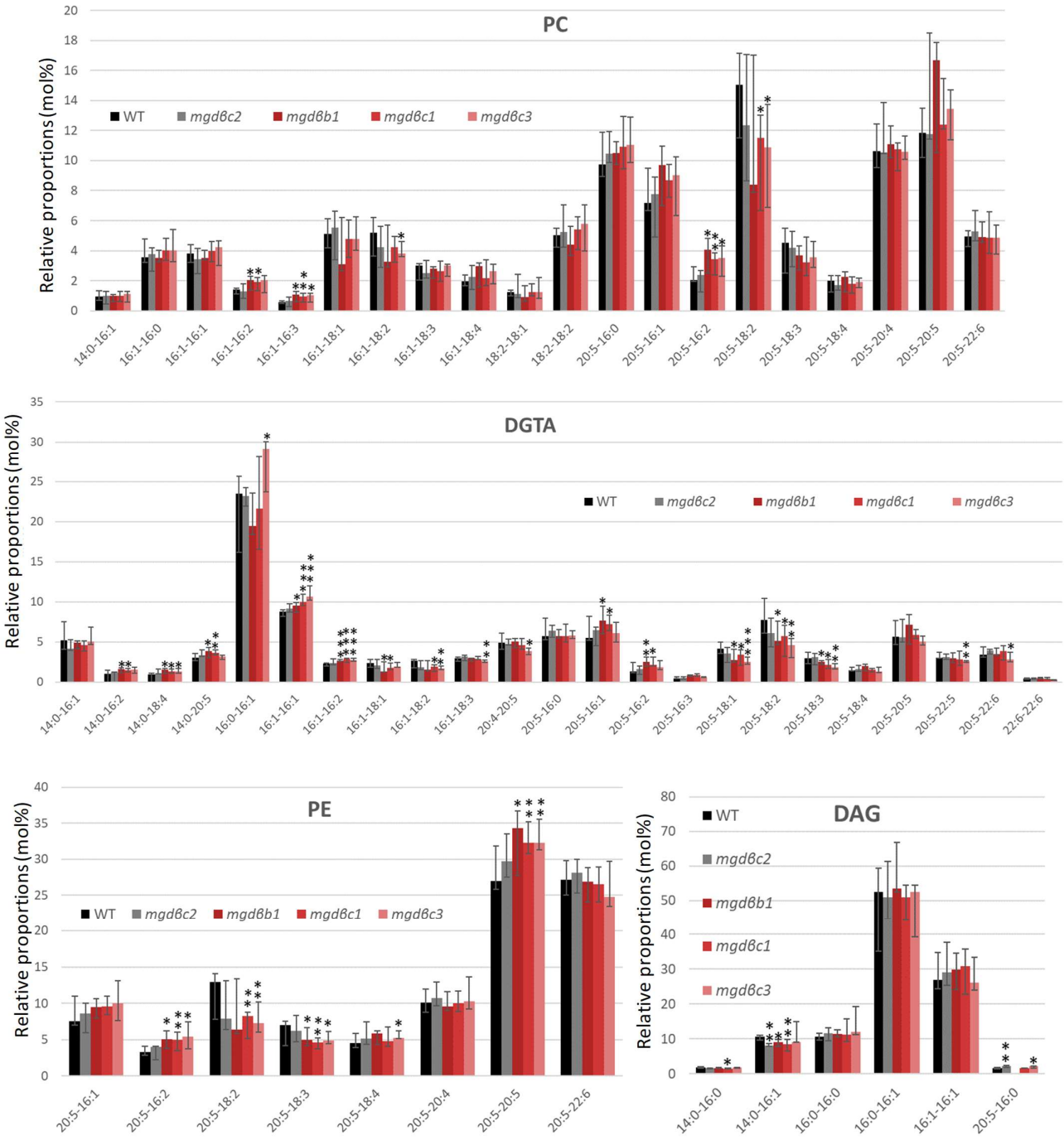
Supplementary Figure S17: Impact of the overexpression of MGD isoforms on the molecular species constituting MGDG, DGDG, SQDG and PG in *P. tricornutum*. Lipids from *P. tricornutum* WT and MGD overexpressing lines grown in 10N10P medium were extracted, and separated as described in Methods. Lipid molecular 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.



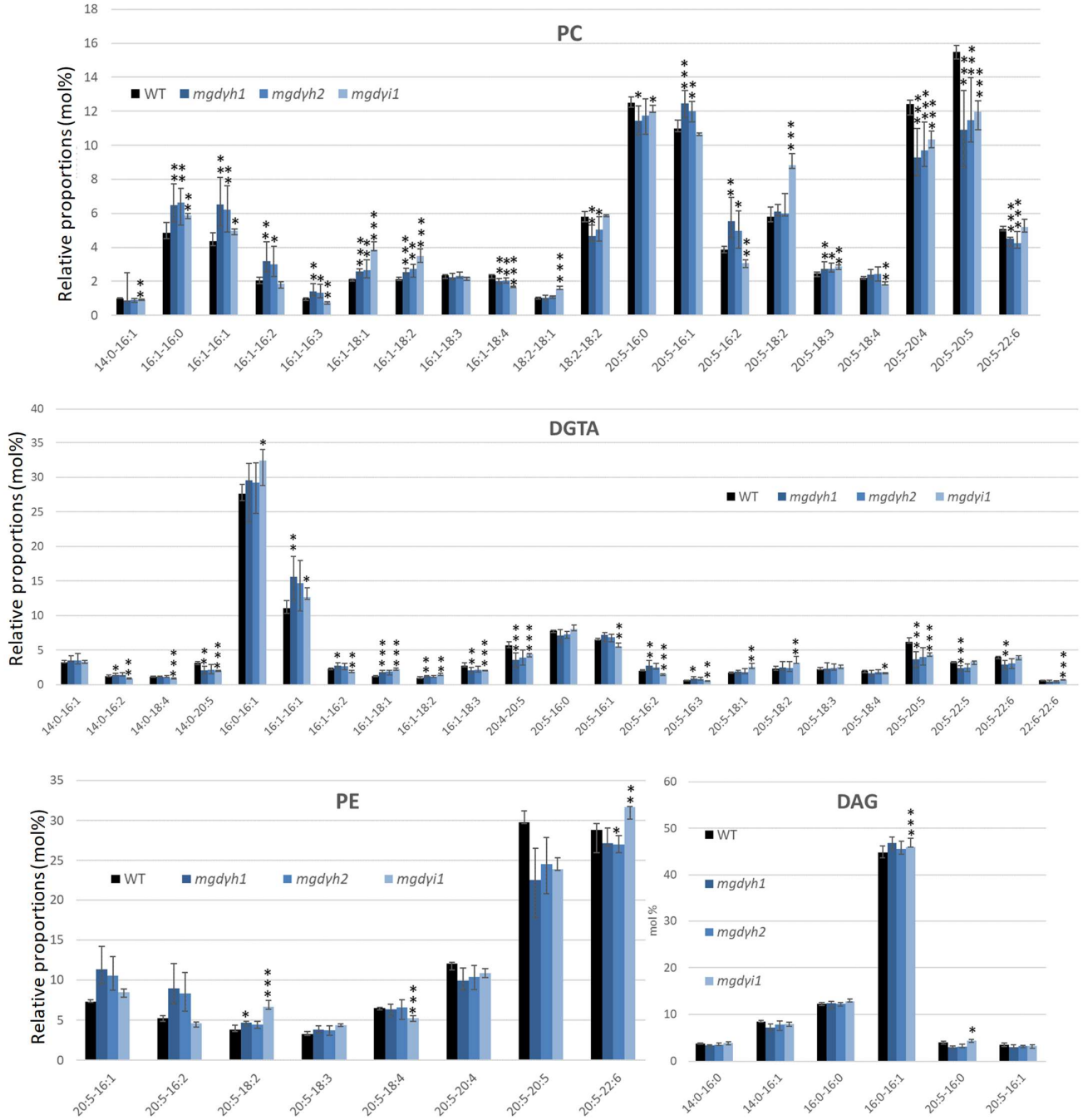
Supplementary Figure S18: Impact of the overexpression of MGD isoforms on the molecular species of endomembrane glycerolipids in *P. tricornutum*. Lipids from *P. tricornutum* WT and MGD overexpressing lines grown in 10N10P medium were extracted, and separated as described in Methods. Lipid molecular species in PC, DGTA, PE, and DAG 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.



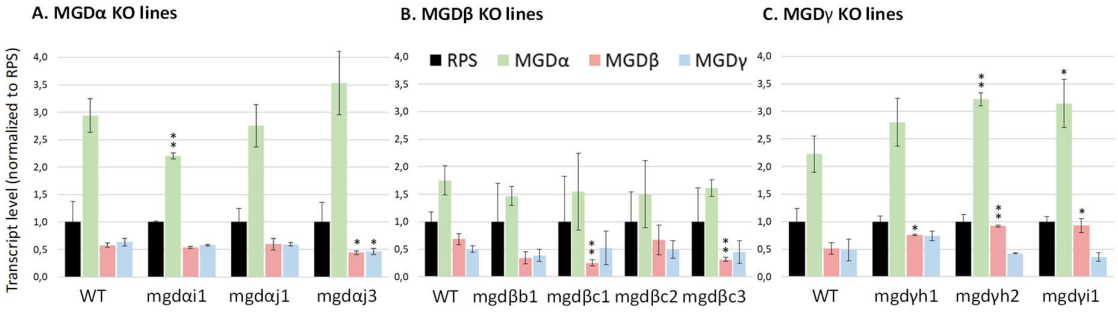
Supplementary Figure S19: Impact of MGDα mutations on the molecular species constituting PC, PE, DGTA and DAG in *P. tricornutum*. Lipids from *P. tricornutum* wild type (WT) and knocked-out (KO) lines grown in 10N10P medium were extracted, and separated as described in Methods. Lipid molecular species 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.



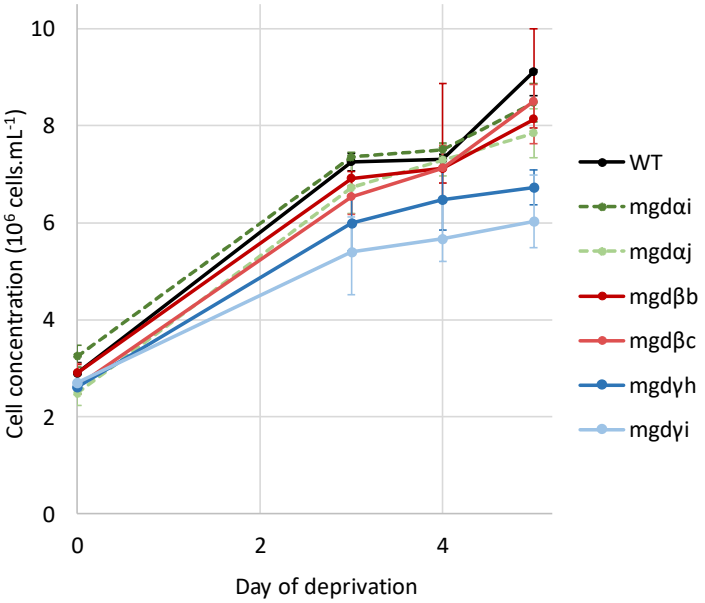
Supplementary Figure S20: Impact of MGDβ mutations on the molecular species constituting PC, PE, DGTA and DAG in *P. tricornutum*. Lipids from *P. tricornutum* wild type (WT) and knocked-out (KO) lines grown in 10N10P medium were extracted, and separated as described in Methods. Lipid molecular species 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.



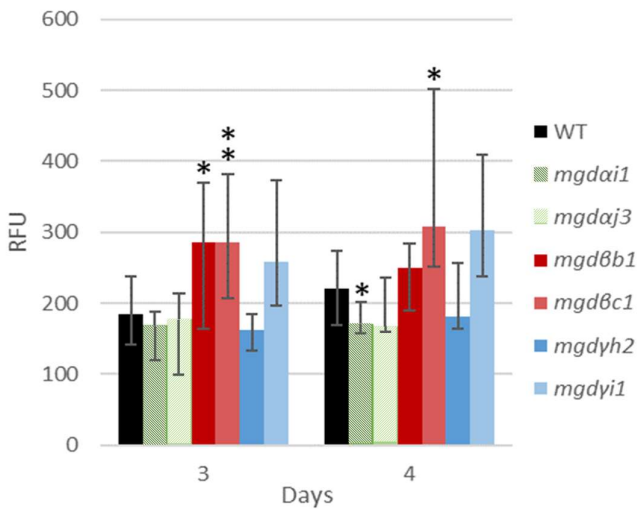
Supplementary Figure S21: Impact of MGDγ mutations on the molecular species constituting PC, PE, DGTA and DAG in *P. tricornutum*. Lipids from *P. tricornutum* WT and KO lines grown in 10N10P medium were extracted, and separated as described in Methods. Lipid molecular species were analysed by LC-MS/MS. 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.



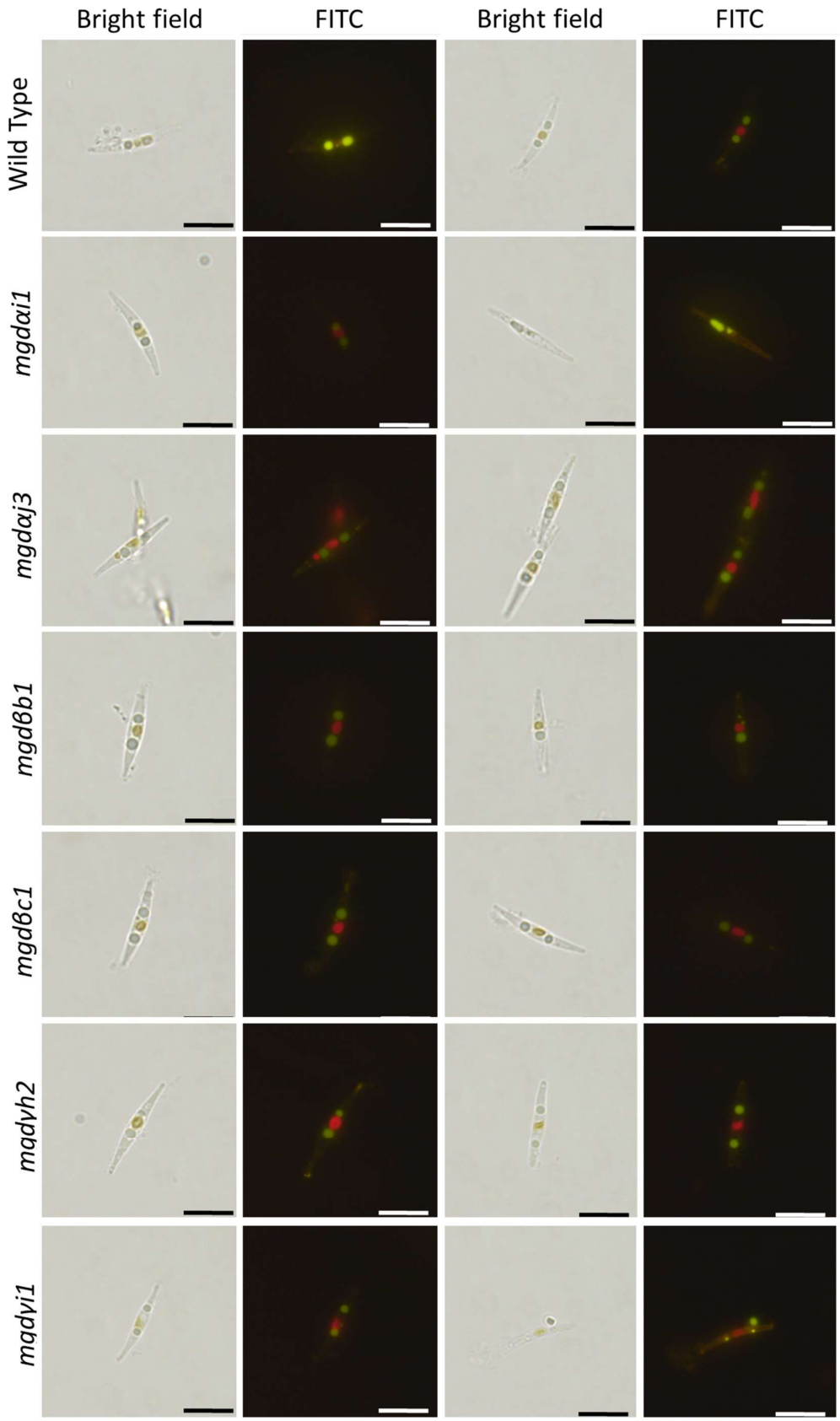
Supplementary Figure S22: *MGD* relative gene expression in mutant lines compared to the WT, grown three days in 10N10P medium. Each graph represents the relative expression of each *MGD* gene in the mutant lines of *MGDα* (A), *MGDβ* (B), and *MGDγ* (C), compared to a WT grown in parallel, normalised to *RPS* housekeeping gene expression (Supplementary Methods). Each result is the average of six biological replicates ± standard error. (*), *P-value* < 5.10^{-2}; (**), *P-value* < 1.10^{-2}; (***), *P-value* < 1.10^{-3}, based on an unpaired multiple t test.



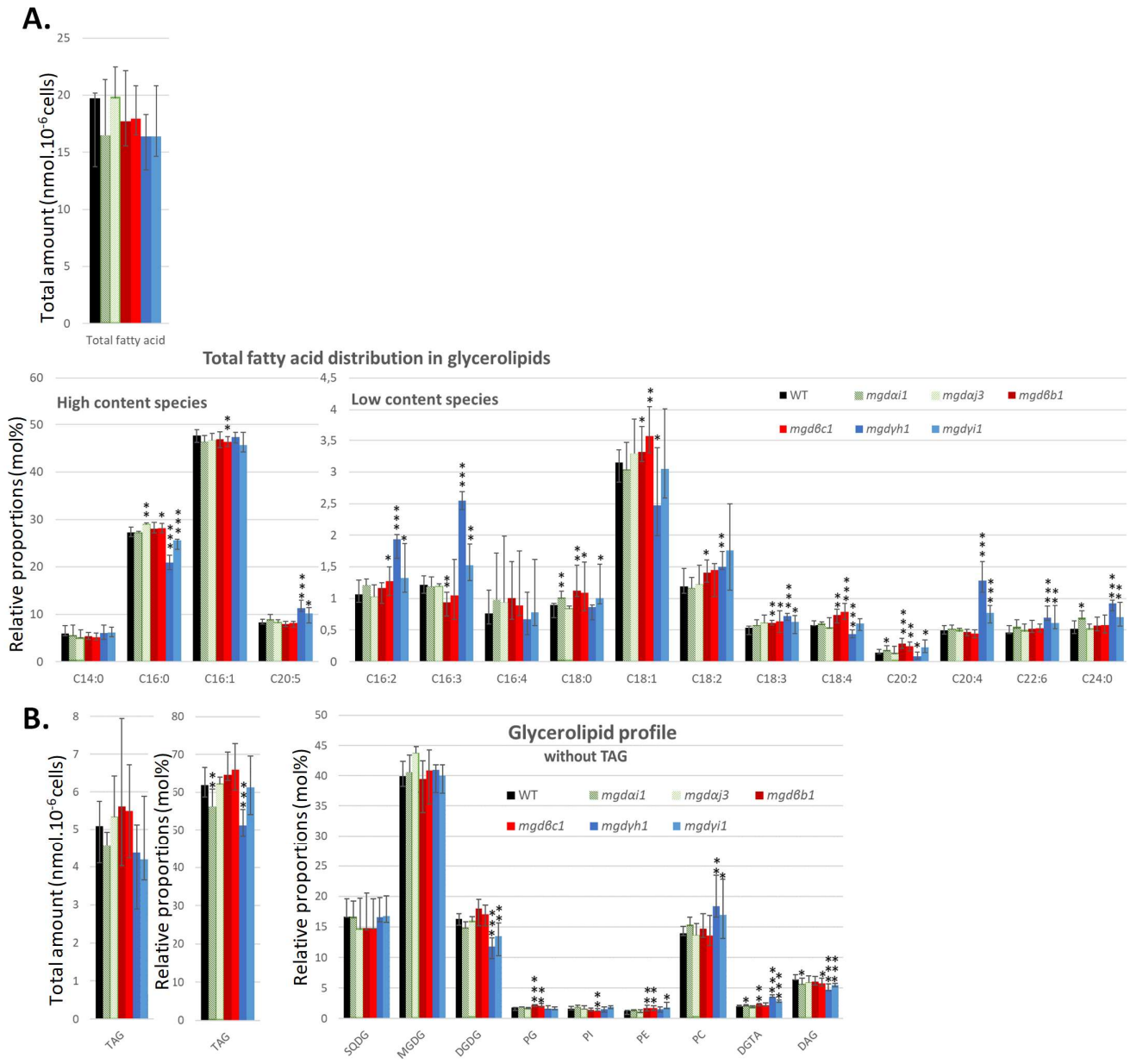
Supplementary Figure S23: Wild type and mutant *P. tricornutum* growth during nitrogen limitation. Cell concentration was monitored in mutant strains cultured at 20 °C in ESAW 0N10P medium, with a WT in parallel (10^6 cells.mL^{-1}). Data are the median of six biological replicates. Error bars represent minimal and maximal values.



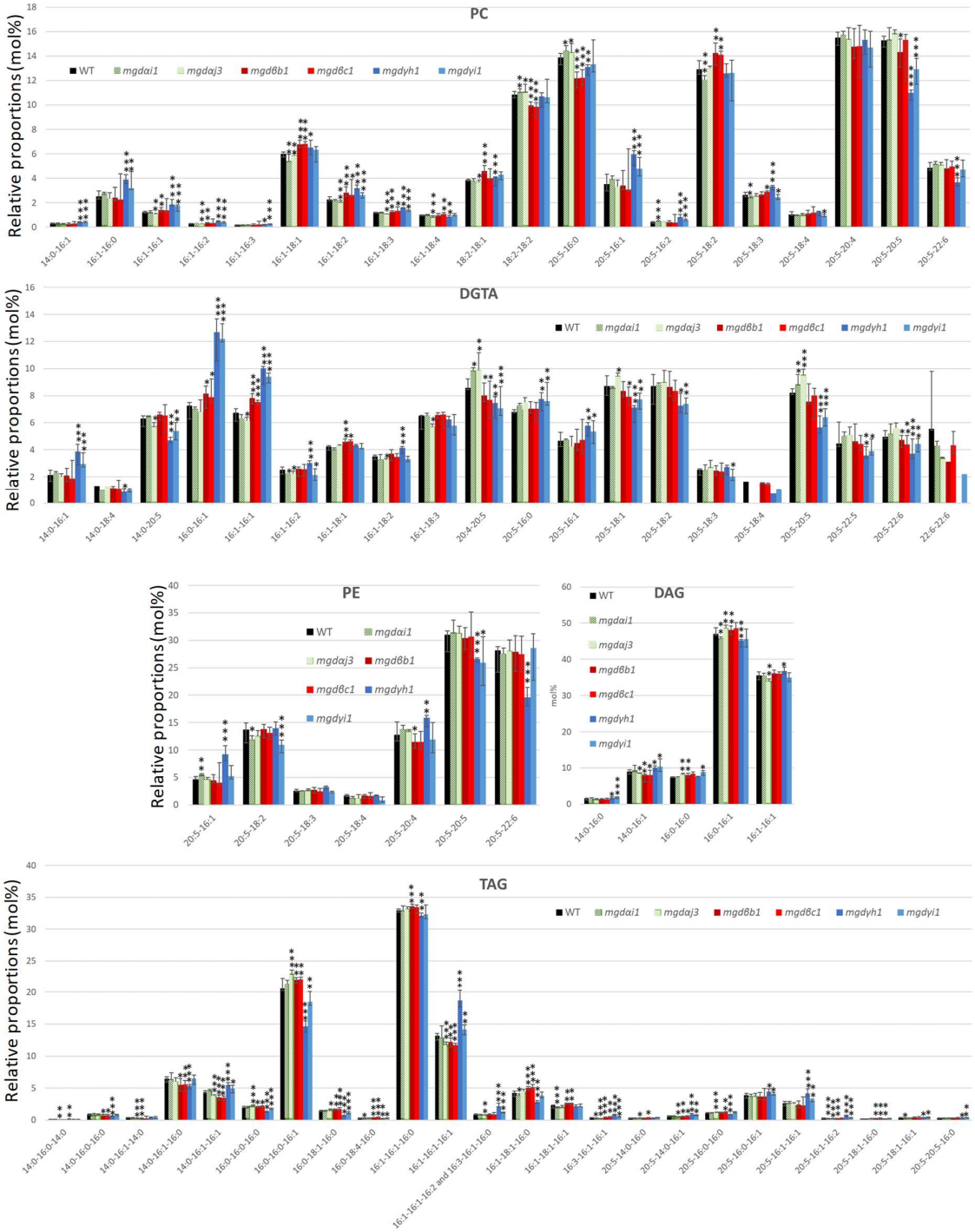
Supplementary Figure S24: Accumulation of non-polar lipids during nitrogen limitation. *P. tricornutum* WT and mutant strains were cultured in ESAW 0N10P medium. Non-polar lipid accumulation was measured after 3, and 4 days by Nile Red staining using a TECAN Infinite M1000 PRO, and expressed as fluorescence intensity normalized by cell number (RFU, relative fluorescence units). Data are the median of six biological replicates. Error bars represent minimal and maximal values. (*), *P-value* < 5.10^{-2}; (**), *P-value* < 1.10^{-2}; (***), *P-value* < 1.10^{-3}, based on an unpaired multiple t test.



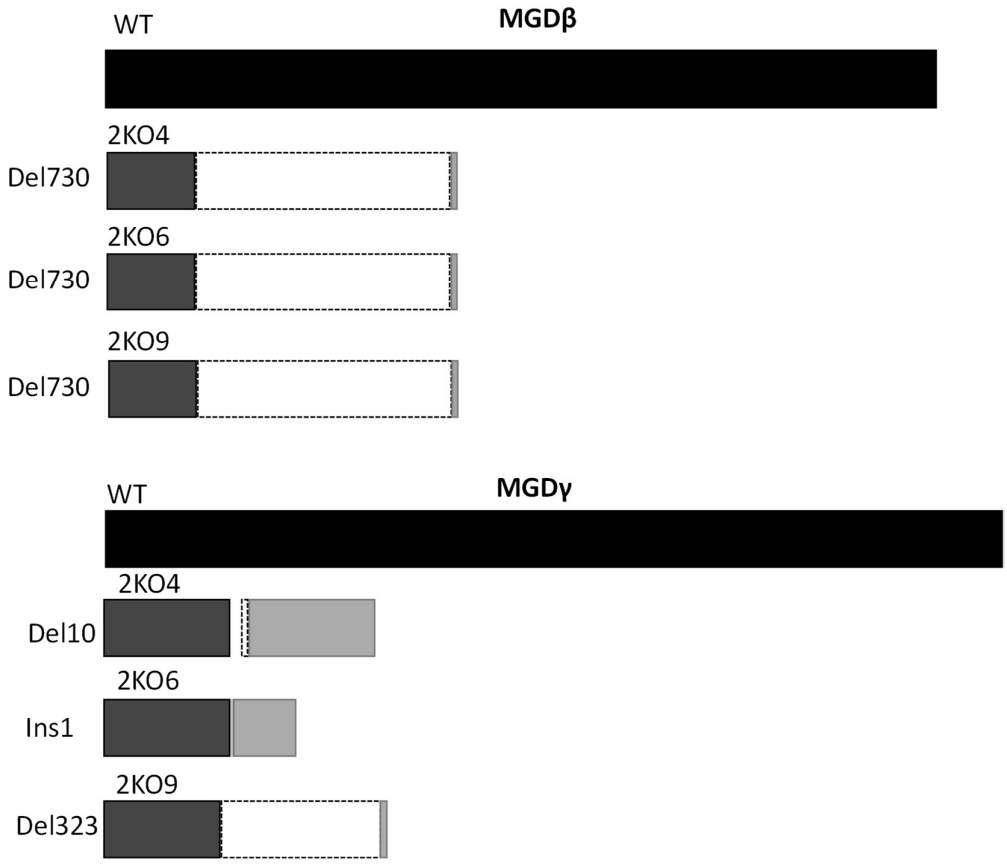
Supplementary Figure S25: Observation of cell morphology and TAG accumulation by epifluorescence microscopy. *P. tricornutum* WT and mutant strains were observed with an epifluorescence microscope after 5 days of culture in ESAW 0N10P medium. Two sets of images were taken per strain. Images were taken with an oil-immersed objective 100x. Chlorophyll autofluorescence and Nile Red staining were observed using a FITC filter. Scale bar: 10 µm.



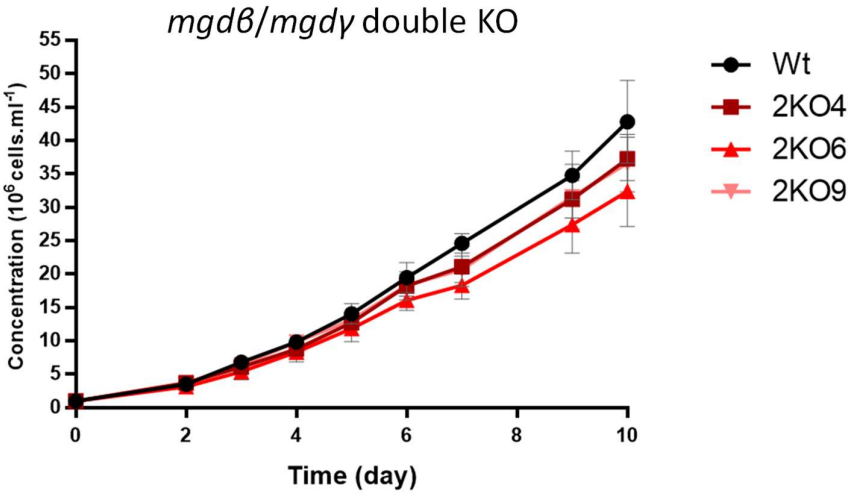
Supplementary Figure S26: Quantitative analysis of FA and glycerolipid content in MGD mutant lines upon nitrogen starvation. Lipids from *P. tricornutum* WT and KO lines grown in 10N10P medium were extracted, and separated as described in Methods. **A**, total FA content (given in nmol.10^{-6} cells) and global FA profile (given in molar percentage) in total glycerolipid extracts. **B**, TAG content (in nmol.10^{-6} cells and molar percent of total glycerolipids), and glycerolipid profile (in molar percent without including TAG). 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.



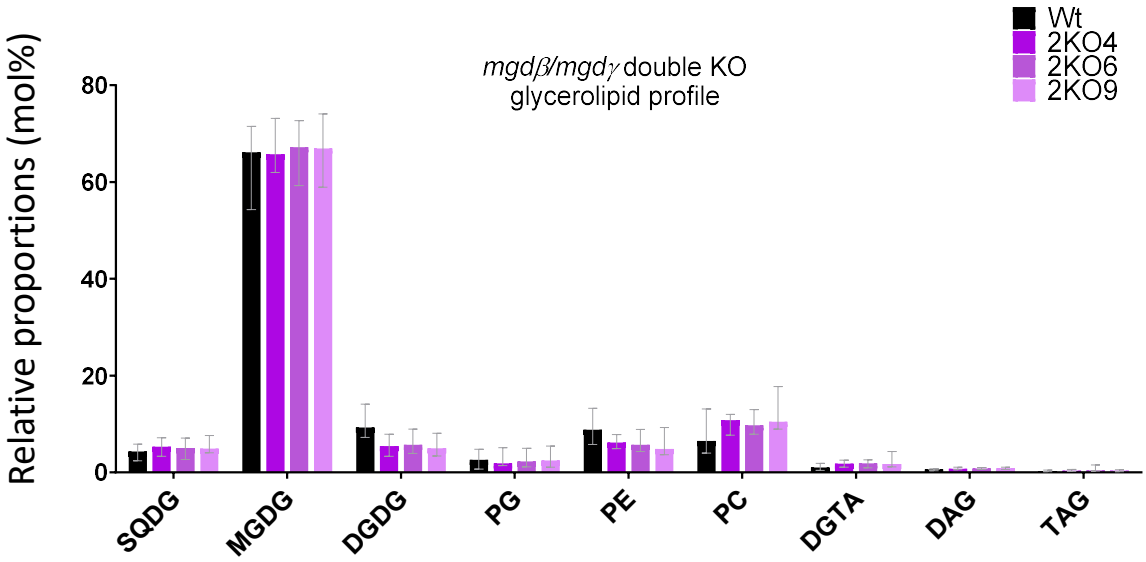
Supplementary Figure S27: Impact of MGD mutations on the molecular species constituting PC, PE, DGTA, DAG and TAG in nitrogen-deprived *P. tricornutum* cells. Lipids from *P. tricornutum* wild type (WT) and knocked-out (KO) lines grown in 10N10P medium were extracted, and separated as described in Methods. Lipid molecular species in PC, DGTA, PE, DAG and TAG 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. DAG, diacylglycerol; DGTA, diacylglyceryl hydroxymethyltrimethyl-β-alanine; PC, phosphatidylcholine; PE, phosphatidylcholine; TAG, triacylglycerol.



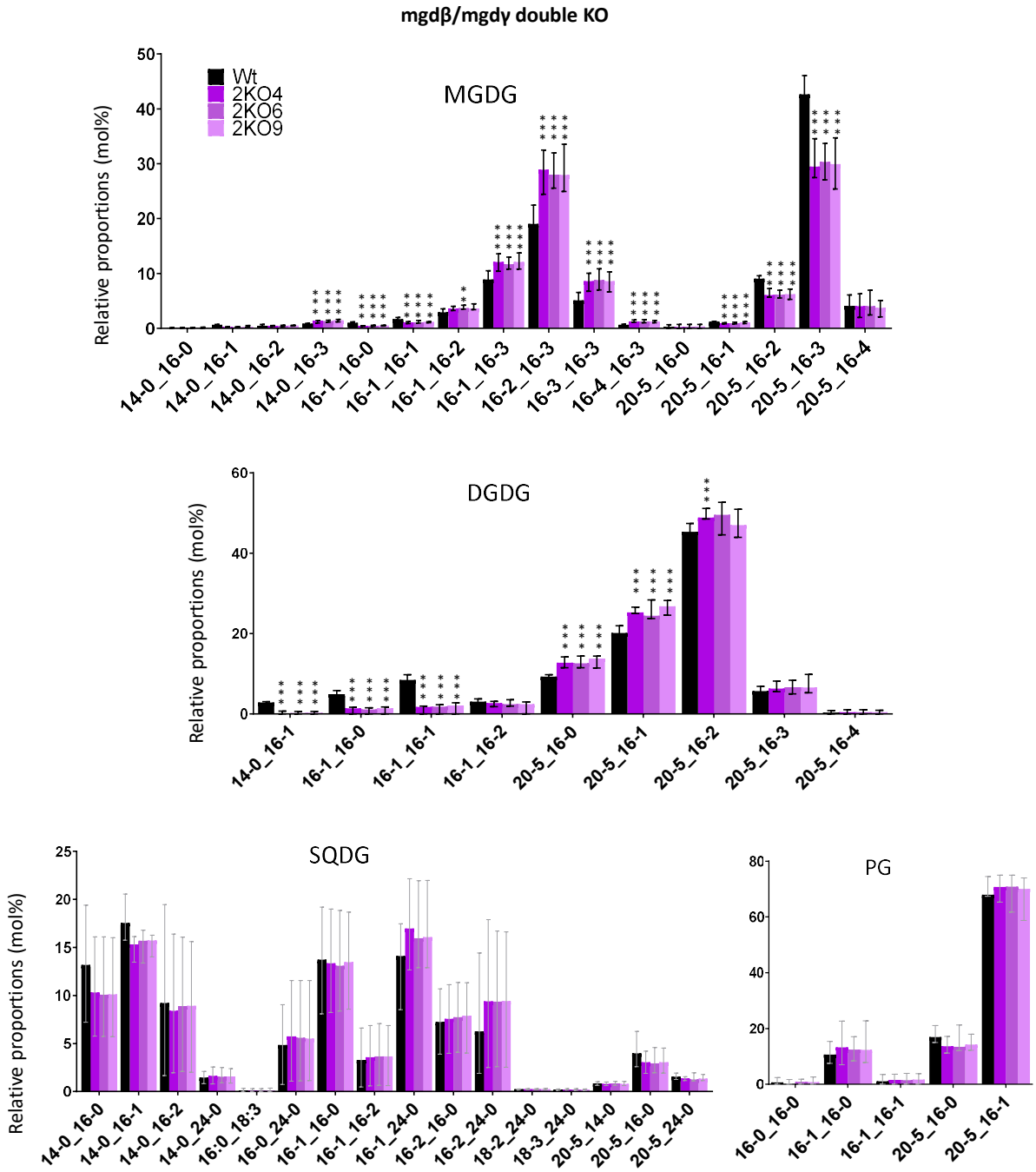
Supplementary Figure S28: Selected MGDβ/MGDγ double knock-out mutants, generated by CRISPR-Cas9 editing. Schematic representation of the impact of the indels introduced on the protein sequences in each mutant lines. Black bars, WT full-length sequence; dark gray, portion of mutant proteins identical to the WT N-terminal part; light gray, portion of mutant proteins differing from the WT C-terminal part. Ins, insertion; Del, deletion, KO, knock-out; WT, wild-type.



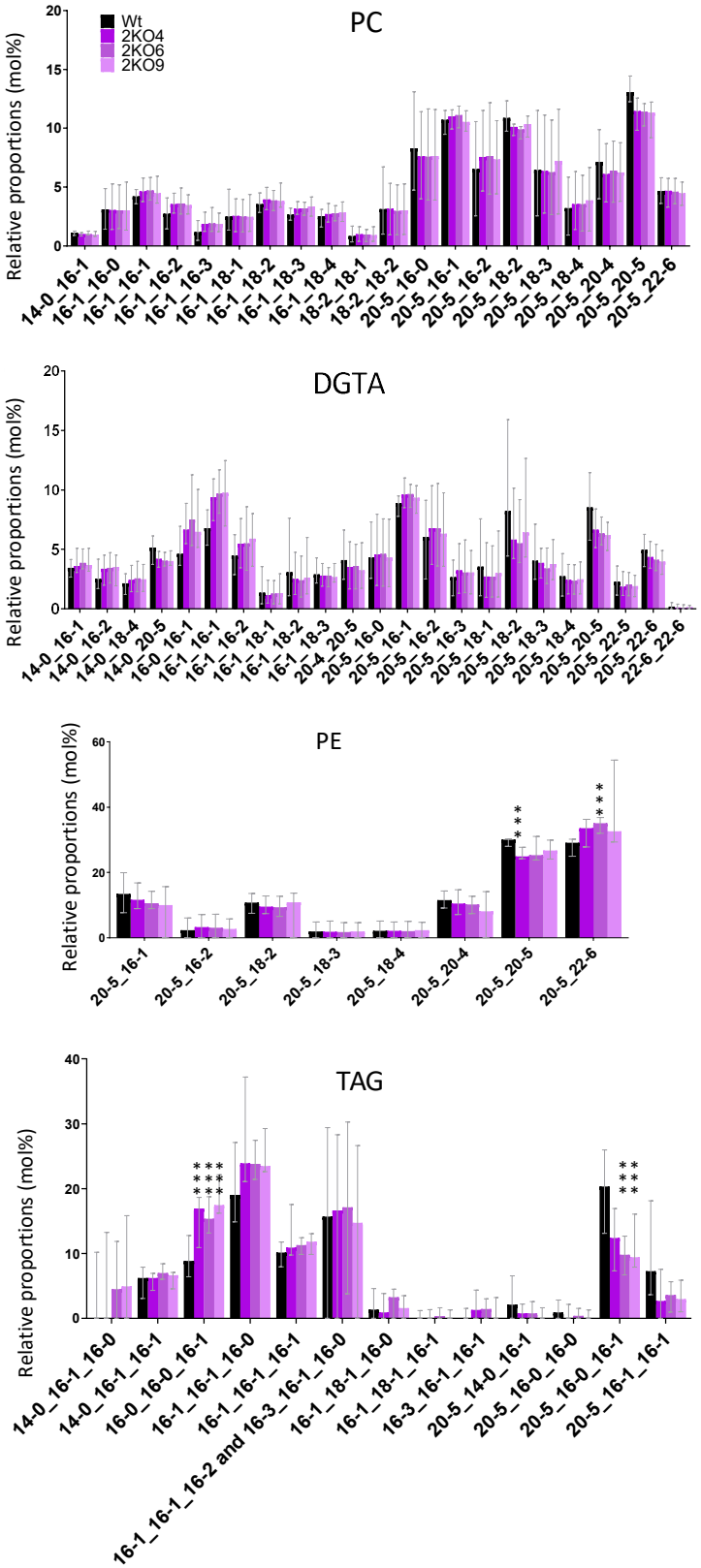
Supplementary Figure S29: Growth curves of *MGDβ*/*MGDγ* double mutants compared to the WT. Mutant (KO) and wild-type (WT) lines were cultivated in parallel at 20°C in 100-mL ESAW 10N10P medium. Cell concentration (10^6 cells.mL^{-1}) was measured every day as indicated, using a TECAN Infinite M1000 PRO. Data are the average of three independent biological replicates ± SD.



Supplementary Figure S30: Quantitative analysis of glycerolipid content in *MGDβ*/*MGDγ* double mutant overexpressing lines. Lipids from *P. tricornutum* wild-type (WT) and double knock-out (KO) lines grown in 10N10P medium were extracted, and separated as described in Methods. Lipids from *P. tricornutum* mutant and the WT lines grown in 10N10P medium were extracted, and separated as described in Methods. Each result is the median of six biological replicates ± min and max values. DAG, diacylglycerol; DGDG, digalactosyldiacylglycerol; DGTA, diacylglyceryl hydroxymethyltrimethyl-β-alanine; MGDG, monogalactosyldiacylglycerol; PC, phosphatidylcholine; PE, phosphatidylcholine; PG, phosphatidylglycerol; SQDG, sulfoquinovosyldiacylglycerol; TAG, triacylglycerol.



Supplementary Figure S31: Impact of *MGDβ*/*MGDγ* double mutations on MGDG, DGDG, SQDG and PG molecular species. Lipids from *P. tricornutum* WT and double KO lines grown in 10N10P medium were extracted, and separated as described in Methods. Lipid molecular species were analyzed by LC-MS/MS. Each result is the median of six biological replicates ± min and max values. (***), *P-value* < 1.10^{-3}, based on an unpaired multiple t test. DGDG, digalactosyldiacylglycerol; MGDG, monogalactosyldiacylglycerol; PG, phosphatidylglycerol; SQDG, sulfoquinovosyldiacylglycerol.



Supplementary Figure S32: Impact of *MGDβ*/*MGDγ* double mutations on PC, DGTA, PE and TAG molecular species. Lipids from *P. tricornutum* WT and double KO lines grown in 10N10P medium were extracted, and separated as described in Methods. Lipid molecular species were analyzed by LC-MS/MS. Each result is the median of six biological replicates ± min and max values. (***), *P-value* < 1.10^{-3}, based on an unpaired multiple t test. DGTA, diacylglyceryl hydroxymethyltrimethyl-β-alanine; PC, phosphatidylcholine; PE, phosphatidylcholine; TAG, triacylglycerol.

Supplementary Table S1: *Phaeodactylum tricornutum* MGD entries in NCBI, EnsemblProtist and Uniprot databases. Gene and cDNA sequences were verified and coding sequences were corrected. The MGDα and MGDγ Uniprot sequences correspond to fragments. The MGDα coding sequence corresponds to the second putative start codon predicted in the Phatr3_J14125 sequence.

	NCBI (gene)	**Ensembl Protists (gene)**	**Uniprot (protein)**	**Curated full length protein sequence**
MGDα	PHATRDRAFT_14125	Phatr3_J14125 / Phatr3_EG02525	B7G3X5 (fragment)	MCKLFFLCFWLSSTAEWSAAFSPNLRPTSTRLDVAAYIRTTTAAIETDEVCTVQ ILMSDTGGGHRASANALRDAFDTLHPGRIQCDIVDIYTEYGPFWPYDSYIELYK FAAKYPITWDIFYHFGATDFGIWLNRLMLELFCFEPFKTCLSRPSGNSGKKADM VVSVHPLTQDIPLRILAELDSNGATRERTGRKTPFCTVVTDLGSAHPTWFNKDV DKCFVPSDALYLAAKKRQLQDSQIVQYGLPIRQGFWANSESAHVAPEKVRKSLR RQLGLDENLPTVLIVGGGDGMGGIVEISKSLGVALGTASTTTQMVVVCGNNQEA KASLEKESWGTTVRVNVQGFVENMDEWMKASDALVTKAGPGTIAEASICGLPCM LFSYLPGQEEGNIPFVEEAGFGKYSGDASVIANTVSSWLLSPEKLEAMRNAALA AARPQATLNIAKDLADAVFAEKKKKLGDAGAKLTRTS
MGDβ	PHATR_43938	Phatr3_J54168	B5Y4U3 (full length)	MVWSRSPSRGRIWLATFMALTSMTGSGFVLLTNAHTGPHHHQPRGGSNTSFRQI SGPPSATSTSVSAPADDGAEEIANNICTDETPSRPIPLQATNEAVASAKMPKNQ GALKVLFLSADTGGGHRASAESLAKQFLIHYPGSTYDLLDVWTEDGVYPYKTLV ESYKHLSAHPQQWKMLYHLSNTRPWEVLMDWHSAFMCEAKIRARIASYNPDVIV SVHPAMQYVPMKSVRHLSRERGRHIPFFTVVTDLGSGHCTWFQKHPDKIYIASE RIRRLTKRRGGTEDCKIVSTGLPIRHDFAVHSKAMGDRTTPSGQAYVQKMKLDL GLPGDKPMVLLMGGGEGVGSLSEIVEQVYRSLVSEGVDATICVVCGRNENLRLS LEQRDWDAVLEARPKFSKRRFFSRILWRRRRSRRLQESLDRAEAYQHDRPDLVN ARATVDVIGLGFVTRMAEYMVAADVLVTKAGPGSIAEAASVGLPIMLTSFLPGQ EAGNVDFVLDAGFGDYNGDPVEIAQELTIWLKDRKLLVAMSKSAQGSGHPTAAE DIVLDIGETTRAWMLLNNK
MGDγ	PHATRDRAFT_9619	Phatr3_J9619	B7FQ04 (fragment)	MATGFSTQIEEDSRIEGDSIEVVSSVEVNAATVVLEQPQMKVGAAATAPTMSSN TMDEPNVSAGLEMAASSIPKSRKEKGLHVLFLSSDTGGGHRASAESLANQFQLL FPGTTYDLLDIVEKDGVAPYNSLVSTYKHLSAHPSQWKLVYTVSNSRAFEMLAD AHLKLMCERAVRKRIQSYNPDVVISVHPLMTNVPVLSCSKISHITGKHLPIFTV VTDLGSAHCLWFANGVEKMFVGSDQIKKLAMARGKVPVEKIILAGLPIRHDFAI QADLLGVRHSEAGRAYQQRVRRELKLPCTDRKTVLVMGGGEGVGSLSNIVDALY VELALQGIDALVLVVCGRNEKLRHKLATRDWQVVFDRWNGALERNGISSSLMSM SFSDACGNGLVTSGCIESRAGTVTSSIRRMLSSKNLQNAVSTPLPDSDHKRTAD SVGEEKKSEVDLNQSIPSDTTIQSPVHDGYESKSLGKVIVTGLGFVTRMAEYMV AADVLVSKAGPGTISEAAAVSLPVMLTSFLPGQEEGNVDYVIDGGFGAYCADTD PIGIGEEVCMWLHDPAKLEMLSNAAKAKGVPNAARDIAQQIGDSTMKWKELNDA THDQSTTELTGHTKIVVSAEEPSR

Supplementary Table S2: Taxonomic classification of the MGD proteins used for phylogenetic analyses. Taxonomic classification according to Cavalier-Smith's system (Turland et al., 2018; Burki et al., 2020). All selected proteins were predicted to be MGD proteins based on similarities with the MGD isoforms from *A. thaliana* and *P. tricornutum* and manual curation.

Kingdom	Division or phylum	Class	Order	Family	Genus and Species	Protein idenfifier
Protozoa	Alveolata	Colpodellida	Chromerida	Vitrellaceae	*Vitrella brassicaformis*	CEM14063.1
						CEM34604.1
		Conoidasida (1)	Eugregarinorida	Gregarinidae	*Gregarina niphandrodes*	XP_011130274.5
Chromista	Stramenopiles	Bacillariophyceae (2)	Bacillariales	Bacillariaceae	*Pseudo-nitzschia multistriata*	VEU37191.1
						VEU39505.1
					Fragilariopsis cylindrus	OEU16730.1
						OEU06991.1
			Naviculales	Stauroneidaceae	*Fistulifera solaris*	GAX23913.1
						GAX10983.1
						GAX09439.1
				Phaeodactylaceae	*Phaeodactylum tricornutum*	XP_002186355
						XP_002176800
						XP_002181685
		Mediophyceae (2)	Thalassiosirales	Thalassiosiraceae	*Thalassiosira pseudonana*	XP_002295865
						XP_002293576.1
						XP_002294242
		Phaeophyceae	Ectocarpales	Ectocarpaceae	*Ectocarpus siliculosus*	CBJ28381.1
						CBJ28372.1
						CBN79326.1
		Pelagophyceae	Pelagomonadales	Pelagomonadaceae	*Aureococcus anophagefferens*	XP_009033523.1
						XP_009038839.1
						XP_009035780.1
		Eustigmatophyceae	Eustigmatales	Monodopsidaceae	*Microchloropsis gaditana/ Nannochloropsis gaditana*	XP_005855249.1
Plantae	Chlorophyta	Trebouxiophyceae	*incertae sedis*	Coccomyxaceae	*Coccomyxa subellipsoidea*	XP_005651388.1
			Chlorellales	Chlorellaceae	*Chlorella sorokiniana*	PRW45643.1
		Mamiellophyceae	Mamiellales	Bathycoccaceae	*Ostreococcus tauri*	OUS42062.1
		Chlorophyceae	Chlamydomonadales	Chlamydomonadaceae	*Chlamydomonas reinhardtii*	PNW74102.1
			Sphaeropleales	Selenastraceae	*Monoraphidium neglectum*	XP_013903204.1
					Raphidocelis subcapitata	GBF88428.1
	Magnoliophyta - Flowering plant	Magnoliopsida - Dicotyledons	Brassicales	Brassicaceae	*Arabidopsis thaliana*	NP_194906.1
						NP_565352.1
						NP_568394.2
			Rubiales	Rubiaceae	*Coffea arabica*	XP_027069738.1
						XP_027084801.1
			Caryophyllales	Chenopodiaceae	*Spinacia oleracea*	XP_021867203.1
						XP_021852153.1
			Nymphaeales	Nymphaeaceae	*Nymphaea colorata*	XP_031495579.1
						XP_031480803.1
			Amborellales	Amborellaceae	*Amborella trichopoda*	XP_006852865.1
						XP_006845407.1
		Liliopsida - Monocotyledons	Cyperales	Poaceae / Gramineae	*Brachypodium distachyon*	XP_010238179.1
						XP_003573675.2
						XP_003570331.1
					Oryza sativa	XP_015611851.1
						XP_015649135.1
						XP_015627712.1
	Bryophyta - Mosses	Bryopsida - True mosses	Funariales	Funariaceae	*Physcomitrium patens*	XP_024403181.1
						XP_024404076.1
	Lycopodiophyta - Lycopods	Lycopodiopsida	Selaginellales	Selaginellaceae	*Selaginella moellendorffii*	XP_024533920.1
	Rhodophyta	Cyanidiophyceae	Cyanidiales	Cyanidiaceae	*Cyanidioschyzon merolae*	XP_005536420.1
				Cyanidioschyzonaceae	*Cyanidiococcus yangmingshanensis*	KAF6002292.1
		Florideophyceae	Gracilariales	Gracilariaceae	*Gracilariopsis chorda*	PXF49423.1
						PXF42956.1
						PXF46603.1
Bacteria	Proteobacteria	Alphaproteobacteria	Hyphomicrobiales	Blastochloridaceae	*Blastochloris viridis*	WP_055036643.1

(1): From Apicomplexa clade
(2): From the diatom group

Supplementary Table S3: **Single guide RNA list**. Single guide RNAs (sgRNA) used for the acquisition of *P. tricornutum* mutants with CRISPR-Cas9. Orientation of the sgRNAs is given taking the target gene orientation as a reference. Forward and reverse sequences correspond to the oligonucleotide sequences ordered for sgRNA insert preparation. Short nucleotide sequences added for cloning purposes are underlined in the sequences. The first possibly variable nucleotide of the associated PAM (protospacer adjacent motif) is written in bold.

Target Gene	Name	Strand	Forward and reverse sequences	PAM sequence	Primers used for sequencing
MGDα	125a	+	TCGACAATTCAACAACCGTGCTCA AAACTGAGCACGGTTGTTGAATTG	**T**GG	125-Fwd-138 & 125-Rev+210
	125i	+	TCGACAGATACTCATGTCCGACAC AAACGTGTCGGACATGAGTATCTG	**T**GG	125-Fwd-138 & 125-Rev+1519
	125j	+	TCGACGACACTGGCGGGGGTCACA AAACTGTGACCCCCGCCAGTGTCG	**G**GG	125-Fwd+205 & 125-Rev+591
MGDβ	168b	+	TCGAGTCCCGAGGACGTATATGGT AAACACCATATACGTCCTCGGGAC	**T**GG	168-Fwd-131 & 168-Rev+359
	168c	-	TCGAGAAGTCGAAGTTGCCGATGG AAACCCATCGGCAACTTCGACTTC	**T**GG	168-Fwd-131 & 168-Rev+359
MGDγ	619a	+	TCGAAGGTTTCTCAACCCAGATAG AAACCTATCTGGGTTGAGAAACCT	**A**GG	619-Fwd-132 & 619-Rev+267
	619c	-	TCGAGGCTGCATTGACTTCAACGG AAACCCGTTGAAGTCAATGCAGCC	**A**GG	619-Fwd-132 & 619-Rev+267
	619h	+	TCGACTTTTCCTCTCGTCCGACAC AAACGTGTCGGACGAGAGGAAAAG	**C**GG	619-Fwd+244 & 619-Rev+713
	619i	+	TCGACTCTCGTCCGACACCGGTGG AAACCCACCGGTGTCGGACGAGAG	**A**GG	619-Fwd+244 & 619-Rev+713

Supplementary Table S4: Primers list. Primers used for the cloning of MGD genes in expression vectors, RT-qPCR analysis of MGD expression, and PCR amplification and sequencing for transformant screening. Enzyme restriction sites are underlined when present.

Primers for constructions in pPha-CG vector

Name	Sequence	Restriction enzyme
MGDα-Fwd	GGATCCATGGACAGACCAAAGGCGTC	BamHI
MGDα-Rev	CTCGAGTGATGTCCGTGTCAATTTAGC	XhoI
MGDβ-Fwd	GGATCCATGGTGTGGTCAAGGTCACC	BamHI
MGDβ-Rev	GTCGACCTTATTGTTGAGAAGCATCC	SalI
MGDγ-Fwd	GCTAGCATGGCTACAGGTTTCTCAACC	NheI
MGDγ-Rev	GTCGACCCGACTTGGTTCTTCGGCG	SalI

Primers for constructions in PIVEX 2.3d vector

Name	Sequence	Restriction enzyme
MGDα-Fwd-NcoI	TACATACCATGGGTTTGCGACCCACGTCA	NcoI
MGDα-Rev-SmaI	TTCAGACCCGGGTGATGTCCGTGTCAATTTAG	SmaI
MGDβ-Fwd-NdeI	GGAATTCCATATGAACGAGGCCGTTGCGAG	NdeI
MGDβ-Rev-SmaI	TTCAGACCCGGGCTTATTGTTGAGAAGCATCC	SmaI
MGDγ-Fwd-NdeI	GGAATTCCATATGGCTGGTTTAGAAATGGCG	NdeI
MGDγ-Rev-SacI	TTCAGAGAGCTCTCCGACTTGGTTCTTCGG	SacI

Primers for Quantitative Real Time PCR

Name	Sequence	Name	Sequence
168-qPCR1-Fwd	TCCTCATTCACTATCCGGGC	619-qPCR5-Fwd	CAGGACTCGGTTTCGTCACA
168-qPCR1-Rev	CTGCGGATGTGCTGATAGGT	619-qPCR5-Rev	AGCATGACGGGAAGAGACAC
168-qPCR2-Fwd	ACCTATCAGCACATCCGCAG	125-qPCR1-Fwd	GGCTTCGTCGAGAACATGGA
168-qPCR2-Rev	GCCCTAATTTTGGCTTCGCA	125-qPCR1-Rev	GCCGGGGAGATACGAAAAGA
168-qPCR3-Fwd	CTGTCGTGACGGATTTGGGA	125-qPCR1bis-Rev	GGCCGGGGAGATACGAAAAG
168-qPCR3-Rev	TCGTTTGGTAAGGCGTCGAA	125-qPCR2-Fwd	TCGTCGAAGAGGCTGGTTTT
168-qPCR4-Fwd	CGGGCTTTGGGGACTACAAT	125-qPCR2-Rev	CTAGCGCAGCATTTCGCATT
168-qPCR4-Rev	CCTTGTGCGCTTTTGCTCAT	125-qPCR3-Fwd	TCTTTTCGTATCTCCCCGGC
168-qPCR5-Fwd	GCGCATGGCTGAATACATGG	125-qPCR3-Rev	CAGCCACGAGCTCACAGTAT
168-qPCR5-Rev	AGTCAACGTTGCCAGCTTCT	125-qPCR4-Fwd	ACGGATATTGGCGGAACTGG
619-qPCR1-Fwd	AGCAGCTACCGCTCCTACA	125-qPCR4-Rev	CACTCCCGAGATCAGTGACG
619-qPCR1bis-Fwd	TGGATGAGCCTAATGTATCGGC	125-qPCR5-Fwd	CGGCAGCGTTCTCACCTAAT
619-qPCR1-Rev	GGTGTCGGACGAGAGGAAAA	125-qPCR5-Rev	TACTGCCGCTATCGAAACCG
619-qPCR2-Fwd	GGCGGCATCAAGTATACCGAA	TUBA-Fwd	CTGGGAGCTTTACTGCTTGGA
619-qPCR2-Rev	AAAGCTGGAACTGATTGGCG	TUBA-Rev	ATGGCTCGAGATCGACGTAAA
619-qPCR3-Fwd	ATTCAAGCCGACCTTCTCGG	RPS-Fwd	AATTCCTCGAAGTCAACCAGG
619-qPCR3-Rev	TCGCCACCTCCCATTACCA	RPS-Rev	GTGCAAGAGACCGGACATAC
619-qPCR4-Fwd	GTCTCTTCCCGTCATGCTTACA	HPRT-Fwd	AGCTTGGAAGTGTGTACTCTC
619-qPCR4-Rev	CCACATGCAAACTTCCTCGC	HPRT-Rev	TGAAGGTTACTCCTGTCGAAG

Primers for transformant screening

Name	Sequence	Name	Sequence
125-Fwd-138	GTGCCGAGGACGAATATGGT	168-Rev+359	GTGAGTACAGCACGAAAAGGT
125-Rev+210	AACGAGTGTACAGGCGACAG	619-Fwd-132	GCTCTCGTCGATGTATTGCC
125-Rev+1519	TCTAGGGCCTCTGAACGGTG	619-Rev+267	ACGAAACAAATAAAACCCGGAG
125-Fwd+205	TCGTTGTCATGTGCAAACTGT	619-Fwd+244	CCTCCGGGTTTTATTTGTTTCG
125-Rev+591	GGTTGCGCCGAAATGATAAA	619-Rev+713	AAACGCCCGGGAATTCGA
168-Fwd-131	CTACCCACTTCGGCATTGAT		

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

To quantify *MGD* expression levels in *P. tricornutum* WT and mutant lines, RT-qPCR were performed after reverse transcription of extracted RNA of biological triplicates, using SuperScript IV VILO Master Mix with ezDNase Enzyme (ThermoFisher Scientific), following manufacturer's instructions. One RT-qPCR primer pair was used for each *MGD* gene: 125-qPCR2-Fwd/Rev for *MGDα*, 168-qPCR4-Fwd/Rev for *MGDβ*, and 619-qPCR5-Fwd/Rev for *MGDγ* (Supplementary Table 4). *RPS* (40S Ribosomal Protein), and *HPRT* (hypoxanthine guanine phosphoribosyltransferase) were used as references. RT-qPCR were performed in hard-shell green shell/white wells 96-well PCR plates (Bio-Rad) in technical triplicates. Power SYBR Green Master Mix (ThermoFisher Scientific) was used for the reaction. Incubation and fluorescence analyses were performed as follows: an amplification at 95°C for 10 min, followed by 40 cycles at 95°C for 10 sec, 60 °C for 10 sec and 72°C for 30 sec. After each cycle, SYBR Green fluorescence was measured and amplifications were monitored by melting curves. Primer pair efficiency was calculated as a function of the slope of the obtained cycle thresholds (Ct) and the logarithm base 10 ($\log_{10}$) of corresponding cDNA dilutions. Ct mean among technical replicates was used to calculate relative gene expression values for each gene of interest (GoI) for each genotype. Relative gene expression values for a given GoI were calculated in each genotype for each biological replicate as:

$E_R = 2^{-\Delta Ct_R}$, with $\Delta Ct_R = Ct_R - Ct_{\overline{ref}}$,

where E_R is the relative gene expression calculated for one biological replicate, ΔCt_R is the difference between the Ct of the GoI in a given biological replicate and $Ct_{\overline{ref}}$, and $Ct_{\overline{ref}}$ is the mean of the three Ct of the biological triplicate for the reference gene. For graphical representation, we plotted the mean of the E_R of a GoI calculated in each biological triplicate for each genotype. Standard deviation for each biological triplicate for each GoI in each genotype was calculated as:

$$SD(\overline{E_R}) = SD(\overline{\Delta Ct_R}) * \left|\frac{\partial \overline{E_R}}{\partial \Delta Ct_R}\right| = SD(\overline{\Delta Ct_R}) * ln2 * 2^{-\overline{\Delta Ct_R}},$$

where *SD* is the standard deviation, $\overline{E_R}$ is the mean of the E_R in a biological triplicate, and $\overline{\Delta Ct_R}$ is the mean of the ΔCt_R in a biological triplicate.

Heterologous expression in yeast

MGD isoforms were expressed in *Saccharomyces cerevisiae* strain BY4741 under the control of the galactose inducible promoter pGAL1, following general transformation procedures described earlier (Billey et al., 2021). The WT yeast strain BY4741 was used as negative control. Sequence of AtMGD2 from

Arabidopsis thaliana was used as positive control. In brief, homologous recombination *in vivo* assembly was done in YCplac33 vector using PCR amplified fragments of the promoter, the ADH1 terminator, the gene of interest fused in frame to a GFP. Eight colonies per construct were screened by PCR and GFP epifluorescence and three positive colonies per construct were used for further analyses. For the AtMGD2 positive control, one colony was used. The selected yeast colonies expressing the transgenes were incubated overnight in 5 ml of liquid CSM-URA (Complete Supplement Mixture lacking uracyl) medium complemented with raffinose at 30 °C and 200 rpm. Transgene expression was induced by inoculating 0.03 OD_{600} in 50 ml of liquid CSM-URA medium complemented with 2% (p/v) galactose final concentration and incubated for 24 hours. An amount of cells corresponding to 30 OD_{600} were collected by centrifugation and total lipids were extracted and analysed as detailed in Methods.

Photosynthesis activity based on chlorophyll fluorescence measurements

Fluorescence-based photosynthetic parameters were derived from imaging of chlorophyll fluorescence emission. Data were acquired with a pulse modulated amplitude fluorimeter (MAXI version of *IMAGING-PAM M-Series* equipped with an IMAG-MAX/L LED-Array Illumination unit and an IMAG-K6 camera, Heinz Walz GmbH, Germany). A 200-µL aliquot of cells cultured in 100-mL flaks with 20 mL of ESAW medium at a concentration of 4-5x10^6 cells.mL^{-1} was loaded into a 96-wells flat bottom black plate (Greiner). Cells were dark acclimated for 15 minutes before measurements. Chlorophyll fluorescence was recorded using two different protocols as indicated. Both protocols start with measurements in the dark. In the first protocol, cells were subjected to a light intensity of 700 µmol photons.m^{-2}.s^{-1} during 9 minutes 30 seconds. Then, cells were left with a low light intensity of 20 µmol photons m^{-2}.s^{-1} for 15 minutes to observe fluorescence relaxation. In the second protocol, cells were subjected to a 2-steps increase of light intensity from 55 to 335 µmol photons.m^{-2}.s^{-1}. Cells were allowed to reach steady state (time exposure of 15 minutes and then 10 minutes 30 seconds for each light intensity) before modification of the photon flux. Cells were then left with a light intensity of 20 µmol photons.m^{-2}.s^{-1} for 15 minutes for fluorescence relaxation measurements. Effective photochemical quantum yield of PSII (ΨII) was calculated as $(F'm-F_t)/F'm$, and Non-Photochemical Quenching (NPQ) was calculated as $(F'm_0-F'm)/F'm$, where F_t is the steady-state fluorescence intensity immediately prior to a saturating pulse of actinic light, while $F'm$ and $F'm_0$ are the maximum fluorescence intensities in light- and dark-acclimated cells after a pulse, respectively (Maxwell and Johnson, 2000).

Bacterial conjugation

For bacterial conjugation, an exponentially growing *mgdγh1* culture was concentrated by centrifugation at 3,000 × g for 15 minutes and suspended in 100 µL of ESAW to reach a concentration of $5 \cdot 10^8$ cells/mL per transformation. Each well of six-well plates filled with conjugation-based solid agar medium (45% ESAW, 1% agar, 5% LB) were inoculated with 5.10^7 cells. The plates were then incubated overnight at 20°C under continuous light. Before the transformation process, cargo episomes were introduced into recipient

Escherichia coli EPI300 carrying the pTA-Mob plasmid, encoding the machinery for conjugal transfer (Karas et al., 2015; Diner et al., 2016) and selected on agar plates for both the cargo and conjugation plasmids. One day prior to transformation, several random colonies from the selection plates were used to inoculate 3 mL of LB containing gentamicin for pTA-Mob and kanamycin for the episome. Following overnight incubation, the preculture was used to initiate a culture with an OD_{600} of 0.1 in 12.5 mL of LB and appropriate antibiotics. Once the OD_{600} reached 0.8–1, the culture was centrifuged at 3,000 × g for 15 minutes. The pellet was gently resuspended in 100 µL of pre-warmed SOC medium at 30 °C using a vortex at low speed. The *E. coli* suspension was dispensed into wells containing *P. tricornutum* and gently agitated to achieve homogeneous dispersion. Immediately following the addition of *E. coli* to all wells, ensuring prevention of desiccation, the 6-well plate was incubated in the dark, at 30°C for 90 minutes. Subsequent to the thermal shock, the plate was transferred to incubation conditions of 20°C under continuous light for 48 hours. After the completion of the incubation period, each well was scraped using a sterile cell spreader, harvested into a 2 mL tube, and then streaked onto selection plates containing appropriate selective media. The plates were then incubated until visible colonies formed.

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. (2021). Characterization of the Bubblegum acyl-CoA synthetase of Microchloropsis gaditana. Plant Physiol **185,** 815-835.

Burki, F., Roger, A.J., Brown, M.W., and Simpson, A.G.B. (2020). The New Tree of Eukaryotes. Trends in ecology & evolution **35,** 43-55.

Diner, R.E., Bielinski, V.A., Dupont, C.L., Allen, A.E., and Weyman, P.D. (2016). Refinement of the Diatom Episome Maintenance Sequence and Improvement of Conjugation-Based DNA Delivery Methods. Front Bioeng Biotechnol **4,** 65.

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.

Maxwell, K., and Johnson, G.N. (2000). Chlorophyll fluorescence--a practical guide. J Exp Bot **51,** 659-668.

Russo, M.T., Aiese Cigliano, R., Sanseverino, W., and Ferrante, M.I. (2018). Assessment of genomic changes in a CRISPR/Cas9 Phaeodactylum tricornutum mutant through whole genome resequencing. PeerJ **6,** e5507.

Turland, N.J., Wiersema, J.H., Barrie, F.R., Greuter, W., Hawksworth, D.L., Herendeen, P.S., Knapp, S., Kusber, W.H., Li, D.Z., Marhold, K., May, T.W., McNeill, J., Monro, A.M., Prado, J., Price, M.J., and Smith, G.F. (2018). International Code of Nomenclature for Algae, Fungi, and Plants. Regnum Veg **159,** 1-2.