the FEBS Journal

Cloning and molecular characterization of a novel acyl-CoA:diacylglycerol acyltransferase 1-like gene (*PtDGAT1*) from the diatom *Phaeodactylum tricornutum*

Freddy Guihéneuf, Stefan Leu, Aliza Zarka, Inna Khozin-Goldberg, Ilkhom Khalilov and Sammy Boussiba

Microalgal Biotechnology Laboratory, French Associates Institute for Agriculture and Biotechnology of Drylands, Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Sede Boker Campus 84990, Israel

Keywords
Bacillariophyceae; diacylglycerol acyltransferase 1-like protein; intron-retention splicing; *Phaeodactylum tricornutum*; triacylglycerol

Correspondence
S. Boussiba, Microalgal Biotechnology Laboratory, French Associates Institute for Agriculture and Biotechnology of Drylands, Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Sede Boker Campus 84990, Israel
Fax: +972 8 6596802
Tel: +972 8 6596795
E-mail: sammy@bgu.ac.il

(Received 29 December 2010, revised 5 July 2011, accepted 29 July 2011)

doi:10.1111/j.1742-4658.2011.08284.x

We have identified and isolated a cDNA encoding a novel acyl-CoA:diacylglycerol acyltransferase (DGAT)1-like protein, from the diatom microalga *Phaeodactylum tricornutum* (*PtDGAT1*). The full-length cDNA sequences of *PtDGAT1* transcripts revealed that two types of mRNA, *PtDGAT1short* and *PtDGAT1long*, were transcribed from the single *PtDGAT1* gene. *PtDGAT1short* encodes a 565 amino acid sequence that is homologous to several functionally characterized higher plant DGAT1 proteins, and 55% identical to the putative DGAT1 of the diatom *Thalassiosira pseudonana*, but shows little homology with other available putative and cloned algal DGAT sequences. PtDGAT1long lacks several catalytic domains, owing to a 63-bp nucleotide insertion in the mRNA containing a stop codon. Alternative splicing consisting of intron retention appears to regulate the amount of active DGAT1 produced, providing a possible molecular mechanism for increased triacylglycerol (TAG) biosynthesis in *P. tricornutum* under nitrogen starvation. DGAT mediates the last committed step in TAG biosynthesis, so we investigated the changes in expression levels of the two types of mRNA following nitrogen starvation inducing TAG accumulation. The abundance of both transcripts was markedly increased under nitrogen starvation, but much less so for *PtDGAT1short*. PtDGAT1 activity of *PtDGAT1short* was confirmed in a heterologous yeast transformation system by restoring DGAT activity in a *Saccharomyces cerevisiae* neutral lipid-deficient quadruple mutant strain (H1246), resulting in lipid body formation. Lipid body formation was only restored upon the expression of *PtDGAT1short*, and not of *PtDGAT1long*. The recombinant yeast appeared to display a preference for incorporating saturated C_{16} and C_{18} fatty acids into TAG.

Database
Nucleotide sequence data are available in the GenBank/EMBL/DDBJ databases under accession number HQ589265, sequence to be released November 15 2011.

Abbreviations
ARA, arachidonic acid; DAG, diacylglycerol; DGAT, acyl-CoA:diacylglycerol acyltransferase; DHA, docosahexaenoic acid; DW, dry weight; EPA, eicosapentaenoic acid; EST, expressed sequence tag; FAME, fatty acid methyl ester; NL, neutral lipid; PtDGAT1, *Phaeodactylum tricornutum* acyl-CoA:diacylglycerol acyltransferase 1; PUFA, polyunsaturated fatty acid; QRT-PCR, quantitative real-time PCR; TAG, triacylglycerol; TFA, total fatty acid; TmDGAT1, *Tropaeolum majus* acyl-CoA:diacylglycerol acyltransferase 1; TpDGAT1, *Thalassiosira pseudonana* acyl-CoA:diacylglycerol acyltransferase 1.

Introduction

Microalgae are prominent candidates for biofuel production, particularly for the production of biodiesel from triacylglycerols (TAGs). Enhancing TAG biosynthesis in microalgae by genetic engineering is thus a tempting approach for increasing the efficiency of algal biodiesel production. TAGs are synthesized *de novo* by sequential transfer of fatty acyl chains from acyl-CoA through the glycerol 3-phosphate pathway, also known as the Kennedy pathway [1]. TAGs are the principal carbon storage compounds in various organisms, including vertebrates, oilseed plants, oleaginous fungi, yeasts, and microalgae. In microalgae, TAGs are mainly accumulated under stressful conditions in extraplastidial oil bodies.

Acyl-CoA:diacylglycerol acyltransferase (DGAT) catalyzes the final and committed step in TAG biosynthesis, and has been identified as the rate-limiting enzyme for oil accumulation in some oil crops [2–8]. Acyl-CoA-independent pathways of TAG formation that may contribute markedly to TAG formation are also known, such as phospholipid:diacylglycerol acyltransferase, which utilizes phosphatidylcholine as the acyl donor [9–11].

Three types of DGAT, types 1, 2, and 3, can take part in the acyl-CoA-dependent formation of TAGs [8]. Two major isoforms, encoded by *DGAT1* and *DGAT2*, have been identified as distinct membrane proteins that are responsible for the bulk of TAG synthesis in most organisms. The role of the third cytosolic DGAT3 awaits clarification [12].

DGAT1s (EC 2.3.1.20) are members of the membrane-bound O-acyltransferase protein superfamily. DGAT1s are larger than DGAT2, and possess at least six transmembrane domains, as compared with the two predicted in DGAT2 [13,14]. It is proposed that these two enzymes have no redundant functions in TAG biosynthesis. DGAT1 plays a dominant role in the determination of oil accumulation and fatty acid composition of seed oils [15]. Indeed, in *Arabidopsis thaliana*, mutations in DGAT1, but not DGAT2, affected seed oil levels [15]; and overexpression of garden nasturtium (*Tropaeolum majus*) *DGAT1* (*TmDGAT1*) in *A. thaliana* and *Brassica napus* resulted in a net increase of 11–30% in seed oil content [15]. Several conserved signature motifs of functional importance, involved in acyl-CoA binding, DAG/phorbol ester binding, the putative thiolase acyl enzyme signature, and a signature motif typical of members of the sucrose-nonfermenting-related protein kinase 1 family, have been identified in DGAT1. Mutagenesis of a serine in a putative sucrose-nonfermenting-related protein kinase 1 target site in TmDGAT1 results in a 38–80% increase in DGAT1 activity, and overexpression of the mutated *TmDGAT1* in *A. thaliana* enhanced oil content by 20–50% per seed [15]. The importance of the conserved phenylalanine in the DAG/phorbol ester-binding site of DGAT1 was demonstrated in maize [16]. A high-oil variant quantitative trait locus (qHO6), affecting maize seed oil and oleic acid contents, appeared to encode the ancestral maize DGAT1 and DGAT2 containing a conserved Phe469, as in DGAT1s from other species. The allele without Phe469 encodes the less active enzyme in domesticated maize, and is the result of a more recent mutation selected by domestication or breeding. Ectopic expression of the high-oil DGAT1/DGAT2 allele containing the extra phenylalanine (Phe469) increased oil and oleic acid contents by up to 41% and 107%, respectively. The emerging role of DGAT2 orthologs appears to be more important for incorporation of unusual fatty acids in the seed storage oils of some plants [13].

It is now well documented that enhanced expression of DGAT can lead to increased oil accumulation in different plant organs, even those that do not normally store TAGs. In particular, overexpression of either *DGAT1* or *DGAT2* was shown to increase the accumulation of TAGs in potato tubers [17] and leaves [18–20]. DGATs typically exhibit a strong capacity for incorporating the most prevalent acyl moieties into the *sn*-3 position [15]. Therefore, manipulation of the expression of this gene may both increase oil content and alter fatty acid composition. Thus, the strategies of metabolic engineering that result in increased oil content in plant tissues may have the potential to increase lipid content in microalgae as well.

So far, two microalgal *DGAT*s have been cloned and characterized on the basis of genome information and annotation from the diatom *Thalassiosira pseudonana* [21] and the chlorophyte *Ostreococcus tauri* [22], both belonging to the *DGAT2* family. Other putative microalgal DGAT genes have been annotated in the genomes of *Chlamydomonas reinhardtii*, *Chlorella variabilis*, and *Phaeodactylum tricornutum*.

In the present study, we report the identification and characterization of a cDNA encoding DGAT1 (*PtDGAT1*) isolated from the diatom microalga *P. tricornutum*, whose function was confirmed *in vivo* by expression in the yeast *Saccharomyces cerevisiae*. Genetic manipulations with such an enzyme might be of value for increased production and improved fatty acid composition of microalgal oils for biodiesel production.

Results

Lipid content and fatty acid composition of *P. tricornutum*

The growth parameters, total fatty acid (TFA) and TAG contents were monitored in axenic batch cultures of *P. tricornutum* on nitrogen-replete and nitrogen-depleted media (Fig. 1; Tables 1 and 2). The TFA composition of *P. tricornutum* (Table 1) was characterized by a large fraction of palmitic acid (16:0), palmitoleic acid (16:1), and eicosapentaenoic acid (EPA) (20:5 n-3). During growth on nitrogen-replete medium, palmitoleic acid showed a substantial increase (from 24% to 32% of TFAs). The levels of saturated fatty acids (palmitic and stearic acids) decreased slightly with culture age, whereas no significant change was observed in the relative proportion of EPA.

During starvation, turbidity and dry weight (DW) contents increased much more slowly than in nitrogen-replete cultures (Fig. 1). This decline in various growth parameters was accompanied by a four-fold increase in the TFA content, which reached 19% of biomass (% DW) after 12 days, with most of the increase occurring during the first 5 days (Table 1; Fig. 1). The main alterations in TFA profile were accounted for by relative increases in palmitic and palmitoleic acids (from 17% to 31% of TFAs, and from 24% to 43% of TFAs, respectively), concomitant with a decrease in the proportion of EPA (from 22% to 7% of TFAs). TAG content increased from 0.8% to 17.3% of DW during nitrogen starvation, representing an increase from 12% to 89% of TFAs accumulated in TAGs (Table 2).

TAG accumulation in lipid bodies under nitrogen starvation was visualized in algal cells with the fluorescent dye Nile Red (Fig. 2). TAGs of *P. tricornutum* contained predominantly palmitic (16:0) and palmitoleic (16:1) acids, constituting more than 75% of TAG fatty acids (Table 2). In addition, 14:0, 18:0, 18:1 and 20:5 n-3 fatty acids were also detected as TAG constituents (< 5–6% of TAG fatty acids). The major changes in TAG fatty acid composition observed under nitrogen starvation were an increase in the proportion of 16:0 acids from 32% to 37% of TFA, correlated with a decrease in the proportion of 16:1 acids from 46% to 38%.

Two types of mRNA from the single *PtDGAT1* gene

We made use of genomic annotation and expressed sequence tag (EST) data to clone a *PtDGAT1* cDNA. cDNA from algal cells grown on nitrogen-replete medium or cDNA from nitrogen-starved cells was amplified by PCR, with a forward primer designed on the basis of the EST (CT880495) and a reverse primer based on the nucleotide sequences of the predicted

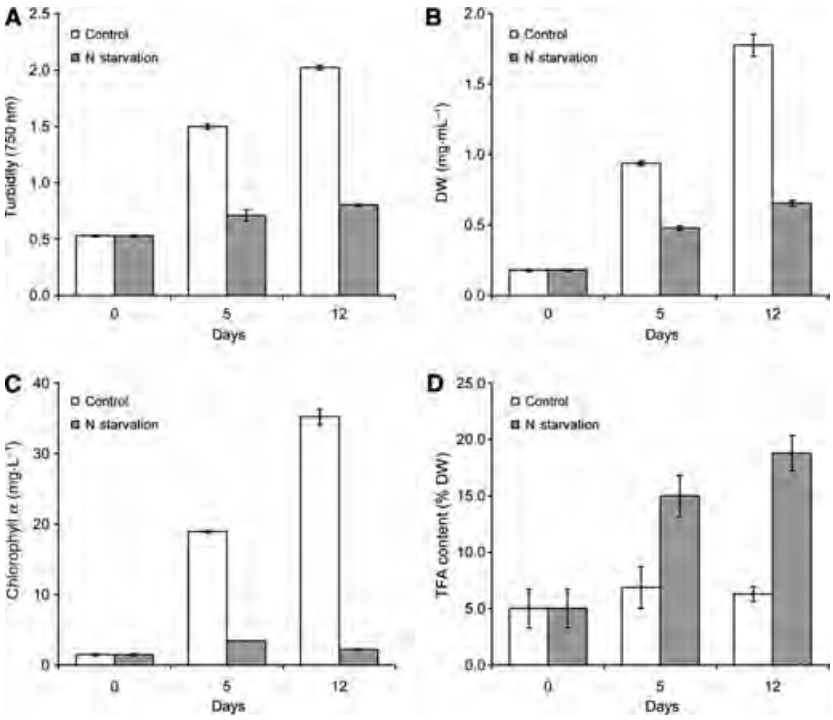

Fig. 1. Turbidity (A), DW (B), chlorophyll *a* content (C) and TFA content (D) measured on days 0, 5 and 12 in *P. tricornutum* cultures grown in full RSE medium (white columns), or under nitrogen starvation (gray columns). Results are expressed as the mean ± standard deviation (n = 3).

Table 1. Total fatty acid composition and content of *P. tricornutum* cultured on nitrogen-replete (+N) and nitrogen-depleted (–N) media. Results are expressed as the mean values of two biological repetitions. tr, traces.

	Time (days)				
	0	5		12	
Fatty acids (% TFA)	+N	+N	–N	+N	–N
Saturated fatty acids (SFAs)					
14:0	5.8	5.8	4.7	5.9	4.5
16:0	17.1	14.7	29.0	14.8	31.1
18:0	4.6	1.2	1.5	1.2	1.5
20:0	0.4	0.2	0.2	tr	0.2
22:0	0.4	0.3	0.2	0.2	0.2
Sum of SFAs	28.3	22.2	35.6	22.2	37.4
Monounsaturated fatty acids (MUFAs)					
16:1	23.9	33.2	39.0	32.0	42.6
18:1 n-9	2.9	2.2	3.8	2.9	4.0
18:1 n-7	1.7	1.1	2.0	0.9	2.6
20:1	0.7	0.2	–	–	–
Sum of MUFAs	29.2	36.7	44.8	35.8	49.2
PUFAs					
16:2 n-4	2.4	2.7	0.6	3.0	0.4
16:3 n-4	4.9	7.3	1.9	5.8	0.9
16:4	0.2	0.6	0.3	0.4	0.2
18:2 n-6	2.6	1.6	1.3	2.3	1.0
18:3 n-6	1.2	0.4	0.8	0.5	0.7
18:3 n-3	3.0	1.0	0.5	1.7	0.2
18:4 n-3	1.0	0.4	0.4	0.5	0.2
20:2	1.3	0.4	0.3	0.4	tr
20:3 n-3	0.4	–	–	–	tr
20:4 n-6 (ARA)	0.8	1.2	1.3	1.3	1.5
20:4 n-3	0.7	–	tr	tr	tr
20:5 n-3 (EPA)	21.9	23.5	10.2	23.6	7.0
22:6 n-3 (DHA)	2.1	2.0	2.0	2.4	1.0
Sum of PUFAs	42.5	41.1	19.6	42.0	13.4
TFAs (% of DW)	6.6	6.9	16.8	6.3	19.4

Table 2. TAG fatty acid composition and content of *P. tricornutum* cultured on nitrogen-depleted (–N) medium. Results are expressed as the mean values of two biological repetitions. tr, traces.

	Time (days)		
Fatty acids (% TFA)	0	5	12
Saturated fatty acids (SFAs)			
14:0	4.9	3.6	3.8
16:0	31.6	33.9	37.4
18:0	3.3	1.7	1.5
20:0	–	–	–
22:0	0.5	–	tr
Sum of SFAs	40.3	39.2	42.7
Monounsaturated fatty acids (MUFAs)			
16:1	46.4	43.6	38.2
18:1 n-9	3.5	3.5	5.6
18:1 n-7	1.5	2.5	1.8
20:1	–	–	–
Sum of MUFAs	51.4	49.6	45.7
PUFAs			
16:2 n-4	0.9	0.8	0.8
16:3 n-4	0.0	0.5	0.5
16:4	–	–	–
18:2 n-6	1.5	0.9	0.7
18:3 n-6	–	0.8	0.7
18:3 n-3	–	–	–
18:4 n-3	–	0.6	0.4
20:2	0.4	0.3	0.2
20:3 n-3	–	0.2	tr
20:4 n-6 (ARA)	–	1.0	1.4
20:4 n-3	–	–	–
20:5 n-3 (EPA)	5.2	5.8	6.0
22:6 n-3 (DHA)	0.4	0.4	0.6
Sum of PUFAs	8.3	11.2	11.6
TAGs (% of DW)	0.8	14.0	17.3
TAGs (% of TFA)	12.3	83.4	89.1

protein (XM_002177717.1) (similar to DGAT1). Two cDNA products of 1695 bp, encoding the full-length coding sequence of *PtDGAT1short*, and of 1768 bp (designated *PtDGAT1long*) (Fig. 3) were obtained. The two sequences differed by a single insert of 63 bp (insert II) (Fig. 4). Importantly, both sequences appeared to be slightly longer than the DNA size of 1635 bp (PhaeoJoinedSeq) predicted from assembling the database sequences of the hypothetical mRNA for the predicted protein (XM_002177717.1) with the EST (CT880495) from *P. tricornutum*.

We also aligned *PtDGAT1short* and *PtDGAT1long* to *P. tricornutum* chromosome 2 genomic DNA [23], confirming that the 63-bp insert in *PtDGAT1long* is an intron and that *PtDGAT1short* is probably subject to regulated splicing during stress induction or other cellular processes (Fig. 4). The amino acid sequences resulting from *PtDGAT1long* mRNA showed that insert II introduces stop codons into the ORF.

These nucleotide sequences were aligned, using CLUSTALW (http://www.ebi.ac.uk/Tools/clustalw2/index.html), to the predicted full-length cDNA (PhaeoJoinedSeq), and to three *P. tricornutum* EST sequences (CT887168, CT881105, and CT880495) showing similarity to the hypothetical *P. tricornutum* protein (XP_002177753.1) and to the coding regions of membrane-bound O-acyltransferase family proteins. The sequences determined by us and the EST sequences differed from the predicted gene sequence (XM_002177717.1) by the presence of a 54-bp fragment that does not affect the ORF, which is apparently an exon erroneously excluded during the gene assembly (XM_002177717.1) and that is also present in the genomic DNA sequence.

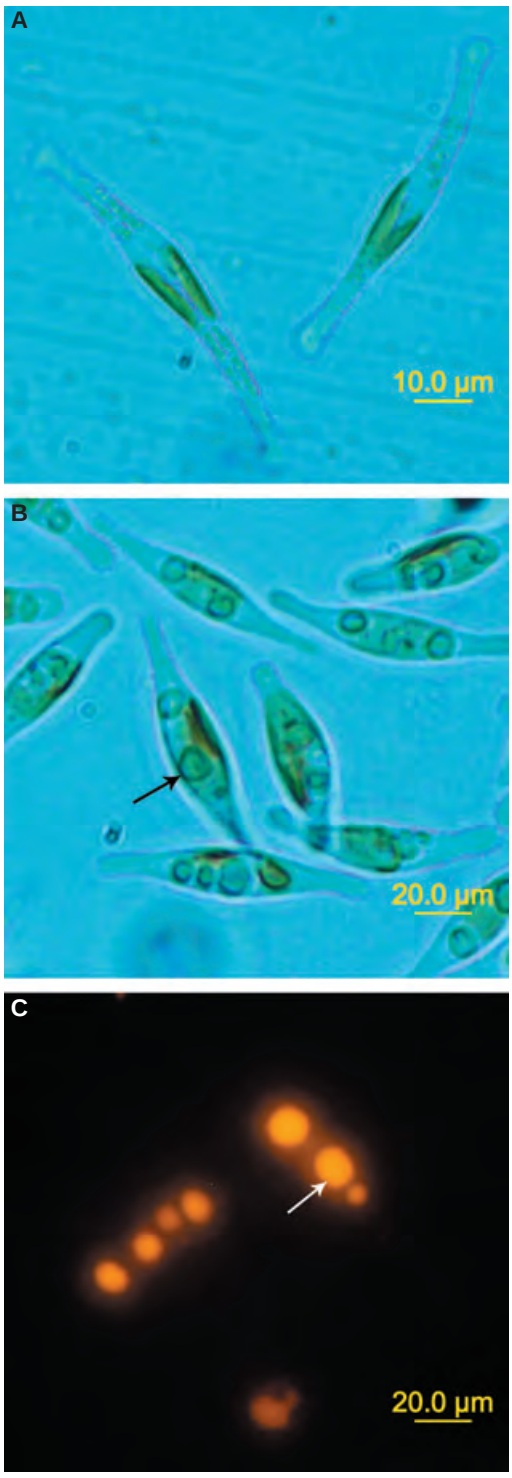

Fig. 2. Lipid body formation in *P. tricornutum* under nitrogen starvation. NL accumulation in lipid bodies was visualized in algal cells with the fluorescent dye Nile Red. (A) Cells grown in standard RSE medium. (B) Cells grown for 10 days in nitrogen-depleted RSE medium. (C) Cells grown for 10 days in nitrogen-depleted RSE medium and stained with Nile Red. The arrows indicate lipid bodies.

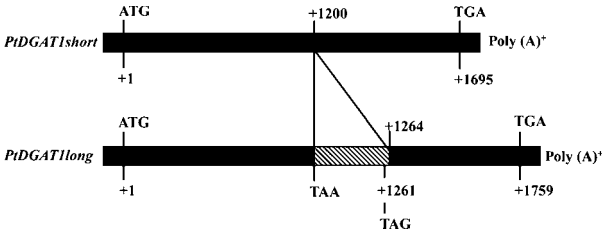

Fig. 3. Schematic diagram of *PtDGAT1short* and *PtDGAT1long* mRNAs transcribed from the *PtDGAT1* gene. The inserted sequence in *PtDGAT1long* is represented by a shaded box. The numbers represent nucleotides from adenine at the start codon. *PtDGAT1short* has a stop codon at nucleotides 1695–1697 (TGA), whereas *PtDGAT1long* has a stop codon (TAA) at the 5′-end (nucleotides 1200–1202) of an inserted sequence.

Gene structure and phylogenetic position of PtDGAT1

The ORF for *PtDGAT1short* was 1695 bp in length, coding for a 565-residue protein, PtDGAT1. The deduced amino acid sequence is 55% identical to that of the putative *T. pseudonana* DGAT1 (TpDGAT1) (XP_002287215); it shares more than 35% identity with DGAT1s of higher plants (Figs 5 and 6), and does not share significant homology with DGAT2s from higher plants and algae. DGAT1s of diatoms (putative TpDGAT1 and PtDGAT1) and some non-photosynthetic protists form a branch on the phylogenetic tree that is separate from that of the higher plant DGAT1s (Fig. 6). Within the green algae, a putative DGAT was also annotated in *C. variabilis*. All putative DGATs identified so far in the genome of *C. reinhardtii* are putative DGAT2 homologs.

At least eight strongly hydrophobic transmembrane regions (Fig. 7) were predicted by various algorithms [DAS (http://www.sbc.su.se/~miklos/DAS/), PHOBIUS (http://www.ebi.ac.uk/Tools/phobius/) and TMHMM (http://www.cbs.dtu.dk/services/TMHMM/)]. Moreover, the conserved motifs or putative signatures found in higher plants were slightly different in PtDGAT1 (Fig. 5). The characteristic basic motif found in the N-terminal amino acids, consisting of three arginines in higher plants [15,19,24], is substituted by KRS in PtDGAT1. The previously reported acyl-CoA-binding signature Ala133–Arg172 and the DAG-binding motif Val458–Val471 involved in the active site were found in PtDGAT1, as in DGAT1s of plants [6,24]. There was also a fatty acid-binding protein signature spanning residues Ala427 to Asn443, which contains a putative tyrosine phosphorylation site, Tyr438 [15]. Within the leucine zipper motif, only two of the six conserved leucines are present in PtDGAT1, besides the previously described critical prolines and serines

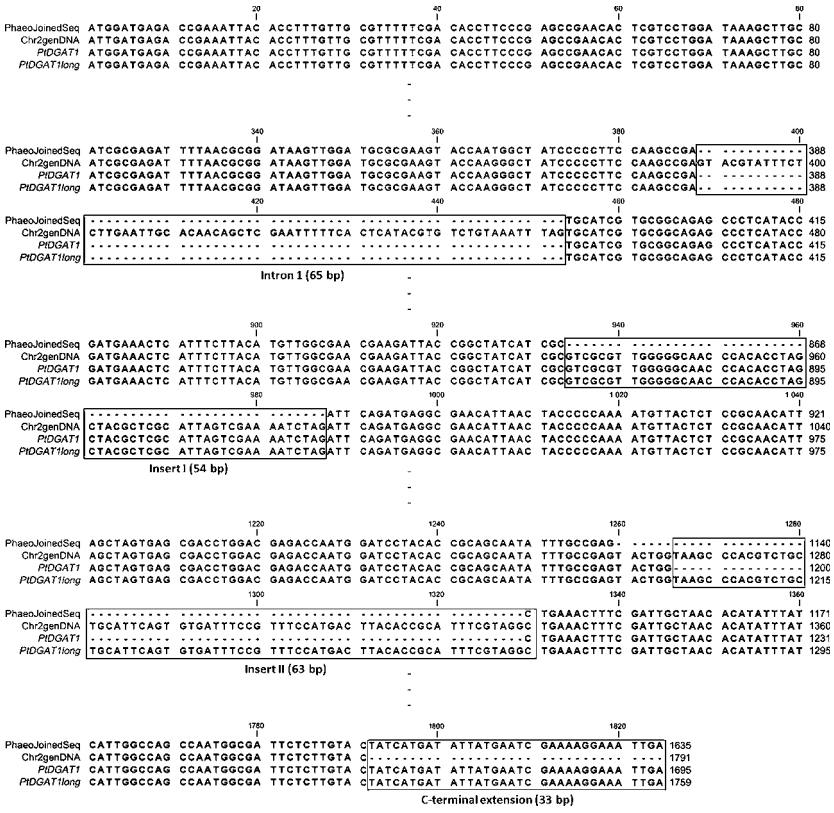

Fig. 4. Alignment of cDNA sequences of *PtDGAT1* (*PtDGAT1short*) and *PtDGAT1long* isolated from *P. tricornutum* with chromosome 2 genomic DNA (Chr2genDNA), the three ESTs and the putative full-length gene sequence for *PtDGAT1* (PhaeoJoinedSeq). Indicated are intron I of 65 bp, which is only observed in Chr2genDNA, insert I of 54 bp, which does not affect the ORF, insert II of 63 bp, which introduces a stop codon in the ORF of *PtDGAT1*, and the C-terminal extension of 33 bp.

coinciding with the thiolase acyl enzyme intermediate-binding signature in higher plants [15,25].

Heterologous expression of *PtDGAT1* in *S. cerevisiae*

To verify whether *PtDGAT1* does indeed encode a protein with DGAT activity, the gene was expressed in the heterologous yeast transformation system by restoring TAG biosynthesis and lipid body formation in an *S. cerevisiae* neutral lipid (NL)-deficient quadruple mutant strain (H1246). BY742 (wild type) and H1246 cells harboring an empty pYES2 vector were used as positive and negative controls, respectively. In addition, a quadruple mutant H1246, expressing the yeast *DGA1* (*DGAT2*) gene, was also used as a positive control to restore TAG formation by the single endogenous gene in the same experimental setup (Fig. 8). Upon expression of *PtDGAT1short*, a prominent spot corresponding to TAG appeared on the TLC plates of the lipid extract of the yeast mutant cells, showing successful restoration of TAG biosynthesis in the deficient phenotype by PtDGAT1 (Fig. 8). As expected, expression of *PtDGAT1long* did not complement the mutation and did not restore the yeast mutant strain's ability to synthesize TAGs.

The fluorescent dye Nile Red was used to stain lipid bodies in the quadruple mutant H1246 after transformation and in the BY742 strain as a control. Expression of *PtDGAT1short* restored the ability to form lipid bodies, as did expression of the yeast *DGA1* gene (Fig. 9D,E). As expected, formation of oil bodies was not observed following expression of *PtDGAT1long* containing the 63-bp insert and harboring stop codons (Fig. 9C), as intron retention introduces in-frame stop codons leading to a truncated DGAT protein and loss of activity.

Fatty acid substrate preferences of PtDGAT1

We tested whether expression of the active form of *PtDGAT1* (*PtDGAT1short*) could facilitate the incorporation of various polyunsaturated fatty acids (PUFAs) into yeast TAG. In parallel, we examined the ability of the endogenous DGAT activity encoded by *DGA1* to incorporate the same set of PUFAs.

In the absence of supplementary fatty acids, the recombinant DGATs utilized the available endogenous diacylglycerol (DAG) and acyl-CoA pool for acylation at the third position of the glycerol backbone. As can be seen in Table 3, PtDGAT1 showed a clear preference for endogenous saturated 16:0 and 18:0 species as

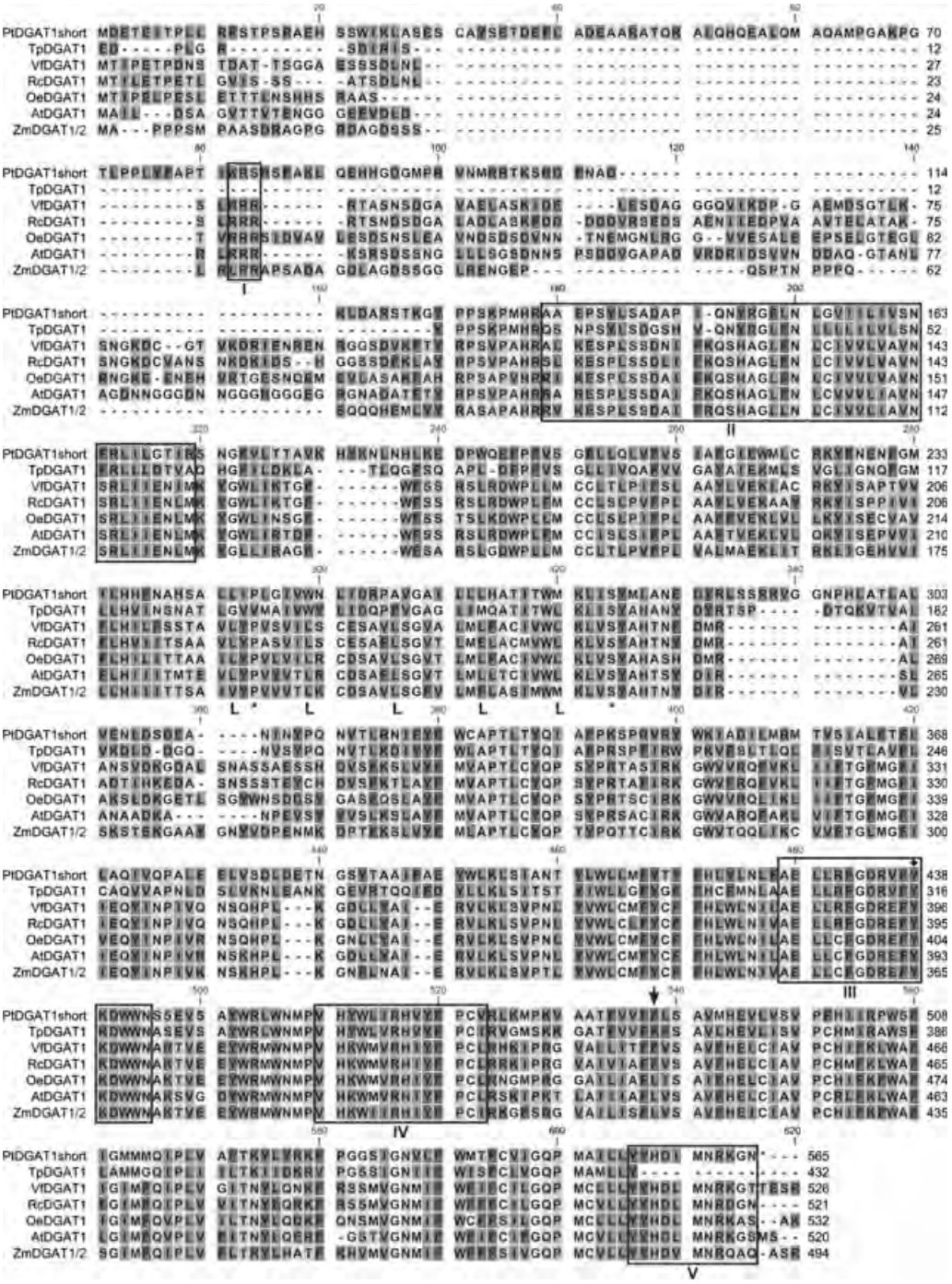

Fig. 5. Sequence comparison of PtDGAT1 with the hypothetical DGAT protein in *T. pseudonana* and related DGAT1 enzymes from higher plants: TpDGAT1 (*T. pseudonana*, accession no. XP002287215.1), OeDGAT1 (*Olea europaea*, accession no. AAS01606.1), VfDGAT1 (*Vernicia fordii*, accession no. ABC94471.1), RcDGAT1 (*Ricinus communis*, accession no. XP002514132.1), AtDGAT1 (*A. thaliana*, accession no. NP179535.1), and ZmDGAT1-2 (*Zea mays*, accession no. ABV91586.1). Conserved motifs or putative signatures (see text for details) are boxed, such as the N-terminal basic motif RRR in higher plants substituted by KRS in PtDGAT1 (I), the acyl-CoA binding signature (II), the fatty acid protein signature (III), which contains a tyrosine phosphorylation site (♦), the DAG-binding site (IV) and two C-terminal motifs – YYHD-like and NRKG-like (V) – as the putative endoplasmic reticulum retrieval motif in the C-terminus. The region containing a conserved leucine repeat (L) in higher plants coincides with a thiolase acyl enzyme intermediate binding signature besides the previously described critical proline and serine residues, which are marked by asterisks. The conserved phenylalanine is designated by an arrow.

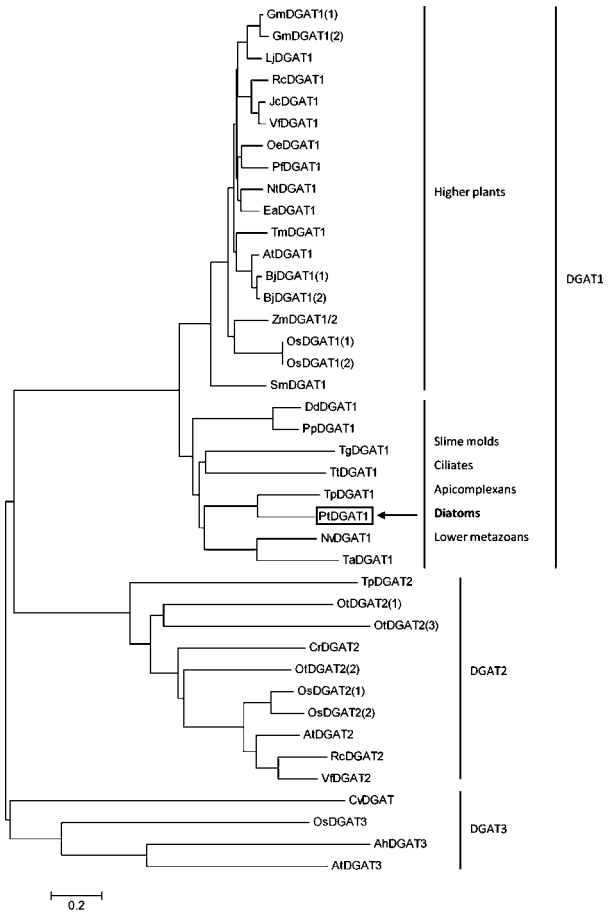

Fig. 6. A phylogram showing relationships among PtDGAT1 and diverse hypothetical and characterized DGAT-like proteins from diatoms and green algae, higher plants, and some primitive eukaryotes. The alignment was generated with CLUSTAL W, and the phylogram was constructed by the neighbor-joining method with MEGA.4: OsDGAT3 (*Oryza sativa* hypothetical protein, accession no. NP001054585.1), AhDGAT3 (*Arachis hypogaea*, accession no. AAX62735.1), AtDGAT3 (*A. thaliana* hypothetical protein, accession no. NP175264.2), CvDGAT (*C. variabilis* hypothetical protein, accession no. EFN50695.1), RcDGAT2 (*R. communis*, accession no. XP002528531.1), VfDGAT2 (*V. fordii*, accession no. ABC94473.1), AtDGAT2 (*A. thaliana*, accession no. NP566952.1), OsDGAT2(1) and OsDGAT2(2) (*Or. sativa* hypothetical protein, accession nos. NP001057530.1 and NP001047917.1), OtDGAT2(1), OtDGAT2(2), and OtDGAT2(3) (*O. tauri*, accession nos. CAL54993.1, CAL58088.1, and CAL56438.1), TpDGAT2 (*T. pseudonana* hypothetical protein, accession no. XP002286252.1), CrDGAT2 (*Ch. reinhardtii*, accession no. XP001693189), TpDGAT1 (*T. pseudonana*, accession no. XP002287215.1), OeDGAT1 (*Ol. europaea*, accession no. AAS01606.1), PfDGAT1 (*Perilla frutescens*, accession no. AAG23696.1), NtDGAT1 (*Nicotiana tabacum*, accession no. AAF19345.1), GmDGAT1(1) and GmDGAT1(2) (*Glycine max*, accession nos. AAS78662.1 and BAE93461.1), LjDGAT1 (*Lotus japonica*, accession no. AAW51456.1), EaDGAT1 (*Euonymus alatus*, accession no. AAV31083.1), JcDGAT1 (*Jatropha curcas*, accession no. ABB84383.1), VfDGAT1 (*V. fordii*, accession no. ABC94471.1), RcDGAT1 (*R. communis*, accession no. XP002514132.1), BjDGAT1(1) and BjDGAT1(2) (*Brassica juncea*, accession nos. AAY40784.1 and AAY40785.1), AtDGAT1 (*A. thaliana*, accession no. NP179535.1). TmDGAT1 (*Tr. majus*, accession no. AAM03340.2), ZmDGAT1/2 (*Z. mays*, accession no. ABV91586.1),OsDGAT1(1) and OsDGAT1(2) (*Or. sativa*, accession nos. NP001054869.2 and AAV10815.1), TgDGAT1 (*Toxoplasma gondii*, accession no. AAP94209.1), DdDGAT1 (*Dictyostelium discoideum*, accession no. XP645633.2), PpDGAT1 (*Polysphondylium pallidum*, accession no. EFA85004.1), TtDGAT1 (*Tetrahymena thermophila*, accession no. XP001014621.2), NvDGAT1 (*Nematostella vectensis* hypothetical protein, accession no. XP001639351.1), TaDGAT1 (*Trichoplax adhaerens* hypothetical protein, accession no. XP002112025.1), and SmDGAT1 (*Selaginella moellendorffii* hypothetical protein, accession no. XP002994237.1).

compared with the yeast DGA1, which was more selective towards monounsaturated 16:1 and 18:1 species. Thus TAGs formed by the action of PtDGAT1 are substantially more saturated than those formed by the yeast DGA1.

A supplementation assay (Table 3) was carried out with various polyunsaturated C_{18} and C_{20} fatty acids of both the n-3 and n-6 groups that are naturally present in *P. tricornutum* but not in *S. cerevisiae*. Exogenously supplemented PUFAs were incorporated into TAGs upon expression of both *PtDGAT1short* and *DGA1*. The expression of yeast *DGA1* resulted in higher levels of both C_{18} and C_{20} PUFAs in TAGs, except for the similar 20:5 n-3 incorporation. Moreover, the expression of both *PtDGAT1short* and *DGA1* resulted in higher incorporation of C_{18} PUFAs than C_{20} PUFAs into TAGs. Incorporation of 18:3 n-3 and 18:3 n-6 into TAGs of the recombinant yeast was associated with a corresponding decrease in the proportions of both monounsaturated 16:1 n-7 and 18:1 n-9 and an increase in the percentage of 16:0. In contrast, TAGs of the transformed yeast supplemented with 20:3 n-3, 20:4 n-3 and 20:5 n-3 showed only a decrease in the proportion of 18:1 n-9. Importantly, TAGs formed by the activity of PtDGAT1 were substantially more saturated, again indicating a preference for saturated acyl species. Whereas cells expressing *DGA1* incorporated n-6 C_{20} PUFAs into TAGs at higher proportions than cells expressing *PtDGAT1short*, incorporation of 20:5 n-3 was similar upon expression of both *PtDGAT1short* and *DGA1*.

Expression patterns of two types of *PtDGAT1* mRNA under nitrogen starvation

Quantitative real-time PCR (QRT-PCR) was performed to quantify the two types of *PtDGAT1* transcript, *PtDGAT1short* and *PtDGAT1long* mRNAs, in

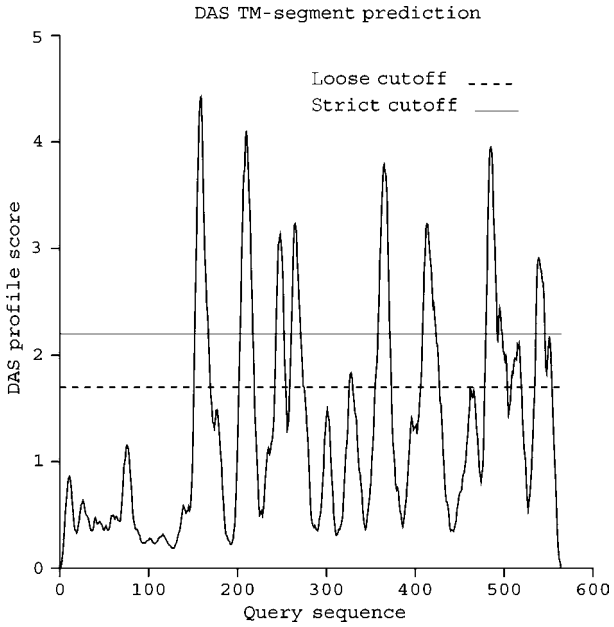

Fig. 7. Hydrophobicity plots indicating at least eight transmembrane (TM) regions of strong hydrophobicity (created using the DAS server).

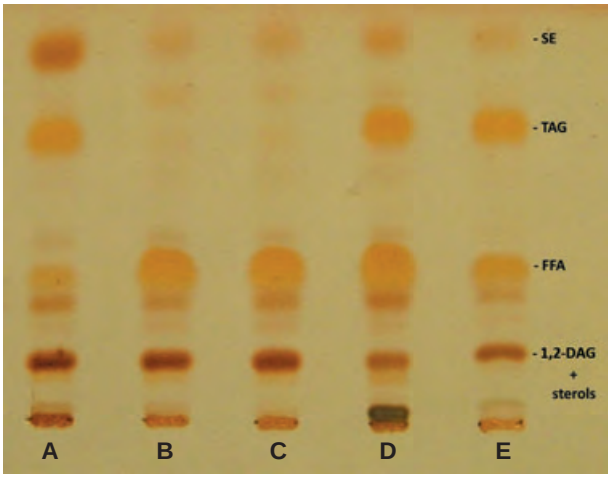

Fig. 8. Complementation of the TAG-deficient phenotype of the yeast mutant H1246 by expression of *PtDGAT1*. Expression was performed for 48 h at 30 °C. Lipid extracts were separated by TLC, and lipid spots were visualized as described in Experimental procedures. The wild-type strain BY742 harboring the empty plasmid was used as a positive control (A). The NL-deficient quadruple mutant strain H1246 harboring the empty plasmid was used as a negative control (B). The NL-deficient quadruple mutant strain H1246 does not express *PtDGAT1*, because of the 63-bp insert introducing stop codons into the ORF (C). Mutant strain H1246 expressing *PtDGAT1* (D). The NL-deficient quadruple mutant strain H1246 expressing yeast *DGA1* was used as a second positive control (E). SE, steryl ester; FFA, free fatty acid; 1,2-DAG, 1,2-diacylglycerol.

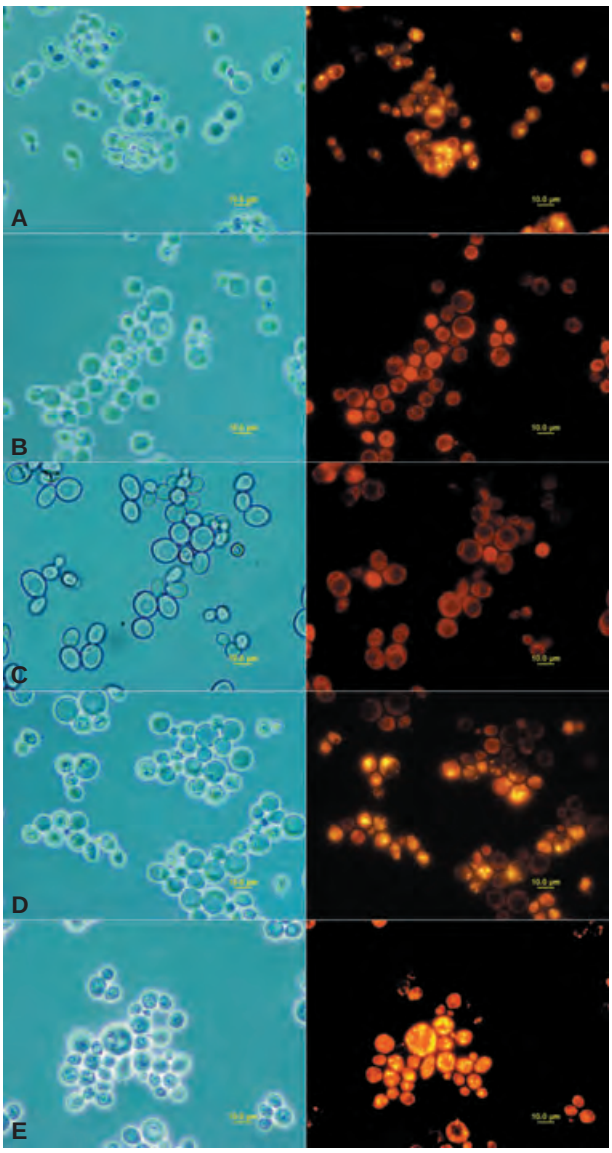

Fig. 9. Lipid body formation is restored upon expression of *PtDGAT1* in the yeast strain H1246. NL accumulation in lipid bodies was visualized in yeast cells with the fluorescent dye Nile Red. The wild-type strain BY742 harboring the empty plasmid was used as a positive control, (A). The NL-deficient quadruple mutant strain H1246 harboring the empty plasmid was used as a negative control (B). The NL-deficient quadruple mutant strain H1246 does not express *PtDGAT1*, because of the 63-bp insert introducing stop codons into the ORF (C). The mutant strain H1246 expresses *PtDGAT1* after intron-retention alternative splicing (D). The NL-deficient quadruple mutant strain H1246 expressing yeast *DGA1* was used as a second positive control (E).

P. tricornutum cells under nitrogen starvation (Fig. 10). The changes in expression level of the two types of transcript in *P. tricornutum* cells during the time course of nitrogen starvation were determined relative to the

Table 3. Fatty acid composition of TAGs isolated from *S. cerevisiae* H1246 cells transformed with pYES2 carrying *PtDGAT1* and yeast *DGA1*. Cultures were supplemented with 250 μM PUFA. Mean values ($n = 3$) are expressed as percentage of TFAs in the TAG fraction. ALA, α-linolenic acid; GLA, γ-linolenic acid; ETA, eicosatrienoic acid; MUFA, monounsaturated fatty acid; SFA, saturated fatty acid.

		Fatty acids in TAGs (% of total)					
PUFA substrates	Gene	16:0	16:1 n-7	18:0	18:1 n-9	Supplementary PUFA	SFA/MUFA[a]
Not supplemented	*PtDGAT1*	17.7	28.6	13.8	31.9	–	0.53
	DGA1	7.4	37.0	7.3	39.6	–	0.19
18:3 n-3 (ALA)	*PtDGAT1*	27.0	21.3	14.9	19.4	12.7	1.03
	DGA1	15.3	29.6	8.3	22.9	18.4	0.45
18:3 n-6 (GLA)	*PtDGAT1*	29.1	17.3	13.4	14.0	19.5	1.36
	DGA1	14.3	23.8	7.6	17.8	30.2	0.53
20:3 n-3 (ETA)	*PtDGAT1*	22.4	28.2	13.7	23.0	5.6	0.70
	DGA1	8.9	40.3	5.3	28.0	12.4	0.21
20:4 n-6 (ARA)	*PtDGAT1*	27.5	26.7	13.7	21.7	3.2	0.85
	DGA1	10.3	38.9	6.5	29.5	9.2	0.25
20:5 n-3 (EPA)	*PtDGAT1*	25.9	22.3	13.8	20.6	10.6	0.93
	DGA1	13.2	31.4	7.6	30.6	11.5	0.35

[a] Ratio of (16:0 + 18:0)/(16:1 + 18:1).

expression level of these transcripts in the log phase (time 0). The transcript abundance was normalized to that of the actin gene. The expression of *PtDGAT1long* mRNA increased to a maximum of 190-fold of control cells after 8–13 days of nitrogen starvation. In comparison, *PtDGAT1short* mRNA expression increased to a maximum of 4.5-fold above control (time 0) after 5 days of nitrogen starvation, and this was followed by a decrease at days 8 and 13. The expression levels of *PtDGAT1short* mRNA were always substantially lower than those of *PtDGAT1long*.

Discussion

We cloned and characterized a novel *DGAT1*-like gene (*PtDGAT1*) from the diatom *P. tricornutum*, encoding the first DGAT1-like protein described from microalgae and, specifically, diatoms. *PtDGAT1short* codes for a 565 amino acid sequence that exhibits similarity to previously identified *DGAT1*s in different plant species, but not to other available cloned algal *DGATs* characterized. Protein hydrophobicity analysis with various algorithms predicted that PtDGAT1 would contain eight putative transmembrane domains that are likely to anchor the protein to the endoplasmic reticulum membrane [13]. This is a characteristic feature of DGAT1s in plants [13,15,26,27], and is in agreement with the putative assessment of PtDGAT1 as DGAT1.

Diatom DGAT1s form a separate branch on the phylogenetic tree, deviating from the higher plant lineage, but similar to that of some heterotrophic protists (Fig. 6). The similarity of PtDGAT1 to the putative DGAT1 of ciliates, lower metazoans, apicomplexans and slime molds in comparison with DGAT1s of higher plants indicates that the origin of the gene encoding this protein is in the heterotrophic organism that diverged from the plant lineage long before the postulated secondary endosymbiosis event leading to the diatoms [28,29].

Heterologous expression studies in the *S. cerevisiae* NL-deficient quadruple mutant confirmed that *PtDGAT1short* encodes a protein that functions like DGAT, restoring TAG biosynthesis and lipid body formation in the NL-deficient quadruple mutant strain H1246 [30]. In two marine microalgae species, *O. tauri* and *T. pseudonana*, the functions of DGAT2 were also confirmed by restoring TAG biosynthesis in this quadruple mutant of *S. cerevisiae*.

Cloning of the full-length cDNA sequences of *PtDGAT1* transcripts revealed that two types of mRNA, *PtDGAT1short* and *PtDGAT1long*, were transcribed from the single *PtDGAT1* gene. PtDGAT1long lacked several catalytic domains, owing to the 63-bp insert, which introduces in-frame stop codons, resulting in loss of DGAT1 activity. QRT-PCR experiments showed that the expression levels of both transcripts were markedly increased under nitrogen starvation, but at a much lower level for *PtDGAT1short*. The production of *PtDGAT1short* could be a result of regulatory alternative splicing involving intron retention. Consequently, one possibility is that *PtDGAT1long* mRNA exists as a premature mRNA form for *PtDGAT1short* and is not translated to protein, whereas nitrogen starvation may activate RNA splicing to produce *PtDGAT1short* mRNA. This mechanism, by which multiple forms of mRNA are

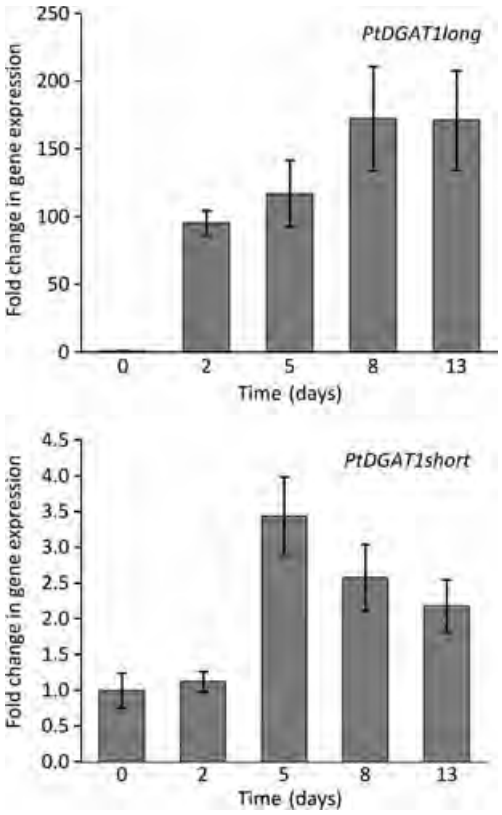

Fig. 10. Expression patterns for *PtDGAT1* gene transcripts under nitrogen starvation. QRT-PCR was performed with primers for amplification of the two types of *PtDGAT1* transcript, *PtDGAT1-short* and *PtDGAT1long* mRNAs. The change in expression level of the two types of transcript in *P. tricornutum* cells grown on nitrogen-free medium was calculated relative to the expression level of these transcripts in the log phase (time 0). The transcript abundance was normalized to that of the housekeeping gene *PtAct* (actin).

produced from a single transcript immediately after transcript synthesis, is the most common alternative splicing in *Arabidopsis* and rice (> 50%) [31,32]. Indeed, mRNAs with intron retention lead to truncated polypeptides, or are subjected to nonsense-mediated mRNA decay, as retained introns often introduce in-frame stop codons [33]. The generation of a truncated transcript and protein product at the expense of the corresponding active form of the enzyme could play a role in regulating the amount of active protein produced, depending on the level of correctly spliced transcript. Similar alternative RNA splicing was reported in the diatom *Chaetoceros compressum* [34], in which a novel heat stress-responsive gene, *HI-5*, encoding two types of transcript, a trypsin-like protease and its related protein, has been cloned.

Oleaginous microalgae cultivated under nutrient-depleted conditions, such as nitrogen starvation, can accumulate high amounts of NLs [35–39]. Global analysis of differential transcript levels in nitrogen-replete and nitrogen-deprived *C. reinhardtii* revealed a strong increase in the transcript level of the gene encoding one of the five putative DGATs (DGTT1, PMI 285889) identified in the genome of this microalga that was almost completely suppressed under nitrogen-replete conditions [39]. Owing to frequent mixing in natural marine environments, planktonic diatoms can encounter rapid changes in nutrient status and light exposure at different depths [40]. A higher expression level of active DGAT1 after a regulatory splicing step could thus provide an explanation, at the molecular level, for the rapid induction of TAG biosynthesis in *P. tricornutum* under nitrogen starvation and high light. Consequently, this mechanism could play a regulatory role during the physically forced mixing events described above in diatom acclimation and optimal adaptation to environmental factors.

Our results for TAG fatty acid composition are consistent with those of Yu *et al.* [41]. The major molecular species of TAG in *P. tricornutum* under starvation conditions are composed of 46:1, 48:1, 48:2 and 48:3 species, with palmitic (16:0), palmitoleic (16:1) and myristic (14:0) acid constituents [41]. The alga can also accumulate a certain percentage of long-chain PUFAs in TAG molecular species with a higher degree of unsaturation, containing 20:4 n-6, 20:5 n-3, and 22:6 n-3, on a lower scale [41]. Our results relating to substrate fatty acid specificity indicated a preference of PtDGAT1 for the production of TAGs with a high level of saturated fatty acids (16:0 and 18:0) in the recombinant *S. cerevisiae*. Indeed, under nitrogen starvation, *P. tricornutum* accumulates high levels of 16:0 and 16:1, and tends mainly to incorporate 16:0 into TAGs. The ability of PtDGAT1 to mediate the incorporation of saturated fatty acids such as 16:0 into TAGs is a beneficial feature for biodiesel production from microalgal and even plant oils, which are often characterized by a level of unsaturation that is higher than necessary for biodiesel standards.

In supplementation experiments, both PtDGAT1 and DGA1 incorporated similar amounts of 20:5 n-3, even though this fatty acid is not naturally present in yeast. The results for C_{20} PUFA supplementation allow us to speculate that PtDGAT1 may prefer n-3 C_{20} PUFA over n-6 C_{20} PUFA. Wagner *et al.* [22] compared *O. tauri* DGAT2 and yeast DGAT2 substrate specificity, and showed that both enzymes are promiscuous towards available acyl-CoA substrate and that they display a similar fatty acid preference, whereby the *T. pseudonana* DGAT2 more efficiently incorporated the n-3 long-chain PUFA DHA

(docosahexaenoic acid, 22:6 n-6) into TAGs than the recombinant DGAT1 of *A. thaliana* [21,22].

In conclusion, a novel *DGAT1*-like gene (*PtDGAT1*) from the diatom *P. tricornutum* was cloned. In yeast expression system, PtDGAT1 restored TAG and lipid body formation, and favored incorporation of saturated fatty acids into TAGs. This feature is of potential value for the genetic engineering of microalgae, and even plants, for enhanced biodiesel production. A possible molecular mechanism for rapid modulation of TAG biosynthesis in *P. tricornutum*, consisting of intron retention alternative splicing, regulates the amount of active protein produced under nitrogen starvation. Future research will shed more light on this molecular mechanism, which induces TAG accumulation under various stress conditions.

Experimental procedures

Growth conditions

Axenic cultures of *P. tricornutum* were obtained from the Culture Collection of Algae and Protozoa (CCAP) at the Scottish Marine Institute (SAMS Research Services Ltd, Oban, UK). *P. tricornutum* were cultivated on RSE medium in 250-mL glass Erlenmeyer flasks in an incubator shaker under an air/CO_2 atmosphere (99 : 1, v/v) and controlled temperature (25 °C) and illumination (100 μmol photons·m^{-2}·s^{-1}) at a speed of 150 r.p.m. RSE medium was composed of 34 g·L^{-1} ReefSalt (Seachem, Madison, GA, USA) complemented with: 18.8 mM KNO_3, 0.51 mM KH_2PO_4, 10.6 mM $Na_2SiO_3.9H_2O$, 0.77 μM $ZnSO_4.7H_2O$, 0.31 μM $CuSO_4.5H_2O$, 1.61 μM $Na_2MoO_4.2H_2O$, 46.3 μM H_3BO_3, 9.15 μM $MnCl_2.4H_2O$, 0.172 μM $Co(NO_3)_2.6H_2O$, 26.8 μM $C_6H_5O_7Fe.5H_2O$, 46.8 μM citric acid, 50 μg·L^{-1} vitamin B_{12}, 50 μg·L^{-1} biotin, and 0.1 mg·L^{-1} thiamine-HCl. To induce nitrogen starvation, daily-diluted cultures were centrifuged (1200 ***g*** for 5 min), washed twice, and resuspended in nitrogen-free RSE medium. Cultures were further maintained under the same conditions before being harvested for RNA isolation. The nitrogen-free medium was prepared by omitting KNO_3 from the RSE medium. Growth parameters (turbidity, chlorophyll *a* and DW content) were evaluated as previously described [42].

RNA isolation and cDNA synthesis

Total RNA was isolated with the procedure described by Bekesiova *et al.* [43] from 35 mL of culture grown in complete or nitrogen-free RSE medium for 5 days. The cells were harvested by centrifugation at 1000 ***g*** for 5 min, flash-frozen in liquid nitrogen, and stored at −80 °C until further use. Total RNA samples were treated with RNA-free Baseline-ZERO DNase (Epicentre Technologies, Madison, WI, USA) before being used for cDNA synthesis. cDNA was prepared from 1 μg of total RNA-template with the Verso cDNA kit (Thermo Fisher Scientific, Epson, UK).

Identification and cloning of the *PtDGAT1* gene

Detailed searches with bioinformatics tools and available databases were employed to find putative *DGAT* candidates in the genome of *P. tricornutum*. A hypothetical *P. tricornutum* protein (XP_002177753.1) containing a 404 amino acid exhibited significant similarity (38% identity and 55% similarity) to the coding region of maize DGAT1/2 protein (EU039830). A putative full-length cDNA encoding PtDGAT1 (534 amino acids) including a putative N-terminus (130 amino acids) was further assembled from the nucleotide sequence of hypothetical protein XM_002177717.1, and the EST (CT880495) from *P. tricornutum*, which shares a common 440-base overlap. The sequences of the oligonucleotide primers used in this study are given in Table 4. The ORF coding for the putative 545-residue PtDGAT1

Table 4. Oligonucleotide primers used in this work.

Primer	Sequence (5′- to 3′)		
*Pt*DGAT*for-Kpn*I	CGCGGT**ACCATG**GATGAGACCGAAATTACAC		
*Pt*DGAT*rev-Xho*I	5GGCCTCGAG**TCA**ATTTCCTTTTCGATTCATAATATCATGATAGTACAAGAGAATCGCCATTGG		
DGA1-f*Bam*HI	GGATCC*ACATA***ATG**TCAGGAACATTCAATGATATAAG		
DGA1-r*Not*I	TGCGGCCGC**TTA**CCCAACTATCTTCAATTCTGCATC		
QRT-PCR primer	Sequence	Amplicon size (bp)	PCR efficiency (%)
PtActF	TATTGTTCATCGCAAGTGCTTCTAA	80	100.7
PtActR	TAATACACCTCCTACAAACGTTGAAGA		
PtDGAT1shortF	TTGCCGAGTACTGGCTGAAA	93	102.6
PtDGAT1shortR	CAAAGAGGTTCAGATACAAATGGAAA		
PtDGAT1longF	TGTGATTTCCGTTTCCATGACT	117	102.3
PtDGAT1longR	AAGAGGTTCAGATACAAATGGAAATATG		

polypeptide was amplified with *PfuUltra* II fusion HS DNA polymerase (Stratagene, La Jolla, CA, USA), with the forward primer (*Pt*DGAT*for-Kpn*I) containing a *Kpn*I restriction site (underlined) and a yeast translation initiation consensus followed by ATG (bold), and the reverse primer (*Pt*DGAT*rev-Xho*I) containing a C-terminal extension, an *Xho*I restriction site (underlined), and a stop codon (in bold). As hypothetical protein XP_002177753.1 lacks at least 10 amino acids that are conserved in higher plant DGAT1s as well as a stop codon, 10 C-terminal amino acids (conserved in the C-terminus of higher plant DGAT1), including two motifs -YYHD-like and NRKG-like as the putative endoplasmic reticulum retrieval motif (ER-DIR), were added during amplification by incorporating a 33-base nucleotide sequence into the reverse primer encoding amino acids YHDIMNRKGN-Stop. The PCR products of the expected size were excised, purified from the gel (Nucleospin Extract II purification kit; Macherey-Nagel, Duren, Germany), and cloned into the pGEM-T Easy Vector System (Promega, Madison, WI, USA), and several clones were sequenced for each condition.

As a positive control in yeast expression assays, the yeast *DGA1* gene encoding DGAT2 was cloned similarly to *PtDGAT1* of *P. tricornutum*, with the forward primer *DGA1*-f*Bam*HI and reverse primer *DGA1*-r*Not*I.

Expression and functional characterization of *PtDGAT1* cDNA by heterologous expression in *S. cerevisiae*

S. cerevisiae strains BY742 (relevant genotype: *MATa his3-Δ1 leu2Δ0 met15Δ0 ura3Δ0*) and H1246 (relevant genotype: *MATα ADE2-1 can1-100 ura3-1 are1-Δ::HIS3 are2-Δ::LEU 2 dga1-Δ::KanMX4 Iro1-Δ::TRP1*) containing knockouts of *DGA1*, *LRO1*, *ARE1* and *ARE2* [30] were kindly provided by S. Stymne (Scandinavian Biotechnology Research, Alnarp, Sweden). Before transformation, yeast cells were cultivated in 1% (w/v) yeast extract, 2% (w/v) peptone and 2% (w/v) glucose (YPG medium) at 30 °C.

Plasmids (pGEM-T Easy Vector System) harboring the assembled full-length sequence encoding PtDGAT1 (1653 bp) were then restricted with *Kpn*I and *Xho*I (NEB, Ipswich, MA, USA). The expected bands were purified from the gel (Nucleospin Extract II purification kit), and ligated into a *Kpn*I–*Xho*I-cut pYES2 vector (Invitrogen, Carlsbad, CA, USA), and the *S. cerevisiae* strains were transformed by the poly(ethylene glycol)/lithium acetate method [44]. H1246 and BY742 cells harboring the empty pYES2 vector were used as negative and positive controls, respectively; as a second positive control, H1246 cells harboring yeast *DGA1* were used. Transformants were selected by uracil prototrophy on yeast synthetic medium (YSM) lacking uracil (Invitrogen).

For functional expression, a minimal selection medium containing 2% (w/v) raffinose was inoculated with the *PtDGAT1* transformants and incubated at 30 °C for 24 h in a water-bath shaker. Sterile YSM (20 mL) was inoculated with raffinose-grown cultures to an attenuance of 0.2 at 600 nm. Expression was induced by adding galactose to a final concentration of 2% (w/v), and cultures were grown at 30 °C for a further 48 h. To determine fatty acid substrate preferences of PtDGAT1 versus *S. cerevisiae* DGA1, supplementation of YSM cultures with α-linolenic acid (18:3 n-3), γ-linolenic acid (18:3 n-6), eicosatrienoic acid (20:3 n-3), arachidonic acid (ARA) (20:4 n-6) and EPA (20:5 n-3) was carried out with 250 μM of the appropriate fatty acid in the presence of 1% (w/v) Tergitol-40 (Sigma-Aldrich, St Louis, MO, USA). Cells were harvested by centrifugation (1200 ***g*** for 5 min), washed twice with 0.1% (w/v) $NaHCO_3$, freeze-dried, and used for lipid analysis.

Nile Red staining and microscopy

The Nile Red staining method described by Greenspan *et al.* [45] was used to visualize the intracellular lipid bodies as an indicator of TAG formation. Aliquots (200 μL) of *P. tricornutum* cultures or yeast cultures grown for 48 h were harvested, stained with 2 μL of Nile Red (0.5 $mg{\cdot}mL^{-1}$ in dimethylsulfoxide), incubated at room temperature for 5 min, and immediately observed by fluorescence microscopy (Carl Zeiss, Göttingen, Germany). A filter allowing maximum excitation at 450–490 nm and a 520-nm cut-off filter were used.

Lipid extraction and analysis

For lipid analysis of *P. tricornutum* and yeast expression cultures, cells were harvested by centrifugation (1200 ***g*** for 5 min), and treated with isopropanol at 80 °C for 10–15 min to stop lipolytic activity. Isopropanol was evaporated under nitrogen gas before lipid extraction according to a modified version of the Bligh and Dyer method [46]. The NL classes were resolved from the total lipid extracts by TLC on Silica Gel 60 plates (Merck, Darmstadt, Germany) in petroleum ether/diethyl ether/glacial acetic acid (70 : 30 : 1, v/v). Lipid standards were included on each TLC plate. TAGs were recovered from the TLC plates for GC analysis following a brief exposure to iodine vapor. For charring, plates were sprayed with a solution of 10% (v/v) H_2SO_4 in methanol, and heated until spots appeared.

Fatty acid analysis

For fatty acid and lipid analyses of *P. tricornutum*, cells were harvested from liquid cultures. Fatty acid methyl esters (FAMEs) from total lipids and TAGs were obtained by transmethylation of the freeze-dried cells or from TAGs extracted with dry methanol containing 2% (v/v) H_2SO_4

and heating at 80 °C for 1.5 h with stirring under an argon atmosphere. Gas chromatographic analysis of FAMEs was performed on a Thermo Ultra gas chromatograph (Thermo Scientific, Italy) equipped with a PTV injector, flame ionization detector, and a fused silica capillary column (30 m × 0.32 mm; Supelco WAX-10; Sigma-Aldrich, Bellefonte, PA, USA). FAMEs were provisionally identified by co-chromatography with authentic standards (Sigma Chemicals, St Louis, MO, USA) and FAMEs of fish oil (Larodan Fine Chemicals, Malmö, Sweden). Results are expressed as the mean values of three replicates ($n = 3$) for two biological repetitions.

QRT-PCR

Template cDNA for QRT-PCR was synthesized from 1 μg of total RNA in a total volume of 20 μL, with oligo-dT primer (Reverse-iT 1st Strand Synthesis Kit; ABgene, Epsom, UK). Each 20-μL cDNA reaction was then diluted 10-fold with PCR-grade water.

QRT-PCR primer design and validation

QRT-PCR primer pairs (Table 4) were designed for *PtDGAT1short*, *PtDGAT1long* and the housekeeping gene *PtAct* (actin) with PRIMER EXPRESS v2.0 (Applied Biosystems, Foster City, CA, USA). *PtDGAT1long* primers were designed to amplify the inserted 63-bp nucleotides of the *PtDGAT1long* cDNA, and *PtDGAT1short* primers to amplify the *PtDGAT1short* cDNA. In order to specifically amplify *PtDGAT1short* cDNA, the forward primer of this pair was designed at the nucleotide junction after splicing. Conditions were set for a primer length of 20–28 bp, a primer melting temperature of 60.0 ± 1.0 °C, and an amplicon length of 50–150. Primer pairs were validated with five serial 80-fold dilutions of cDNA samples, and standard curves were plotted to test for linearity of the response. The primer pairs and primer concentrations with reaction efficiencies of 100% ± 10% were chosen for QRT-PCR analysis of relative gene expression. The nucleotide sequences and characteristics of primers used for QRT-PCR analysis are presented in Table 4.

Gene expression profiling

Gene expression profiling was performed by QRT-PCR, using triplicate reactions for each sample of two independent RNA isolations with a gene-specific primer pair and SsoFast EvaGreen Supermix (Bio-Rad, Hercules, CA, USA) in a CFX96 Real-Time System (Bio-Rad). The amplification procedure was 95 °C for 20 s, 40 cycles of 95 °C for 3 s, and 60 °C for 30 s. A melting curve was obtained for each pair of primers to confirm that a single, specific product was produced in each reaction.

Calculation of gene transcript levels

The mean changes in gene expression, as multiples of the original values, were calculated according to the $2^{-\Delta\Delta Ct}$ method [47], using the average of threshold cycle (Ct) values from triplicate cDNA–primer samples. The ΔCt followed by the ΔΔCt was calculated from the average Ct values of the target and the endogenous genes. The transcript abundance of the two obtained sequences, *PtDGAT1short* and *PtDGAT1long*, were normalized to the endogenous control *PtAct*. The changes in gene expression were calculated with $2^{-\Delta\Delta Ct}$ to find the expression level of the target gene, which was normalized to the endogenous gene relative to the expression of the target gene at time 0.

Phylogenetic analysis

The encoded protein sequences of known or putative *DGAT* genes were aligned with CLUSTALW (http://www.ebi.ac.uk/Tools/msa/clustalw2/), and the phylogram was constructed by the neighbor-joining method with MEGA.4 (http://www.megasoftware.net/).

Acknowledgements

This work was supported in part by the Jacob Blaustein Institutes for Desert Research - Jacob Blaustein Center for Scientific Cooperation (BCSC), and by the FP 7 European Project GIAVAP 'Genetic Improvement of Algae for Value Added Products'.

References

1 Kennedy EP & Weiss SB (1956) The function of cytidine coenzymes in the biosynthesis of phospholipids. *J Biol Chem* **222**, 193–214.
2 Ichihara K, Takahashi T & Fujii S (1988) Diacylglycerol acyltransferase in maturing safflower seeds: its influences on the fatty acid composition of triacylglycerol and on the rate of triacylglycerol synthesis. *Biochim Biophys Acta* **958**, 125–129.
3 Perry HJ & Harwood JL (1993) Changes in lipid content of developing seeds of *Brassica napus*. *Phytochemistry* **32**, 1411–1415.
4 Perry HJ & Harwood JL (1993) Radiolabelling studies of acyl lipids in developing seeds of *Brassica napus*: use of [1-^{14}C]acetate precursor. *Phytochemistry* **33**, 329–333.
5 Perry HJ, Bligny R, Gout E & Harwood JL (1999) Changes in Kennedy pathway intermediates associated with increased triacylglycerol synthesis in oil-seeds rape. *Phytochemistry* **52**, 799–804.
6 Jako C, Kumar A, Wei Y, Zou J, Barton DL, Giblin EM, Covello PS & Taylor DC (2001) Seed-specific

over-expression of an arabidopsis cDNA encoding a diacylglycerol acyltransferase enhances seed oil content and seed weight. *Plant Physiol* **126**, 861–874.
7 Weselake RJ (2005) Storage lipids. In *Plant Lipids: Biology, Utilisation and Manipulation* (Murphy DJ, ed.), pp. 402. CRC Press, Blackwell Publishing, Oxford.
8 Lung SC & Weselake RJ (2006) Diacylglycerol acyltransferase: a key mediator of plant triacylglycerol synthesis. *Lipids* **41**, 1073–1088.
9 Dahlqvist A, Ståhl U, Lenman M, Banas A, Lee M, Sandager L, Ronne H & Stymne S (2000) Phospholipid:diacylglycerol acyltransferase: an enzyme that \catalyzes the acyl-CoA-independent formation of triacylglycerol in yeast and plants. *Proc Natl Acad Sci USA* **97**, 6487–6492.
10 Ståhl U, Carlsson AS, Lenman M, Dahlqvist A, Huang B, Banas W, Banas A & Stymne S (2004) Cloning and functional characterization of a phospholipid:diacylglycerol acyltransferase from Arabidopsis. *Plant Physiol* **135**, 1324–1335.
11 Zhang M, Fan J, Taylor DC & Ohlrogge JB (2009) DGAT1 and PDAT1 acyltransferases have overlapping functions in Arabidopsis triacylglycerol biosynthesis and are essential for normal pollen and seed development. *Plant Cell* **21**, 3885–3901.
12 Saha S, Enuguttim B, Rajakumari S & Rajasekharan R (2006) Cytosolic triacylglycerol biosynthetic pathway in oilseeds. Molecular cloning and expression of peanut cytosolic diacylglycerol acyltransferase. *Plant Physiol* **141**, 1533–1543.
13 Shockey JM, Gidda SK, Chapital DC, Kuan JC, Dhanoa PK, Bland JM, Rothstein SJ, Mullen RT & Dyer JM (2006) Tung tree DGAT1 and DGAT2 have nonredundant functions in triacylglycerol biosynthesis and are localized to different subdomains of the endoplasmic reticulum. *Plant Cell* **18**, 2294–2313.
14 Yen CLE, Stone SJ, Koliwad S & Harris C-RV Jr (2008) DGAT enzymes and triacylglycerol biosynthesis. *J Lipid Res* **49**, 2283–2301.
15 Xu J, Francis T, Mietkiewska E, Giblin EM, Barton DL, Zhang Y, Zhang M & Taylor DC (2008) Cloning and characterization of an acyl-CoA-dependent diacylglycerol acyltransferase 1 (DGAT1) gene from *Tropaeolum majus*, and a study of the functional motifs of the DGAT protein using site-directed mutagenesis to modify enzyme activity and oil content. *Plant Biotechnol J* **6**, 799–818.
16 Zheng H, Allen WB, Roesler K, Williams ME, Zhang S, Li S, Glassman K, Ranch J, Nubel D, Solawetz W *et al.* (2008) A phenylalanine in DGAT is a key determinant of oil content and composition in maize. *Nat Genet* **40**, 367–372.
17 Klaus D, Ohlrogge JB, Neuhaus HE & Dormann P (2004) Increased fatty acid production in potato by engineering of acetyl-CoA carboxylase. *Planta* **219**, 389–396.
18 Zou J, Katavic V, Giblin EM, Barton DL, MacKenzie SL, Keller WA, Hu X & Taylor DC (1997) Modification of seed oil content and acyl composition in the Brassicaceae by expression of a yeast sn-2 acyltransferase gene. *Plant Cell* **9**, 909–923.
19 Bouvier-Navé P, Benveniste P, Oelkers P, Sturley SL & Schaller H (2000) Expression in yeast and tobacco of plant cDNAs encoding acyl CoA:diacylglycerol acyltransferase. *Eur J Biochem* **267**, 85–96.
20 Andrianov V, Borisjuk N, Pogrebnyak N, Brinker A, Dixon J, Spitsin S, Flynn J, Matyszczuk P, Andryszak K, Laurelli M *et al.* (2010) Tobacco as a production platform for biofuel: overexpression of Arabidopsis DGAT and LEC2 genes increases accumulation and shifts the composition of lipids in green biomass. *Plant Biotechnol J* **8**, 277–287.
21 Zou J, Xu J & Zhen Z (2009) Diacylglycerol acyltransferase 2 genes and proteins encoded thereby from algae. WIPO Patent Application WO/2009/085169.
22 Wagner M, Hoppe K, Czabany T, Heilmann M, Daum G, Feussner I & Fulda M (2010) Identification and characterization of an acyl-CoA:diacylglycerol acyltransferase 2 (DGAT2) gene from the microalga *O. tauri*. *Plant Physiol Biochem* **48**, 407–416.
23 Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, Kuo A, Maheswari U, Martens C, Maumus F, Otillar RP *et al.* (2008) The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature* **456**, 239–244.
24 Mañas-Fernández A, Vilches-Ferrón M, Garrido-Cárdenas JA, Belarbi EH, Alonso DL & García-Maroto F (2009) Cloning and molecular characterization of the acyl-CoA:diacylglycerol acyltransferase 1 (DGAT1) gene from *Echium*. *Lipids* **44**, 555–568.
25 Zou J, Wei Y, Jako C, Kumar A, Selvaraj G & Taylor DC (1999) The *Arabidopsis thaliana* TAG1 mutant has a mutation in a diacylglycerol acyl-transferase gene. *Plant J* **19**, 645–653.
26 Deng X, Li Y & Fei X (2009) Microalgae: a promising feedstock for biodiesel. *Afr J Microbiol Res* **3**, 1008–1014.
27 Hobbs DH, Lu C & Hills MJ (1999) Cloning of a cDNA encoding diacylglycerol acyltransferase from *Arabidopsis thaliana* and its functional expression. *FEBS Lett* **452**, 145–149.
28 Kroth P (2006) Molecular biology and the biotechnology potential of diatoms. *Adv Exp Med Biol* **616**, 23–33.
29 Keeling PJ, Burger G, Durnford DG, Lang BF, Lee RW, Pearlman RE, Roger AJ & Gray MW (2005) The tree of eukaryotes. *Trends Ecol Evol* **20**, 670–676.
30 Sandager L, Gustavsson MH, Ståhl U, Dahlqvist A, Wiberg E, Banas A, Lenman M, Ronne H & Stymne S (2002) Storage lipid synthesis is non-essential in yeast. *J Biol Chem* **277**, 6478–6482.

31 Ner-Gaon H, Halachmi R, Savaldi-Goldstein S, Rubin E, Ophir R & Fluhr R (2004) Intron retention is a major phenomenon in alternative splicing in Arabidopsis. *Plant J* **39**, 877–885.
32 Kim E, Magen A & Ast G (2007) Different levels of alternative splicing among eukaryotes. *Nucleic Acids Res* **35**, 125–131.
33 Maquat LE (2004) Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. *Nat Rev* **5**, 89–99.
34 Kinoshita S, Kaneko G, Lee JH, Kikuchi K, Yamada H, Hara T, Itoh Y & Watabe S (2001) A novel heat-stress responsive gene in the diatom *Chaetoceros compressum* encoding two types of transcripts, a trypsin-like protease and its related protein, by alternative RNA splicing. *Eur J Biochem* **268**, 4599–4609.
35 Piorreck M, Baasch KH & Pohl P (1984) Biomass production, total protein, chlorophylls, lipids and fatty acids of freshwater green and blue-green algae under different nitrogen regimes. *Phytochemistry* **23**, 207–216.
36 Larson TR & Rees TAV (1996) Changes in cell composition and lipid metabolism mediated by sodium and nitrogen availability in the marine diatom *Phaeodactylum tricornutum* (Bacillariophyceae). *J Phycol* **32**, 388–393.
37 Illman AM, Scragg AH & Shales SW (2000) Increase in *Chlorella* strains calorific values when grown in low nitrogen medium. *Enzyme Microb Technol* **27**, 631–635.
38 Scragg AH, Illman AM, Carden A & Shales SW (2002) Growth of microalgae with increased calorific values in a tubular bioreactor. *Biomass Bioenerg* **23**, 67–73.
39 Miller R, Wu G, Deshpande RR, Vieler A, Gaertner K, Li X, Moellering ER, Zäuner S, Cornish A & Liu B *et al.* (2010) Changes in transcript abundance in *Chlamydomonas reinhardtii* following nitrogen-deprivation predict diversion of metabolism. *Plant Physiol* **154**, 1737–1752.
40 Allen AE, Vardi A & Bowler AC (2006) An ecological and evolutionary context for integrated nitrogen metabolism and related signaling pathways in marine diatoms. *Curr Opin Plant Biol* **9**, 264–273.
41 Yu ET, Zendejas FJ, Lane PD, Gaucher S, Simmons BA & Lane TW (2009) Triacylglycerol accumulation and profiling in the model diatoms *Thalassiosira pseudonana* and *Phaeodactylum tricornutum* (Baccilariophyceae) during starvation. *J Appl Phycol* **21**, 669–681.
42 Solovchenko AE, Khozin-Goldberg I, Cohen Z & Merzlyak MN (2009) Carotenoid-to-chlorophyll ratio as a proxy for assay of total fatty acids and arachidonic acid content in the green microalga *Parietochloris incise*. *J Appl Phycol* **21**, 361–366.
43 Bekesiova I, Nap JP & Mlynarova L (1999) Isolation of high quality DNA and RNA from leaves of the carnivorous plant *Drosera rotundifolia*. *Plant Mol Biol Rep* **17**, 269–277.
44 Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K, Albright LM, Cohen DM & Varki A (1995) *Current Protocols in Molecular Biology*. Wiley, New York.
45 Greenspan P, Mayer EP & Fowler SD (1985) Nile red: a selective fluorescent stain for intracellular lipid droplets. *J Cell Biol* **100**, 965–973.
46 Bligh EG & Dyer WJ (1959) A rapid method of lipid extraction and purification. *Can J Biochem Physiol* **37**, 911–917.
47 Livak KJ & Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. *Methods* **25**, 402–408.