Soybean oil biosynthesis: Role of diacylglycerol acyltransferases

ARTICLE *in* FUNCTIONAL & INTEGRATIVE GENOMICS · JANUARY 2013

Impact Factor: 2.48 · DOI: 10.1007/s10142-012-0306-z · Source: PubMed

6 AUTHORS, INCLUDING:

Tomoko Hatanaka
Kobe University
28 PUBLICATIONS **398** CITATIONS
SEE PROFILE

Hirotada Fukushige
University of Kentucky
21 PUBLICATIONS **805** CITATIONS
SEE PROFILE

Funct Integr Genomics (2013) 13:99–113
DOI 10.1007/s10142-012-0306-z

ORIGINAL PAPER

Soybean oil biosynthesis: role of diacylglycerol acyltransferases

Runzhi Li · Tomoko Hatanaka · Keshun Yu · Yongmei Wu · Hirotada Fukushige · David Hildebrand

Received: 6 July 2012 / Revised: 17 December 2012 / Accepted: 26 December 2012 / Published online: 16 January 2013

Abstract Diacylglycerol acyltransferase (DGAT) catalyzes the acyl-CoA-dependent acylation of *sn*-1,2-diacylglycerol to form seed oil triacylglycerol (TAG). To understand the features of genes encoding soybean (*Glycine max*) DGATs and possible roles in soybean seed oil synthesis and accumulation, two full-length cDNAs encoding type 1 diacylglycerol acyltransferases (*GmDGAT1A* and *GmDGAT1B*) were cloned from developing soybean seeds. These coding sequences share identities of 94 % and 95 % in protein and DNA sequences. The genomic architectures of *GmDGAT1A* and *GmDGAT1B* both contain 15 introns and 16 exons. Differences in the lengths of the first exon and most of the introns were found between *GmDGAT1A* and *GmDGAT1B* genomic sequences. Furthermore, detailed in silico analysis revealed a third predicted *DGAT1*, *GmDGAT1C*. GmDGAT1A and GmDGAT1B were found to have similar activity levels and substrate specificities. Oleoyl-CoA and *sn*-1,2-diacylglycerol were preferred substrates over vernoloyl-CoA and *sn*-1,2-divernoloylglycerol. Both transcripts are much more abundant in developing seeds than in other tissues including leaves, stem, roots, and flowers. Both soybean *DGAT1A* and *DGAT1B* are highly expressed at developing seed stages of maximal TAG accumulation with *DGAT1B* showing highest expression at somewhat later stages than *DGAT1A*. *DGAT1A* and *DGAT1B* show expression profiles consistent with important roles in soybean seed oil biosynthesis and accumulation.

Keywords *Glycine max* · Triacylglycerol metabolism · DGAT1 · Genomic architecture · Gene expression · DGAT specificity

Introduction

Soybean (*Glycine max* L. Merr) is a major oilseed crop, contributing about 28 % of the world's vegetable oil production (USDA 2012a, b). In addition to its direct usage for human and animal consumption, soy oil is a major renewable resource for biofuel and chemical production, but the biosynthesis of soy oil has yet to be elucidated.

Most seed oils, including soybean, are stored largely in the form of triacylglycerol (TAG) as a major carbon and energy reserve (Bewley and Black 1994). In plant seeds, TAG biosynthesis occurs in certain membranes of the endoplasmic reticulum (ER) and accumulates in oil bodies that are generated through budding of the outer ER membrane (Huang 1992). TAG bioassembly is traditionally thought to be catalyzed by the membrane-bound enzymes of the Kennedy pathway, which sequentially transfer acyl chains from acyl-CoAs to *sn*-1,-2 and -3 positions of a glycerol backbone (Ohlrogge and Browse 1995). New evidence indicates the importance of phosphatidylcholine in generating the diacylglycerol (DAG) pool for TAG biosynthesis (Lu et al. 2009). Diacylglycerol acyltransferase (DGAT) (EC 2.3.1.20) functions in the final step of the pathway by transferring an acyl group from acyl-CoA to the *sn*-3 position of *sn*-1,2-diacylglycerol. DGAT has been proposed to be the rate-limiting enzyme in plant storage lipid accumulation (Ichihara et al. 1988; Perry and Harwood

Electronic supplementary material The online version of this article (doi:10.1007/s10142-012-0306-z) contains supplementary material, which is available to authorized users.

R. Li · Y. Wu
Shanxi Agricultural University, Taigu 030801, China

T. Hatanaka
Kobe University, Kobe 657-8501, Japan

R. Li · K. Yu · Y. Wu · H. Fukushige · D. Hildebrand
University of Kentucky, Lexington, KY 40546-0312, USA

D. Hildebrand (✉)
Department of Plant and Soil Science, University of Kentucky,
1405 Veterans Dr.,
Lexington, KY 40546, USA
e-mail: dhild@uky.edu

1993). It was discovered that many organisms have two distinct classes of DGATs with no homology to each other, designated DGAT1 and DGAT2 (Cases et al. 1998; Oelkers et al. 1998; Routaboul et al. 1999; Zou et al. 1999) that are non-redundantly responsible for the bulk of TAG synthesis. For example, in the model plant *Arabidopsis thaliana*, only mutations in DGAT1 and phosphatidylcholine—diacylglycerol acyltransferase (PDAT) but not DGAT2—have been reported to affect seed oil levels (Katavic et al. 1995; Routaboul et al. 1999; Zhang et al. 2009; Zou et al. 1999). DGAT2 in some plants, however, appears to incorporate unusual fatty acids in the seed oils (Burgal et al. 2008; Shockey et al. 2006). We found both DGAT1 and DGAT2 contribute to unusual fatty acid accumulation in seed oil (Li et al. 2010a, b). Lardizabal et al. (2008) reported that transgenic expression of a modified fungal DGAT2 increased total oil level by 1.5 % in soybean seeds. Taylor et al. (2009) reported that overexpression of *A. thaliana* and *Brassica napus* DGAT1s in canola can increase oil contents from 2.5 % to 7 %.

In addition to DGAT1 and DGAT2, several other enzymes are identified to synthesize TAG. Recently, Durrett et al. (2010) reported that a distinct DGAT (EaDAcT) with *sn*-3 acetyltransferase activity could synthesize 3-acetyl-1,2-diacyl-*sn*-glycerols (unusual, reduced viscosity oils) in *Euonymus* and transgenic Arabidopsis seeds. Another type of DGAT is a soluble DGAT enzyme identified from peanut *Arachis hypogaea* (Saha et al. 2006), but functional orthologs have yet to be characterized in other species. A bifunctional DGAT/wax ester synthase ADP1 from the bacterium *Acinetobacter calcoaceticus* can use both fatty alcohols and DAGs as acyl acceptors to synthesize wax esters and TAGs, respectively (Kalscheuer and Steinbuchel 2003; Stoveken et al. 2005). ADP1 homologs are also identified in petunia and Arabidopsis, whereas these proteins exhibit either absent or lower DGAT activity compared to the wax synthase activity (King et al. 2007; Li et al. 2008). Phospholipid/diacylglycerol acyltransferase (PDAT) (EC2.3.1.43), an acyl-CoA-independent enzyme, was found to transfer *sn*-2 acyl group from a phospholipid such as phosphatidylcholine or phosphatidylethanolamine to the *sn*-3 position of DAG for TAG formation in plants and yeast (Dahlqvist et al. 2000; Oelkers et al. 2000). PDAT and DGAT1 showed overlapping functions in Arabidopsis TAG biosynthesis (Zhang et al. 2009).

Settlage et al. (1998) showed a strong positive correlation between DGAT activity and the rate of oil accumulation in soybean seeds, suggesting important roles of DGAT in soy oil synthesis. Recently, a full-length cDNA of a soybean DGAT was cloned, and the data of its expression pattern and sequence polymorphism among various genotypes were reported by Wang et al. (2006); however, its expression in developing embryos during seed development or its functional activity was not addressed critically. Further studies are needed to elucidate detailed roles of DGAT in soy oil accumulation. There is growing interest in increasing soy oil content and engineering soybeans to produce high levels of unusual fatty acids with industrial applications (Li et al. 2010a, b; Yu et al. 2006, 2008). Here, we present the cloning of two full-length cDNAs encoding type 1 diacylglycerol acyltransferases from soybean, designated *GmDGAT1A* and *GmDGAT1B*. Their TAG biosynthesis activities were demonstrated using a yeast expression system, and their expression profiles in different organs and during seed development in relation to TAG accumulation characterized. The genomic architectures of these two DGAT1 genes were also analyzed and compared with other known DGAT1s from soybean. The current data suggest roles of GmDGAT1A and GmDGAT1B in soybean oil accumulation.

Materials and methods

General experimental procedures

Trizol used for RNA isolation, SuperScript II RT kit for reverse transcription-polymerase chain reaction (RT-PCR) analyses and yeast expression vector pYES2 were from Invitrogen, CA. The Smart RACE cDNA amplification kit for RACE (rapid amplification of cDNA ends) was from BD Biosciences, NJ. A gel extraction kit for DNA extraction from agarose gels was from Qiagen, Inc., CA. The pGEM-T Easy vector for subcloning of cDNAs was from Promega, WI. The BigDye Terminators v3.1 Cycle Sequencing Kit for DNA sequencing was from Applied Biosystems, CA. The PCR DIG Probe Synthesis Kit for Southern Blot analyses was from Roche Applied Science, IN.

Biological materials

Soybean, cv. "Jack," seeds were planted year round in a greenhouse at the University of Kentucky in Lexington, KY with two to three times per day of watering depending on the weather conditions. The yeast strain INVSc1 and *Escherichia coli* strain DH5α for yeast and *E. coli* transformation were from Invitrogen, CA.

Dicot embryogenesis is divided into five general stages: globular, heart, cotyledon, maturation, and dormancy. In this study, we focused on cotyledon to maturation stages and further classified this maturation phase into six different developmental stages (Fig. 1b) according to days after flowering (AF), seed fresh weight (milligrams), and seed morphology. Six stages are simply assigned as stage 1 (12 days AF/30–70 mg), stage 2 (25 days AF/100–150 mg), stage 3 (35 days AF/200–250 mg), stage 4 (45 days AF/300–350 mg), stage 5 (55 days AF/400–

480 mg), and stage 6 (yellowing and beginning dry down) (65 days AF/360–300 mg).

Cloning of GmDGAT1 cDNAs

For soybean DGAT cloning, a BLAST search of the sequence database using the Arabidopsis protein sequence identified soybean expressed sequence tag (EST) (Gm-c1036-7949). The EST was fully sequenced in both directions. Since the EST lacked the 5′ end of the cDNA, it was obtained by 5′ RACE with appropriate nested primers using a Smart RACE cDNA Amplification kit (Clontech, BD Biosciences). A cDNA was synthesized from poly(A)+ RNA of developing seeds of soybean (cultivar, "Jack"). We then designed the following two primers from the sequence information of the cDNA of the EST; SoyD5-1: 5′-GCGTAAAGAAGG TTTCCCTTGAGAGGATGC-3′ and SoyD3-1: 5′-GTTGCCCCTACATTATGTTACCAGCCAAGC-3′. The 5′-half and 3′-half of the cDNAs were amplified using the PCR conditions described in the user manual of the kit. In order to obtain other possible DGAT sequences, another set of primers were designed: SoyD5-2: 5′-GAAAACACGCTCGGTC TTCTTC-3′ and SoyD3-2: 5′-TACAATTGCCAGAGGA GAGTTG-3′. Fractionation of the amplified fragments (1.5 kb), cloning and sequencing were carried out as described above.

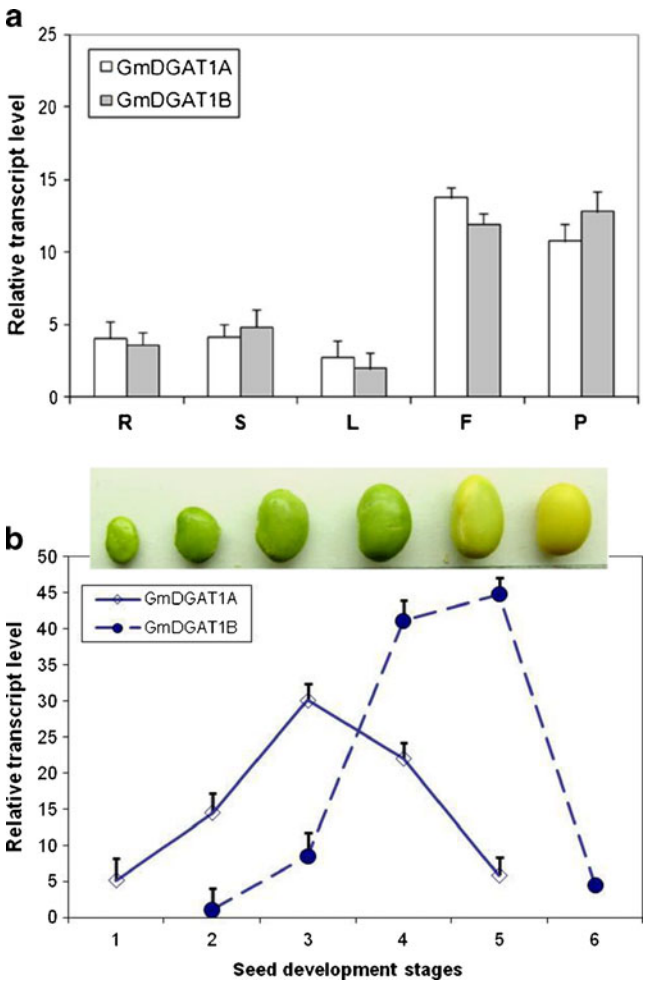

Fig. 1 Expression profiles of *GmDGAT1A* and *GmDGAT1B* genes in different organs and seed developmental stages of soybean plants by quantitative real-time PCR. **a** *GmDGAT1* gene expression in young roots (*R*), stems (*S*), leaves (*L*), flower buds (*F*), and young pods (containing developing seeds) (*P*) at 20 days AF. **b** *GmDGAT1* gene expression during seed development. Total RNA were extracted from different organs and developing seeds. The first strain *cDNA* was used as template to amplify the target gene. The *actin* gene was amplified as an internal control. The seed development fresh weights were: stage 1, 30–70 mg; stage 2, 100–150 mg; stage 3, 200–250 mg; stage 4, 300–350 mg; stage 5, 400–480 mg; and stage 6, 360–300 mg

Southern blot analysis

Genomic DNA of soybean was isolated from young leaves using a modified CTAB (*N*-cetyl-*N*,*N*,*N*-trimethylammonium bromide) procedure as described previously (Hatanaka et al. 2004). Aliquots of genomic DNA (10 μg) were digested overnight with four restriction enzymes, *Xba*1, *Bam*H1, *Eco*RV, and *Not*1, individually. The digested DNA was fractionated in a 0.8 % (*w*/*v*) agarose gel and transferred to a positively charged nylon membrane (Hybond N+, Amersham Biosciences, NJ) overnight in 20× SSC (3 M NaCl, 0.3 M sodium citrate, pH 7.0). The membrane was hybridized to a digoxigenin (DIG)-labeled probe representing the protein encoding region of *GmDGAT1A* and *1B* cDNAs. The membrane was washed with 2× SSC, 0.1 % SDS; 0.2× SSC, 0.1 % SDS, and 0.1× SSC, 0.1 % SDS for 15 min at 65 °C. The hybridized DNA was detected with alkaline phosphatase conjugated anti-DIG antibody and its chemiluminescent substrate, CDP-Star, following the manufacturer's protocol.

Quantitative real-time PCR

Total RNA was isolated from young roots, stems, leaves of three-leaf soybean seedlings, and flower buds and developing seeds at six developmental stages. After extraction, RNA samples were treated with DNaseI (Promega) to remove contaminating DNA. First-strand cDNA was synthesized using equal amounts of RNA as templates following the manufacturer's instructions. PCR controls were performed in the absence of added reverse transcriptase to ensure RNA samples were free of DNA contamination. All real-time reactions were performed in an iCycler iQ detection system (Bio-Rad) using the intercalation dye SYBR Green I Master Mix kit (Applied Biosystems) as a fluorescent reporter.

PCR reactions were performed in triplicate in 25-μL volumes using 1 μL of each forward and reverse primer (500 nM), 12.5 μL of SYBR green master mix, 5 μL of a 1:10 (*v*/*v*) dilution of cDNA, and 5.5 μL of DEPC water.

Reactions were performed in MicroAmp 96-well plates (Applied Biosystems) covered with optical adhesive covers (Applied Biosystems). The following program was applied: initial polymerase activation at 95 °C for 10 min, then a two-temperature thermal cycle consisting of denaturation at 95 °C for 15 s, followed by annealing extension at 60 °C for 1 min, for a total of 40 cycles.

The quantification of PCR products was performed via a calibration curve procedure using *actin* RNA as an internal standard. PCR products were analyzed using melting curves as well as agarose gel electrophoresis to ensure single product amplification. The ratio of gene-specific expression to *actin* signal was defined as a relative expression. Primers for specific amplification of each cDNA were designed using the Primer Express software (Applied Biosystems), taking into account criteria such as product length (around 500 bp), optimal PCR annealing temperature, and likelihood of primer self-annealing. The primers for *GmDGAT1A* were 5′-GAAGAGAAGACTGAGTTAGTAAACACG-3′ (forward) and 5′-ACAACTGGTTCGGGTATAACATTCTGC-3′ (reversed). The primers for *GmDGAT1B* were 5′-ACTCTTCCCTGCGCCGCCGTCCC-3′ and 5′-ACAACTGGTTCGGGTATACACTTCCGT-3′ (reverse). The primers for the *actin* gene were 5′-AAGCTGTTCTCT CCTTGTACGCC-3′ (forward) and 5′-GCACAGTGTGAGA CACACCATCA-3′ (reverse).

Yeast expression of GmDGAT1A and GmDGAT1B

The *ORF* of *GmDGAT1A* and *1B* were cloned into the pYES2 vector as described previously for *Vernonia* DGAT1A and *Vernonia* DGAT1B except that a Kozak translation initiation sequence of ANNATGG was generated after PCR of the target genes based on the vector manufacturer's recommendations (Yu et al. 2008). Because the *Vernonia* DGAT1A and *Vernonia* DGAT1B happen to contain a Kozak translation initiation sequence, the manipulation as described above was not needed for their yeast vector construction. After the generation of the Kozak translation initiation sequence for *GmDGAT1A* and *GmDGAT1B*, the ORFs were not changed but the −3 bp was changed from "C" to "A" for both genes. Yeast transformation, confirmation of the transformed yeast, yeast microsomal extraction and microsome protein determination were performed as described previously for *Vernonia* DGAT1s (Yu et al. 2008).

Yeast microsomal extraction and DGAT in vitro activity assays were also performed as described previously with some modifications (Yu et al. 2008). Microsomal protein concentrations were determined by a modified Lowry method (Wang et al. 2005). The substrate combinations used for the yeast microsomal assays were also [^{14}C]oleoyl-CoA or [^{14}C]Va-CoA with *sn*-1,2-dioleoylglycerol (DODAG) or *sn*-1,2-divernoloylglycerol (DVDAG). From a preliminary study, we found that the highest DGAT activity for GmDGAT1A and GmDGAT1B was also from the substrate combination of [^{14}C]oleoyl-CoA with *sn*-DODAG. Therefore, this substrate combination was again used to determine the linear range of microsomal protein levels for our yeast microsomal assays. The following microsomal protein levels for each assay were used for the linear range determination: 10, 20, 40, 80, 160, 320, 640 (preceding numbers in nanograms), 1.28, 2.56, 5.12, 10.24, 20.48, and 40.96 (preceding numbers in micrograms). Since we found that the level of 40-ng microsomal protein was within the linear range of DGAT activity response for the highest substrate combination, it was used for microsomal DGAT assays. The lipids from the microsomal assay reactions were extracted and analyzed as previously described (Yu et al. 2006). For the linear range of microsomal protein level analysis, there were three replicates for each treatment. For microsomal DGAT assays, there were five replicates for each treatment, and the assays were performed three times each using a separate batch of isolated microsomes. Statistical analyses were performed using SAS PROC GLM and least significant difference means separation after a significant ANOVA (Carmer and Swanson 1973).

Results

Isolation and sequence analysis of two cDNA clones encoding type 1 DGATs from soybean

To clone full-length cDNAs encoding soybean DGAT1(s), a BLAST search of the sequence database using the Arabidopsis *AtDGAT1* (AJ131831) amino acid sequence identified a soybean EST homolog (GenBank no. BM309699). Based on the partial sequence of this EST cDNA, two pairs of primers were designed to isolate the 5′- and 3′-ends of candidate soybean DGAT cDNAs from the total RNAs prepared from developing seeds of soybean (cv. "Jack") using a RACE protocol. Finally, two full-length cDNAs encoding soybean DGAT1 were obtained, designated *GmDGAT1A* and *GmDGAT1B*, and deposited in GenBank under accession numbers AB257589 and AB257590 (deposited 17 June 2006).

GmDGAT1A is 1,888 bp in length with a 59-bp 5′-leader sequence and 333 bp of the 3′-untranslated region (UTR). The complete open reading frame (ORF) of 1,497 bp encodes a protein of 498 amino acids (Fig. 2) with a predicated molecular weight of 57.23 kDa and a calculated isoelectric point of 8.96 (http://www.expasy.ch/cgi-bin/pi_tool). The full-length of G*mDGAT1B* is 1,960 bp, containing a 55-bp 5′-leader sequence and 391 bp 3′-UTR. The *GmDGAT1B* ORF of 1,515 bp is predicted to encode a protein of 504 amino acids (Fig. 2) with a theoretical pI/MW of 8.88/58.02 kDa. GmDGAT1A and 1B share 95.8 % identity with only 19 amino acid differences+six gaps throughout the coding regions (Fig. 2). Nine of these

changes are within the same amino acid groups based on the polarity and charge, similar to what was reported by Wang et al. (2006) for the amino acid changes in DGATs within 14 soybean accessions (Wang et al. 2006). GmDGAT1A is missing three amino acid residues at position 24 and three more between positions 60 and 70 in the N-terminal region compared to GmDGAT1B (Fig. 2).

In order to further investigate genomic locations and copy number of *GmDGAT1A* and *1B*, a BLAST search (http://www.phytozome.net/search.php?show=blast) was performed using *GmDGAT1* cDNAs and predicted amino acid sequences. A draft of the soybean genome is now available (Schmutz et al. 2010). ORFs corresponding to *GmDGAT1A* and *1B* are Glyma13g16560 and Glma17g06120, respectively. Amino acid identities are 99.6 % (496/498) for GmDGAT1A and Glyma13g16560, and 99.8 % (503/504) for GmDGAT1B and Glyma17g06120. GmDGAT1A and 1B are clearly two different loci. *GmDGAT1A* is located on chromosome 13, whereas *GmDGAT1B* is on chromosome 17. Only two amino acid residues were different between GmDGAT1A and Glyma13g16560. Amino acids at positions 366 and 479 are glycine (G^{366}) and histidine (H^{479}) for GmDGAT1A while amino acids at the same positions are aspartic acid and glutamine for Glyma13g16560. Glyma13g16560 coding sequence is an exact match with that of AY496439. Similarly, an amino acid difference at the 406th position was found for GmDGAT1B (tyrosine) and Glyma17g06120 (phenylalanine). These differences are likely due to allelic differences between the different genotypes from which these sequences were obtained.

Though we did not isolate additional clones other than GmDGAT1A and 1B using the primers designed from the EST (BM309699), a recent BLAST search of soybean genomic sequences posted at Soybase (http://www.soybase.org/) with *AtDGAT1* identified two additional ORFs with DGAT annotations: Glyma09g07510 and Glyma09g07520. The Glyma09g07510 sequence is very short (859 bp) and we originally assumed it was a pseudogene. We tentatively designated Glyma09g07520 as *GmDGAT1C*; however, an alignment of the deduced amino acid sequence from *GmDGAT1C* showed a close alignment with a partial cDNA from another soybean DGAT1 (GenBank accession no. AY652765) while the former is truncated at the C-terminus end indicating a potential error in the gene prediction for Glyma09g07520. A further BLAST search of soybean genomic sequences with AY652765 showed the 3′-end of AY652765 matches the

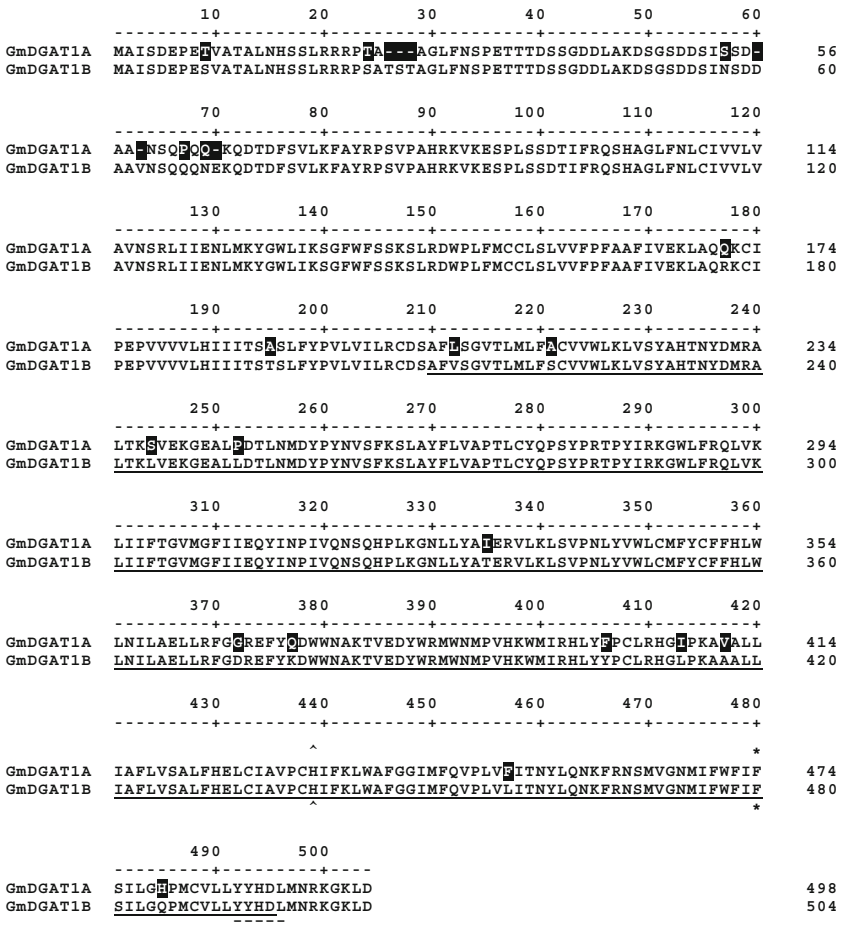

Fig. 2 Alignment of deduced amino acid sequences of GmDGAT1A and GmDGAT1B. The alignment was generated by Lasergene (DNAStar) with the CLUSTAL W method. The amino acid differences and gaps between GmDGAT1A and GmDGAT1B are *shaded in black*. A conserved phenylalanine (*F*) and the putative active site histidine (*H*) located near the C-terminus is indicated by an *asterisk* and an *accent*. The MBOAT domain and an ER retrieval motif are *underlined* by a *solid line* and a *dashed line*

sequence of Glyma09g07510 indicating that Glyma09g07510 and Glyma09g07520 together encode a predicted *DGAT1C* with four additional exons in an intergenetic region. A BLAST search of the GenBank soybean EST database with Glyma09g07510 sequence identified two ESTs (GenBank no. CO981711 and EH038988), sequences of which contain those of Glyma09g07510 and Glyma09g07520 and confirmed the new sequence assignment of the putative *DGAT1C*. Amino acid identities between DGAT1C and Glyma13g16560 (DGAT1A) or Glma17g06120 (DGAT1B) are 79.5 % or 78.9 %, respectively.

Genomic organization of GmDGAT1s

In order to investigate the copy numbers of DGAT1A and 1B genes in the soybean genome, Southern blot analysis was employed. Genomic DNA from soybean cv "Jack" was digested with the restriction enzymes *Xba*1, *BamH* 1, *EcoR* V, and *Not*1. The entire coding regions of GmDGAT1A and 1B cDNAs were labeled by DIG and used as probes. The results are shown in Fig. 3, suggesting each of them exist as single copies in the soybean genome because only one major hybridization band was detected. Similarly, using a 0.8-kb fragment of *GmDGAT* as a probe, Wang et al. (2006) performed Southern blot with *Taq*I-digested soybean genomic DNA and found two major bands and a few weaker bands consistent with multiple *Taq*I sites.

For understanding *GmDGAT1* genomic structures, we subsequently cloned genomic DNA sequences for *GmDGAT1A* and *GmDGAT1B*. The genomic DNA length of *GmDGAT1A* is 7,701 bp while the *GmDGAT1B* is 7,964 bp, both of which are larger than that of *AtDGAT1* (3,020 bp), *OsDGAT1* (6,220 bp) and *LjDGAT1* (5,759 bp) and the possible *DGAT1* (7,311 bp) of *Medicago truncatula* (AC174465) (Fig. 4). Their corresponding loci, Glyma13g16560 and Glyma17g06120, are 7,813 bp and 7,698 bp in the soybean genotype (cv. William 82). All four soybean *DGAT1s*, as well as a predicted *GmDGAT1C* (*Glyma17g07510+Glyma17g07520*), have 15 introns and 16 exons. Fifteen of the exons (no. 2 to no. 16) are of the same length and have similar sequences for all five *GmDGAT1s*. Only exon no. 1 shows a difference in length with 300 bp for *GmDGAT1A*/Glyma13g16560, 318 bp for *GmDGAT1B*/Glyma17g06120 and 357 bp for Glyma17g07510+Glyma17g07520 (Fig. 4).

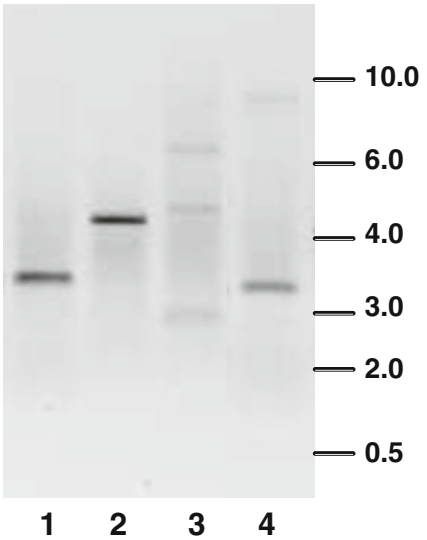

Fig. 3 Southern blot of soybean *GmDGAT1* genes. The genomic DNA was digested with restriction enzymes *Xba*l (lane *1*), *Bam*H1 (lane *2*), *Eco*RV (lane *3*), and *Not*1 (lane *4*). The DNA blot was hybridized with a dioxigenin-labeled *cDNA* encoding the ORF of *GmDGAT1s* as a prob. The blot was washed at high stringency after hybridization at 0.1× SSC/0.1 % SDS at 65 °C

Identification of other genes involved in TAG formation from soybean genome databases

A summary of information available on DGAT1s, DGAT2s, and PDAT1s in the soybean genome is shown in Table 1 and at our Soybean Lipid Biosynthesis Pathway website (http://www.uky.edu/SoyLipidGenes). The EST data indicate that all three *GmDGAT1s* are transcribed. BLAST analysis of the soybean draft genome sequence data using *AtDGAT2* (At3g51520), also reveals five possible soy DGAT2 genes, Glyma09g32790 (65.1 % identity, 80.5 % positive) (*GmDGAT2A*), Glyma16g21960 (63.8 % identity, 79.9 % positive) (*GmDGAT2B*), Glyma16g21970 (57.6 % identity, 77.1 % positive) (*GmDGAT2C*), Glyma01g36010 (71.3 % identity, 84.4 % positive) (*GmDGAT2D*) and Glyma11g09410 (47.9 % identity, 61.3 % positive) (*GmDGAT2E*) (Table 1). These proteins are annotated as diacylglycerol O-acyltransferase. A third type of DGAT is reported in peanut (Saha et al. 2006) that is now designated as DGAT3. BLAST analysis of the soybean genome with the peanut (*A. hypogaea*) AhDGAT3 indicates two DGAT3s, Glyma17g04650 (*GmDGAT3A*) and Glyma13g17860 (*GmDGAT3B*) (Table 1).

The same approach using AtPDAT1 (At5g13640) detected six possible PDATs in the soybean genome, Glyma17g05910 (76.9 % identity, 88.0 % positive) (*GmPDAT1A*), Glyma13g16790 (76.9 % identity, 87.7 % positive) (*GmPDAT1B*), Glyma07g04080 (76.9 % identity, 87.8 % positive) (*GmPDAT1C*), Glyma16g00790 (77.7 % identity, 87.8 % positive) (*GmPDAT1D*), Glyma12g08920 (53.5 % identity, 67.6 % positive) (*GmPDAT1E*), and Glyma11g19570 (50.2 % identity, 63.9 % positive) (*GmPDAT1F*) (Table 1). These proteins are annotated as phosphatidylcholine-sterol O-acyltransferases. The EST data indicate GmPDAT1A–GmPDAT1D are transcribed, but no EST data were uncovered for GmPDAT1E and GmPDAT1F. To date, no data are available of possible functions of these additional TAG biosynthetic enzyme encoding genes in soybean biology.

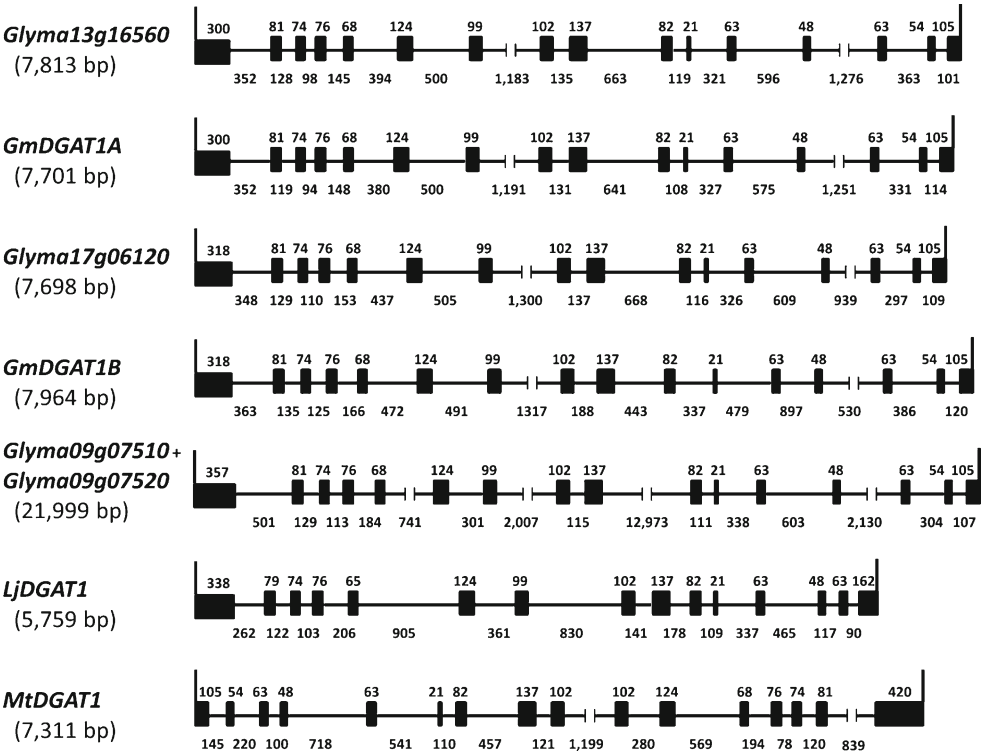

Fig. 4 Comparison of the genomic structure of Gm*DGAT1*s and *L. japonicus* and *M. truncatula* DGAT1s. The *filled boxes* indicate exons, and the *lines* indicate introns. The size (base pair) of each intron/exon was indicated by *numbers*. The full-length genomic sequences of *Glyma13g16560*, *Glyma17g06120*, *Glyma09g07510*+*Glyma09g07520*, and *M. truncatula DGAT1* (AC174465) were obtained from soybean and *Medicago* genome sequence databases. The genomic sequence of *L. japonicus* was obtained from PlantdGDB (http://www.plantgdb.org/LjGDB/cgi-bin/getRecord.pl?dbid=0;resid=5;chrUID=173761)

Expression analysis of GmDGAT1s in different organs and seed development

To further investigate the potential role of GmDGAT1A and B in soybean oil accumulation, we analyzed the temporal and tissue-specific expression patterns of both *GmDGAT1A* and *GmDGAT1B* transcripts by quantitative real-time PCR using an equal amount of total RNA from the sample tissues (Fig. 1). The *actin* gene, a plant housekeeping gene, was used as an internal control. Both *GmDGAT1s* show low expression in flowers (not opened) and pods (excluding seeds) with *GmDGAT1A* higher in flowers and *GmDGAT1B* higher in pods (Fig. 1a). Weak expression is also detected in young roots, stems, and leaves (Fig. 1a). However, the two genes are highly expressed in developing soybean seeds (Fig. 1b). Notably, *GmDGAT1A* is mainly expressed in earlier stages (stages 2, 3, 4) while *1B* is expressed maximally somewhat later (stages 3, 4, 5). Similarly, Wang et al.'s (2006) data also indicated soy DGAT (e.g., soy *DGAT1A*) expression was high at a relatively early stage (30 DAF, equivalent to the period between stages 2 and 3 here) of pod development.

Although TAG synthesis mainly occurs in seeds, fatty acid synthesis also occurs in other plant tissues. Our previous study showed that the maximal increase of soy oil was observed from stage 3 to stage 5 (Li et al. 2010a, b). Both soy DGAT1A and B expression patterns closely correlate with soybean oil accumulation suggesting that GmDGAT1s are major contributors to TAG synthesis and oil accumulation in developing soybean seeds with GmDGAT1B being possibly more important in this regard.

Functional characterization of GmDGAT1A and GmDGAT1B by yeast expression and in vitro assay

The coding regions of *GmDGAT1A* and *GmDGAT1B* were transferred into the yeast expression vector pYES2 under the control of the galactose-inducible *GAL1* promoter, and a Kozak translation initiation sequence was incorporated based on the vector manufacturer's recommendations. The construct was used to transform yeast. Yeast cells harboring an empty pYES2 vector plasmid were used as controls. Yeast microsomes containing the DGAT protein were isolated from the transformed yeast cells and then used to analyze the DGAT activity and substrate specificity in vitro.

For accurate microsomal DGAT assay results, the substrate combination of [^{14}C]oleoyl-CoA and *sn*-DODAG was used for both GmDGAT1A and GmDGAT1B to determine the linear range of microsomal protein levels for yeast microsomal assays (Fig. 5). The linear microsomal protein range was analyzed using 10 ng to 40.96 μg of microsomal protein levels. At the levels of microsomal protein from 10 to 640 ng, GmDGAT1A and GmDGAT1B activities proportionally increased as the microsomal protein level increased as shown in Fig. 5. To leave some margin, the 40-ng microsomal protein level (data point no. 6 in Fig. 5; encircled) was used for the subsequent DGAT activity and substrate specificity analyses on GmDGAT1A and GmDGAT1B along

Table 1 Putative DGATs and PDATs in the soybean genome (see also: http://www.uky.edu/SoyLipidGenes)

Locus ID	Abbreviation	Subcelluarlocation[a]	Evidence of activity	ESTs from dbEST/NCBI												
				Plants at various developmental stages	Globular-stage embryos	Green seeds	Seed coats	Germinating shoots	Detriorating cotyledons	One or 3-week old whole seedlings	Leaves	Stems	Roots	Immature flowers	Flowers	Young pods
Glyma13g16560	GmDGAT1A	Endomembrane	Activity in yeast			+			+			+				+
Glyma17g06120	GmDGAT1B	Endomembrane	Activity in yeast									+	+			
Glyma09g07520	GmDGAT1C	Endomembrane	Sequence similarity to AtDGAT1							+						
Glyma09g32790	GmDGAT2A	Endomembrane	Sequence similarity to AtDGAT2							+	+					
Glyma16g21960	GmDGAT2B	Endomembrane	Sequence similarity to AtDGAT2						+	+	+		+	+		
Glyma16g21970	GmDGAT2C	Endomembrane	Sequence similarity to AtDGAT2	+								+				
Glyma01g36010	GmDGAT2D	Endomembrane	Sequence similarity to AtDGAT2	+	+		+			+						
Glyma11g09410	GmDGAT2E	Endomembrane	Sequence similarity to AtDGAT2	+												
Glyma17g04650	GmDGAT3A	Unknown	Sequence similarity to AhDGAT3			+			+			+	+		+	+
Glyma13g17860	GmDGAT3B	Unknown	Sequence similarity to AhDGAT3										+			
Glyma17g05910	GmPDAT1A	Endomembrane	Sequence similarity to AtPDAT1	+[b]				+[b]								
Glyma13g16790	GmPDAT1B	Endomembrane	Sequence similarity to AtPDAT1	+				+								
Glyma07g04080	GmPDAT1C	Endomembrane	Sequence similarity to AtPDAT1			+		+		+	+	+			+	+
Glyma16g00790	GmPDAT1D	Endomembrane	Sequence similarity to AtPDAT1		+					+				+		
Glyma12g08920	GmPDAT1E	Endomembrane	Sequence similarity to AtPDAT1	N/F												
Glyma11g19570	GmPDAT1F	Endomembrane	Sequence similarity to AtPDAT1	N/F												

Locus ID	ESTs from dbEST/NCBI															
	Senescing leaves	Callus grown in dark	Apical meristems	Drought-stressed leaves	Salt-stressed leaves	*Pseudomonas*-infected leaves	Soybean rust-infected leaves	*Phytophythora sojae*-infected hypocotyls	Salicylic acid-treated seedlings	Hypersensitive response-induced seedlings	Drought-stressed roots	Tissue culture suspensions	Etiolated seedlings	Auxin-treated	Somatic embryos on MSD20	
Glyma13g16560		+	+												+	
Glyma17g06120				+	+	+						+	+			
Glyma09g07520							+									
Glyma09g32790									+							
Glyma16g21960				+					+							
Glyma16g21970				+	+	+						+	+			
Glyma01g36010																
Glyma11g09410																
Glyma17g04650	+	+	+				+			+	+					
Glyma13g17860															+	

Table 1 (continued)

Locus ID	ESTs from dbEST/NCBI													
	Senescing leaves	Callus grown in dark	Apical meristems	Drought-stressed leaves	Salt-stressed leaves	*Pseudomonas*-infected leaves	Soybean rust-infected leaves	*Phytophthora sojae*-infected hypocotyls	Salicylic acid-treated seedlings	Hypersensitive response-induced seedlings	Drought-stressed roots	Tissue culture suspensions	Etiolated seedlings	Auxin-treated	Somatic embryos on MSD20
Glyma17g05910															
Glyma13g16790															
Glyma07g04080			+					+			+			+	
Glyma16g00790															
Glyma12g08920															
Glyma11g19570															

EST information was obtained from GenBank EST databases. Glyma13g16560=GmDGAT1A; Glyma17g06120=GmDGAT1B. Sequence similarity is not good evidence of activity. AtDGAT2 apparently is not active at least in developing seeds. Some GmDGATs may have no activity in vivo as well

NCBI National Center for Biotechnology Information, *N/F* not found, *AhDGAT3* peanut DGAT3, *AtDGAT1* Arabidopsis DGAT1, *AtDGAT2* Arabidopsis DGAT2, *CDS* coding sequence

[a] Predicted

[b] The nucleotide sequence identity between GmPDAT1A and GmPDAT1B is higher than 95 %, and the ESTs assigned to GmPDAT1B could potentially be of GmPDAT1A

with the vector control. This level is three magnitudes lower than 50–100 μg used by other studies (Bouvier-Nave et al. 2000; He et al. 2004; Kroon et al. 2006). The DGAT assays using 40-ng microsomal protein equivalents have greatly reduced the background on the phosphorimages. Also, microsome levels needed for each reaction is greatly reduced.

As shown in Fig. 6, GmDGAT1A and GmDGAT1B exhibited significantly higher activity relative to the vector control at the linear range assay conditions. The activities of the two soybean DGAT1s are similar. Among the substrates examined, activity was highest with the 18:1-CoA/*sn*-DODAG substrate combination and lowest with vernoloyl-CoA/*sn*-DVDAG. The two substrate combinations of vernoloyl-CoA/*sn*-DODAG and 18:1-CoA/*sn*-DVDAG gave intermediate activity. Within these middle levels, the substrate combination of 18:1-CoA/*sn*-DVDAG had higher DGAT activity than the substrate combination vernoloyl-CoA/*sn*-DODAG. Vernoloyl-CoA and *sn*-DVDAG were selected as substrates in the assay to identify if the two DGATs can preferentially incorporate substrates with vernolic moieties into TAG. If not, other DGATs with such substrate specificity will be needed to co-express with epoxygenase so as to increase vernolic acid accumulation in the engineered soybean seeds. These results indicate that GmDGAT1s are not very helpful for vernolic acid accumulation into TAG and can explain the limited incorporation of vernolic acid into TAG when epoxygenase enzymes catalyzing vernolic acid biosynthesis are expressed in soybeans without DGATs that do have specificity for vernolic acid (Li et al. 2010a, b).

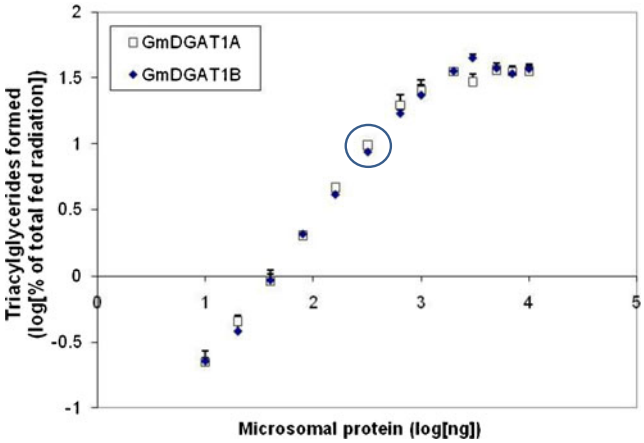

Fig. 5 Linear range of soybean DGAT1A and B activities in yeast microsomal assays. [^{14}C]Triglycerides formed in yeast microsomal assays expressing GmDGAT1A and GmDGAT1B were examined at various microsomal protein levels to determine the linear range of activity. Yeast microsomes (10 ng to 41 μg protein equivalents) were fed with 5 μM of [^{14}C]oleoyl-CoA together with 100 μM of *sn*-1,2-dioleoylglycerol. *Bars* are means±STD of three replicates

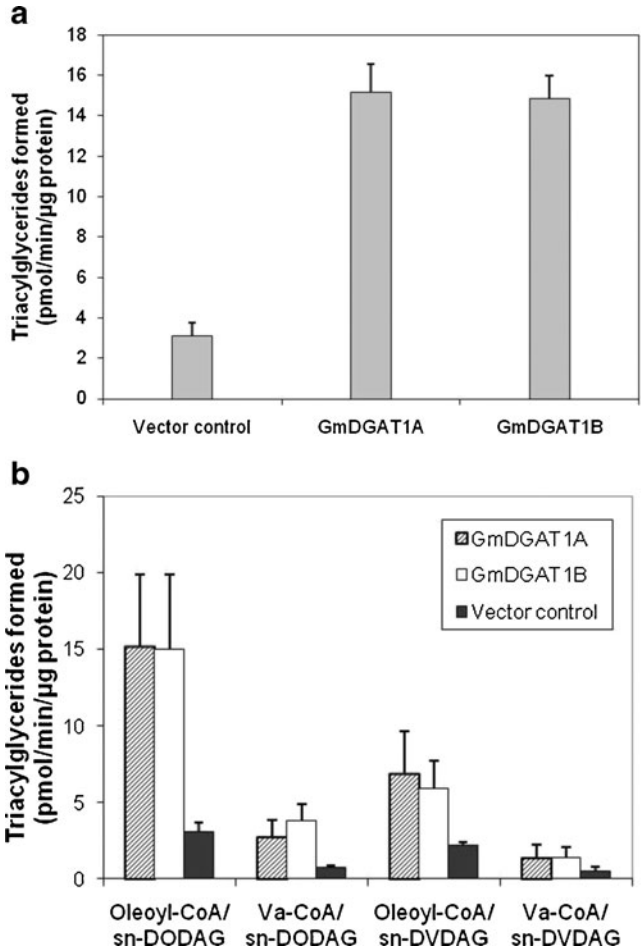

Fig. 6 GmDGAT1A and 1B enzymatic activities and substrate specificities in vitro. **a** GmDGAT1A and 1B enzymatic activities. **b** GmDGAT1A and 1B substrate specificities. [^{14}C]Triglycerides formed when microsomes from yeast expressing GmDGAT1A, GmDGAT1B, and vector control (40 ng microsomal protein equivalents) were fed 5 µM of [^{14}C]oleoyl-CoA or [^{14}C]vernoloyl-CoA (Va-CoA) together with 100 µM of *sn*-DODAG or *sn*-1,2-divernoloylglycerol. *Bars* are means±STD (n=6). Both GmDGAT1A and GmDGAT1B transformed yeast had significantly higher TAG synthesis activity than the vector control (P<.0001)

Discussion

Identification of predicted functional motifs in GmDGAT1s

In plants, DGAT1 has been shown to be localized in the ER (2006), where the Kennedy pathway mainly occurs, and to contain a number of membrane spanning domains. The putative C-terminus ER retrieval motifs found in other plant DGAT1s (e.g., tobacco DGAT1, -YYHDV-; Arabidopsis DGAT1, -YYHDL-) was also detected in GmDGAT1s (-YYHDL-) (Fig. 7a). These putative pentapeptide ER retrieval motifs (-Φ-X-X-K/R/D/E-Φ-, where Φ is any large hydrophobic amino acid residue) were originally shown to be at the extreme C-termini (McCartney et al. 2004) but demonstrated to function even upstream from their C-termini by Shockey et al. (2006). Recently, McFie et al. (2010) showed that murine DGAT1 has three transmembrane spans using a protease protection assay though several transmembrane prediction programs indicated 6 to 9 transmembrane domains. They showed that the murine DGAT1 has a short stretch at the N-terminus end exposed to cytosol while the majority of the protein including a potential active site histidine and its C-terminus end resides in lumen of ER. On the contrary, both the N- and C-terminus ends of tung DGAT1 (VfDGAT1) are predicted to be located in the cytosol by Shockey et al. (2006). Multiple alignments of murine DGAT1 and plant DGAT1s including *GmDGAT1A* and *1B* show a good similarity in the first transmembrane domain (Fig. 7a) while the putative second and third transmembrane domains differ significantly. The latter may indicate that this region, at least for plant DGAT1s, is imbedded in the membrane, but does not cross the membrane as McFie et al. (2010) proposed as an alternative model. Further detailed experiments such as protease protection assays should be conducted with plant DGAT1s in the future. Interestingly, an exception is that no transmembrane domains are predicted for a peanut cytosolic DGAT (Saha et al. 2006).

Like many other plant DGAT1s, *GmDGAT1A* and *B* belong to a membrane-bound O-acyltransferase (MBOAT) super family, and a conserved sequence is underlined in Fig. 2 (Hofmann 2000). The putative active site histidine, identified by McFie et al. (2010), is conserved among plant DGAT1s except for OsDGAT1-2 (GenBank no. AP003714) (Fig. 7a). Another conserved feature is a phenylalanine located near the C terminal (positions 460-490) among known plant DGAT1s, which was found to be an important determinant of oil content and composition in maize (Zheng et al. 2008). This phenylalanine also exists at positions 474 and 480 in GmDGAT1A and 1B (Fig. 2).

Scanning the protein sequence against the Prosite database (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_proscan.html) identified a number of putative functional motifs including leucine zipper motif, N-glycosylation, cAMP-, and cGMP-dependent protein kinase phosphorylation, protein kinase C phosphorylation, casein kinase II phosphorylation, and N-myristoylation sites for both GmDGAT1s (Table 2). The difference is that only GmDGAT1B has a predicted tyrosine kinase phosphorylation site. A notable motif among those detected is a leucine zipper motif, which is involved in protein–protein interaction, such as dimerization (van Heeckeren et al. 1992); however, a further examination of the residues other than leucine indicated that this region may not form a coiled coil structure as expected from known leucine zipper motifs (Mason and Arndt 2004). It remains to be determined whether this motif and other sites identified by a Prosite database search are important in the regulation of the functions of these enzymes in vivo. *DGAT1* from *Lotus*

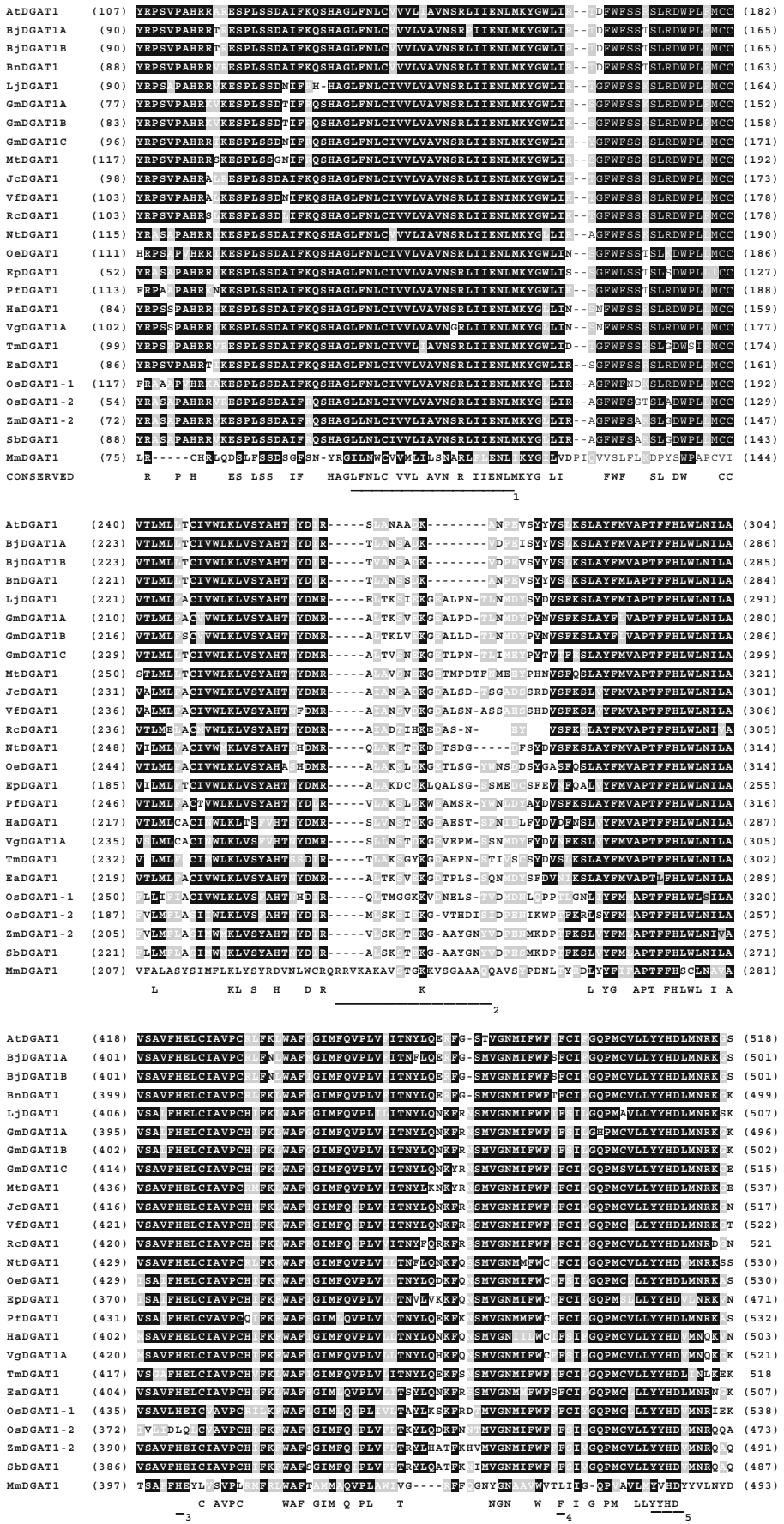

Fig. 7 Sequence comparison of known soybean DGATs and plant DGAT1s. Multiple alignment (partial). Alignments were generated by Lasergene (DNAStar) with the CLUSTAL W method. Identical amino acid residues are *shaded in black* and amino acid residues in the same functional group are *shaded in gray*. *Underlined* regions/residues are: (1) Predicted first transmembrane domain of murine DGAT1; (2) Predicted second and third transmembrane domain region of murine DGAT1; (3) the conserved histidine residue within a predicted active site; (4) the conserved phenylalanine residue thought to be involved in a high oil content; (5) the ER retrieval motif demonstrated by Shockey et al. (2006)

Table 2 Putative functional motifs in GmDGAT1A and B

Functional site	GmDGAT1A		GmDGAT1B	
	Position	Amino acid	Position	Amino acid
N-glycosylation	15–18	NHSS	15–18	NHSS
N-(Aach et al.)-[ST]-{P}	256–259	NVSF	262–265	NVSF
cAMP-/ cGMP-dependent	21–24	RRPT	21–24	RRPS
protein kinase phosphorylation				
RK](2)-x-[ST]				
Protein kinase C	18–20	SLR	18–20	SLR
phosphorylation	139–141	SSK	145–147	SSK
[ST]-x-[RK]	142–144	SLR	148–150	SLR
	258–250	SFK	264–266	SFK
			335–337	TER
Casein kinase II	34–37	TTD	37–40	TTTD
phosphorylation	39–42	SGDD	41–44	SSGD
[ST]-x(2)-[DE]	47–50	SGSD	50–53	SGSD
	142–145	SLRD	148–151	SLRD
	228–231	TNYD	234–237	TNYD
	378–381	TVED	384–387	TVED
N-myristoylation	48–53	GSDDSI	51–56	GSDDSI
G-{EDRKHPFYW}-x(2)	406–411	GIPKAV	412–417	GLPKAA
-[STAGCN]-{P}			484–489	GQPMCV
Leucine zipper pattern[a]	199–220	LRCDSAF	205–226	LRCDSAF
L-x(6)-L-x(6)-L-x(6)-L		LSGVTLM		VSGVTLM
		LFACVVWL		LFSCVVWL
Tyrosine kinase			369–376	RFGDREFY
phosphorylation				
[RK]-x(2,3)-[DE]-x(2,3)-Y				

[a] Substitution of an individual leucine by hydrophobic residues such as isoleucine and valine is tolerated (van Heeckeren et al. 1992)

japonicus (AY859489) is closely aligned with *GmDGAT1A* and *GmDGAT1B* (Supplementary Figure) on an amino acid sequence basis. This *LjDGAT1* genomic sequence also has a similar exon structure with *GmDGAT*1s except for the last exon being the combination of two exons found in *GmDGAT1*s, hence the 15 exon/14 intron structure overall. *M. truncatula*, a model legume plant species, also show the 15 intron/16 exon structure, but the organization of exon structures is completely different from that of *GmDGAT1s* and *LjDGAT1* (Fig. 4). Similarly, rice (*Oryza sativa*) and Arabidopsis *DGAT1s* (At2G19450 and AY858584) both contain 15 introns and 16 exons. The structure of 14 introns/15 exons found in *LjDGAT1* was also detected in soybean (cv. 8904) *DGAT1* (AY496439), and *Glycine latifolia* (cv. PW0031) *DGAT1* (Wang et al. 2006).

Though a similar organization is found in exons, a great variation was found in intron sequence and length between the two soy *DGAT1* genes *GmDGAT1A* and *GmDGAT1B*. For example, the tenth intron is 108 and 337 bp for *GmDGAT1A* and *1B*. Moreover, there are many small differences (1 to 7 bp) between the two DGATs. The differences in intron size for the two loci are also apparently due to differences between the two genotypes (Jack vs. William 82). As for *GmDGAT1C*, the ninth intron is rather large and its length is 12,973 bp. The intron size may determine the gene size since most differences in the introns were insertions/deletions from 10 bp to a few hundred base pairs; however, the biological significance of the variation of the intron size or sequence remains to be determined. These results are consistent with the allotetraploid soybean genome origin with each *DGAT1* gene coming from each different soybean ancestor. With both genes apparently functional in soybean oil biosynthesis, the combined genes may at least partially account for the higher oil content of *G. max* compared with *Glycine* soybean ancestors (Chaven et al. 1982).

As is reported in peanuts and Arabidopsis, potential products of these GmDGAT3 genes are not predicted to localize to membranes unlike other DGATs (Rani et al. 2010; Saha et al. 2006). Arabidopsis has a predicted DGAT3 gene, *At1g48300*, with high homology to the reported peanut DGAT3, but its function is yet to be investigated. A recent report indicates that

another Arabidopsis gene, *At5g23940*, has TAG biosynthetic activity in vitro and apparently is involved in cutin formation in Arabidopsis leaves and stems (Rani et al. 2010; Saha et al. 2006) (see Supplemental Figure showing phylogenetic relations of soybean and other related DGATs). It is not clear whether this additional gene is involved in seed or other tissue TAG biosynthesis *in planta*.

Two DGAT1s can play a role in soybean oil biosynthesis

The high TAG biosynthetic activities with normal substrates in developing seeds suggest roles in common fatty acid incorporation into oil. The results with multiple substrates indicate that neither GmDGAT1A or B has much activity for vernolic acid accumulation into TAG and can explain the limited incorporation of vernolic acid into TAG when epoxygenase enzymes catalyzing vernolic acid biosynthesis are expressed in soybeans without DGATs that do have specificity for vernolic acid (Li et al. 2010a, b). The DGAT-specific activities in the present study are also two magnitudes higher than those from soybean developing seed microsomes reaching ca. 1 pmol/min/μg protein (or 10 pmol/min/nmol phosphatidylcholine) for the optimum substrate combination of oleoyl-CoA with *sn*-1,2-DODAG (Yu et al. 2006) which is probably due to a much higher proportion of GmDGAT1A and GmDGAT1B in the overall microsomal proteins when overexpressed in yeast.

As mentioned in the introduction, DGAT1, DGAT2 and PDATs all have been shown to be capable of TAG biosynthesis, but the precise roles of such enzymes and their corresponding genes in TAG accumulation in different organisms is only recently being unraveled. In yeast, three genes/enzymes can catalyze TAG biosynthesis, but a DGAT2 is responsible for most TAG biosynthesis in yeast (Oelkers et al. 2002). Arabidopsis has only one DGAT1, one DGAT2, and a number of PDAT-like sequences, and it has been known that DGAT1 is important in seed oil biosynthesis but other enzymes also contribute (e.g., Katavic et al. 1995). Zhang et al. (2009) recently showed that RNAi silencing of *PDAT1* in a *dgat1-1* background or *DGAT1* in *pdat1-1* background led to a 70 % to 80 % reduction in oil content but *DGAT2* and other *PDAT* mutants exhibited no such effects in Arabidopsis, indicating that DGAT1 and PDAT1 have overlapping function in Arabidopsis TAG biosynthesis. Zhang et al. (2009) suggest that the residual TAG seen in *DGAT1*, *PDGAT1* double RNAi lines of Arabidopsis may be due to incomplete suppression of the expression of these two genes by RNAi. DGAT2 is reported to play an important role in directing unusual fatty acids into TAG (Shockey et al. 2006; Li et al. 2010a, b). Co-expression of *DAGT2* with hydroxylase or epoxygenase in Arabidopsis (Burgal et al. 2008) and soybean (Li et al. 2010a, b) greatly increased accumulation of hydroxy and epoxy fatty acids in seeds. The roles of DGAT1 and DGAT2 in oil production appear to be species-dependent. In some plants, DGAT1 may play a more dominant role, whereas in plants containing unusual fatty acids, such as castor (*Ricinus communis* L.) and tung trees (*Vernicia fordii* (Hemsl.) Airy Shaw), DGAT2 appears to have an important function. A recent report suggests that in olives, DGAT1 contributes to most of the TAG deposited in seeds, but both DGAT1 and DGAT2 contribute to fruit TAG biosynthesis (Banilas et al. 2011).

The current knowledge of soybean oil biosynthesis is limited. Although several reports indicate a role of DGAT in oil accumulation in developing soybean seeds (Kwanyuen and Wilson 1986, 1990; Settlage et al. 1998), it is not yet clear what contributions if any DGAT1, DGAT2, and PDAT make in soy oil synthesis. Our previous data showed microsomal DGAT activity from soybean developing seeds was correlated with soy oil accumulation (Yu et al. 2006), and the data reported here show that two soybean DGAT1s have expression profiles consistent with their playing important roles in soybean oil biosynthesis. Further studies are needed to determine the specific roles of DGAT1s and other enzymes in soybean oil biosynthesis.

In conclusion, two functional *DGAT1* cDNAs were cloned and characterized from developing soybean seeds. This is the first characterization of all DGATs and PDAT1s in the soybean genome and cloning and characterization of a *GmDGAT1B* cDNA. Also, activity and substrate specificity of the two DGAT genes highly expressed during soybean seed development were assessed in a yeast expression system and in vitro assays. Their expression profiles support a function in soy oil accumulation in developing seeds with GmDGAT1B possibly being more important in this regard. Further studies are needed to verify roles of these DGATs and other possible enzymes in soybean seed oil biosynthesis.

Acknowledgments This work was supported by the United Soybean Board; the Kentucky Science and Engineering Foundation; Ashland, Inc.; and the Consortium for Plant Biotechnology Research as well as National Natural Science Foundation of China (No.30971806). The editorial input of Jessime Kirk and Richard Joost is much appreciated.

References

Banilas G, Karampelias M, Makariti I, Kourti A, Hatzopoulos P (2011) The olive DGAT2 gene is developmentally regulated and shares overlapping but distinct expression patterns with DGAT1. J Exp Bot 62:521–532. doi:10.1093/jxb/erq286

Bewley J, Black M (1994) Seeds: physiology of development and germination. Plenum Press, New York

Bouvier-Nave P, Benvenise P, Oelkers P, Sturley S, Schaller H (2000) Expression in yeast and tobacco of plant cDNAs encoding acyl CoA:diacylglycerol acyltransferase. Eur J Biochem 267:85–96

Burgal J, Shockey J, Lu CF, Dyer J, Larson T, Graham I, Browse J (2008) Metabolic engineering of hydroxy fatty acid production in

plants: RcDGAT2 drives dramatic increases in ricinoleate levels in seed oil. Plant Biotechnol J 6:819–831. doi:10.1111/j.1467-7652.2008.00361.x

Carmer SG, Swanson MR (1973) Evaluation of 10 pairwise multiple comparison procedures by Monte-Carlo methods. J Am Stat Assoc 68:66–74

Cases S, Smith SJ, Zheng Y, Myers HM, Lear SR, Sande E, Novak S, Collins C, Welch CB, Lusis AJ, Erickson SK, Farese RV Jr (1998) Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis. P Natl Acad Sci U S A 95:13018–13023

Chaven C, Hymowitz T, Newell C (1982) Chromosome number, oil and fatty acid content of species in the genus *Glycine* subgenus *Glycine*. J Am Oil Chem Soc 59:23–25. doi:10.1007/bf02670062

Dahlqvist A, Stahl 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 U S A 97:6487–6492

Durrett TP, McClosky DD, Tumaney AW, Elzinga DA, Ohlrogge J, Pollard M (2010) A distinct DGAT with sn-3 acetyltransferase activity that synthesizes unusual, reduced-viscosity oils in Euonymus and transgenic seeds. P Natl Acad Sci U S A 107:9464–9469. doi:10.1073/pnas.1001707107

Hatanaka T, Shimizu R, Hildebrand D (2004) Expression of a *Stokesia laevis* epoxygenase gene. Phytochemistry 65:2189–2196

He X, Chen GQ, Lin JT, McKeon TA (2004) Regulation of diacylglycerol acyltransferase in developing seeds of castor. Lipids 39:865–871

Hofmann K (2000) A superfamily of membrane-bound O-acyltransferases with implications for Wnt signaling. Trends Biochem Sci 25:111–112. doi:10.1016/s0968-0004(99)01539-x

Huang AHC (1992) Oil bodies and oleosins in seeds. Annu Rev Plant Physiol Plant Mol Biol 43:177–200

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

Kalscheuer R, Steinbuchel A (2003) A novel bifunctional wax ester synthase/acyl-CoA:diacylglycerol acyltransferase mediates wax ester and triacylglycerol biosynthesis in *Acinetobacter calcoaceticus* ADP1. J Biol Chem 278:8075–8082

Katavic V, Reed DW, Taylor DC, Giblin EM, Barton DL, Zou J, Mackenzie SL, Covello PS, Kunst L (1995) Alteration of seed fatty acid composition by an ethyl methanesulfonate-induced mutation in Arabidopsis thaliana affecting diacylglycerol acyltransferase activity. Plant Physiol 108:399–409

King A, Nam JW, Han JX, Hilliard J, Jaworski JG (2007) Cuticular wax biosynthesis in petunia petals: cloning and characterization of an alcohol-acyltransferase that synthesizes wax-esters. Planta 226:381–394. doi:10.1007/s00425-007-0489-z

Kroon JTM, Wei W, Simon WJ, Slabas AR (2006) Identification and functional expression of a type 2 acyl-CoA:diacylglycerol acyltransferase (DGAT2) in developing castor bean seeds which has high homology to the major triglyceride biosynthetic enzyme of fungi and animals. Phytochemistry 67:2541–2549

Kwanyuen P, Wilson RF (1986) Isolation and purification of diacylglycerol acyltransferase from germinating soybean cotyledons. Biochim Biophys Acta 877:238–245

Kwanyuen P, Wilson RF (1990) Subunit and amino acid composition of diacylglycerol acyltransferase from germinating soybean cotyledons. Biochim Biophys Acta 1039:67–72

Lardizabal K, Effertz R, Levering C, Mai J, Pedroso MC, Jury T, Aasen E, Gruys K, Bennett K (2008) Expression of *Umbelopsis ramanniana DGAT2A* in seed increases oil in soybean. Plant Physiol 148:89–96. doi:10.1104/pp.108.123042

Li F, Wu X, Lam P, Bird D, Zheng H, Samuels L, Jetter R, Kunst L (2008) Identification of the wax ester synthase/acyl-coenzyme a: diacylglycerol acyltransferase WSD1 required for stem wax ester biosynthesis in Arabidopsis. Plant Physiol 148:97–107. doi:10.1104/pp.108.123471

Li R, Yu K, Hatanaka T, Hildebrand DF (2010a) *Vernonia* DGATs increase accumulation of epoxy fatty acids in oil. Plant Biotechnol J 8:184–195

Li R, Yu K, Hildebrand D (2010b) DGAT1, DGAT2 and PDAT expression in seeds and other tissues of epoxy and hydroxy fatty acid accumulating plants. Lipids 45:145–157

Lu C, Xin Z, Ren Z, Miquel M, Browse J (2009) An enzyme regulating triacylglycerol composition is encoded by the ROD1 gene of *Arabidopsis*. Proc Natl Acad Sci 106:18837–18842. doi:10.1073/pnas.0908848106

Mason JM, Arndt KM (2004) Coiled coil domains: stability, specificity, and biological implications. ChemBioChem 5:170–176. doi:10.1002/cbic.200300781

McCartney AW, Dyer JM, Dhanoa PK, Kim PK, Andrews DW, McNew JA, Mullen RT (2004) Membrane-bound fatty acid desaturases are inserted co-translationally into the ER and contain different ER retrieval motifs at their carboxy termini. Plant J 37:156–173. doi:10.1111/j.1365-313X.2004.01949.x

McFie PJ, Stone SL, Banman SL, Stone SJ (2010) Topological orientation of acyl-coa:diacylglycerol acyltransferase-1 (DGAT1) and identification of a putative active site histidine and the role of the N terminus in dimer/tetramer formation. J Biol Chem 285:37377–37387. doi:10.1074/jbc.M110.163691

Oelkers P, Behar A, Cromley D, Billheimer J, Sturley S (1998) Characterization of two human genes encoding acyl coenzyme A: cholesterol acyltransferase-related enzymes. J Biol Chem 273:26765–26771

Oelkers P, Tinkelenberg A, Erdeniz N, Cromley D, Billheimer JT, Sturley SL (2000) A lecithin cholesterol acyltransferase-like gene mediates diacylglycerol esterification in yeast. J Biol Chem 275:15609–15612. doi:10.1074/jbc.C000144200

Oelkers P, Cromley D, Padamsee M, Billheimer JT, Sturley SL (2002) The DGA1 gene determines a second triglyceride synthetic pathway in yeast. J Biol Chem 277:8877–8881

Ohlrogge JB, Browse J (1995) Lipid biosynthesis. Plant Cell 7:957–970

Perry H, Harwood J (1993) Changes in the lipid content of developing seeds of *Brassica napus*. Phytochemistry 32:1411–1415

Rani SH, Krishna THA, Saha S, Negi AS, Rajasekharan R (2010) Defective in cuticular ridges (DCR) of *Arabidopsis thaliana*, a gene associated with surface cutin formation, encodes a soluble diacylglycerol acyltransferase. J Biol Chem 285:38337–38347. doi:10.1074/jbc.M110.133116

Routaboul J-M, Benning C, Bechtold N, Caboche M, Lepiniec L (1999) The TAG1 locus of Arabidopsis encodes for a diacylglycerol acyltransferase. Plant Physiol Biochem 37:831–840

Saha S, Enugutti 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. doi:10.1104/pp.106.082198

Schmutz J, Cannon SB, Schlueter J, Ma JX, Mitros T, Nelson W, Hyten DL, Song QJ, Thelen JJ, Cheng JL, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya MK, Sandhu D, Valliyodan B, Lindquist E, Peto M, Grant D, Shu SQ, Goodstein D, Barry K, Futrell-Griggs M, Abernathy B, Du JC, Tian ZX, Zhu LC, Gill N, Joshi T, Libault M, Sethuraman A, Zhang XC, Shinozaki K, Nguyen HT, Wing RA, Cregan P, Specht J, Grimwood J, Rokhsar D, Stacey G, Shoemaker RC, Jackson SA (2010) Genome sequence of the palaeopolyploid soybean. Nature 463:178–183. doi:10.1038/Nature08670

Settlage SB, Kwanyuen P, Wilson RF (1998) Relation between diacylglycerol acyltransferase activity and oil concentration in soybean. J Am Oil Chem Soc 75:775–781

Shockey JM, Gidda SK, Chapital DC, Kuan J-C, 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. doi:10.1105/tpc.106.043695

Stoveken T, Kalscheuer R, Malkus U, Reichelt R, Steinbuchel A (2005) The wax ester synthase/acyl coenzyme A:diacylglycerol acyltransferase from Acinetobacter sp strain ADP1: characterization of a novel type of acyltransferase. J Bacteriol 187:1369–1376. doi:10.1128/Jb.187.4.1369-1376.2005

Taylor DC, Yan Z, Kumar A, Francis T, Giblin EM, Barton DL, Ferrie JR, Laroche A, Shah S, Weiming Z, Snyder CL, Hall L, Rakow G, Harwood JL, Weselake RJ (2009) Molecular modification of triacylglycerol accumulation by over-expression of DGAT1 to produce canola with increased seed oil content under field conditions. Botany 87:533–543

USDA (2012a) In: Service FA (ed) Oilseeds: world markets and trade, FOP 10-09. USDA, Washington, DC

USDA (2012b) In: Service ER (ed) World agricultural supply and demand estimates, WASDE-477. USDA, Washington, DC

van Heeckeren WJ, Sellers JW, Struhl K (1992) Role of the conserved leucines in the leucine zipper dimerization motif of yeast GCN4. Nucleic Acids Res 20:3721–3724. doi:10.1093/nar/20.14.3721

Wang Z, Triezenberg SJ, Thomashow MF, Stockinger EJ (2005) Multiple hydrophobic motifs in Arabidopsis CBF1 COOH-terminus provide functional redundancy in trans-activation. Plant Mol Biol 58:543–559

Wang H-W, Zhang J-S, Gai J-Y, Chen S-Y (2006) Cloning and comparative analysis of the gene encoding diacylglycerol acyltransferase from wild type and cultivated soybean. Theor Appl Genet 112:1086–1097

Yu K, McCracken C Jr, Li R, Hildebrand DF (2006) Diacylglycerol acyltransferase from *Vernonia* and *Stokesia* prefer substrates with vernolic acid. Lipids 41:557–566

Yu K, Li R, Hatanaka T, Hildebrand D (2008) Cloning and functional analysis of two type 1 diacylglycerol acyltransferases from *Vernonia galamensis*. Phytochemistry 69:1119–1127

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. doi:10.1105/tpc.109.071795

Zheng P, Allen WB, Roesler K, Williams ME, Zhang S, Li J, Glassman K, Ranch J, Nubel D, Solawetz W, Bhattramakki D, Llaca V, Deschamps S, Zhong G-Y, Tarczynski MC, Shen B (2008) A phenylalanine in DGAT is a key determinant of oil content and composition in maize. Nat Genet 40:367–372

Zou J, Wei Y, Taylor DC (1999) The *Arabidopsis thaliana* TAG1 mutant has a mutation in a diacylglycerol acyl transferase gene. Plant J 19:645–654