



# Structure-Function Analysis of Diacylglycerol Acyltransferase Sequences from 70 Organisms

Cao



### **RESEARCH ARTICLE**



Open Access

# Structure-Function Analysis of Diacylglycerol Acyltransferase Sequences from 70 Organisms

Heping Cao

### Abstract

**Background:** Diacylglycerol acyltransferase families (DGATs) catalyze the final and rate-limiting step of triacylglycerol (TAG) biosynthesis in eukaryotic organisms. Understanding the roles of DGATs will help to create transgenic plants with value-added properties and provide clues for therapeutic intervention for obesity and related diseases. The objective of this analysis was to identify conserved sequence motifs and amino acid residues for better understanding of the structure-function relationship of these important enzymes.

**Results:** 117 DGAT sequences from 70 organisms including plants, animals, fungi and human are obtained from database search using tung tree DGATs. Phylogenetic analysis separates these proteins into DGAT1 and DGAT2 subfamilies. These DGATs are integral membrane proteins with more than 40% of the total amino acid residues being hydrophobic. They have similar properties and amino acid composition except that DGAT1s are approximately 20 kDa larger than DGAT2s. DGAT1s and DGAT2s have 41 and 16 completely conserved amino acid residues, respectively, although only two of them are shared by all DGATs. These residues are distributed in 7 and 6 sequence blocks for DGAT1s and DGAT2s, respectively, and located at the carboxyl termini, suggesting the location of the catalytic domains. These conserved sequence blocks do not contain the putative neutral lipid-binding domain, mitochondrial targeting signal, or ER retrieval motif. The importance of conserved residues has been demonstrated by site-directed and natural mutants.

**Conclusions:** This study has identified conserved sequence motifs and amino acid residues in all 117 DGATs and the two subfamilies. None of the completely conserved residues in DGAT1s and DGAT2s is present in recently reported isoforms in the multiple sequences alignment, raising an important question how proteins with completely different amino acid sequences could perform the same biochemical reaction. The sequence analysis should facilitate studying the structure-function relationship of DGATs with the ultimate goal to identify critical amino acid residues for engineering superb enzymes in metabolic engineering and selecting enzyme inhibitors in therapeutic application for obesity and related diseases.

### Background

The complete genomes of many organisms including human, mouse, *Arabidopsis* and rice have been sequenced. The immediate challenge of post-genomic biology is to determine the biological functions of proteins encoded by unknown genes. Many endogenous proteins occur in extremely low abundance (such as the anti-inflammatory protein tristetraprolin/zinc-finger protein 36, TTP/ZFP36) [1] and are labile (such as omega-3

Correspondence: Heping.Cao@ars.usda.gov

fatty-acid desaturase, FAD3) [2], which complicates characterization of those proteins.

One approach to gain clues about the structure-function relationship of proteins is to perform comprehensive amino acid sequence analysis. It is generally accepted that critical amino acid residues and sequence motifs in the same family of proteins are evolutionarily conserved. We previously used a protein sequence analysis approach to identify conserved sequence motifs and critical amino acid residues in several families of proteins from diverse organisms. The protein sequences we analyzed previously include the TTP/ZFP36 family involved in mRNA binding and destabilization [3,4], adenylate translocators [5], starch/glycogen synthases



© 2011 Cao et al; licensee BioMed Central Ltd. This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Commodity Utilization Research Unit, Southern Regional Research Center, Agricultural Research Service, U.S. Department of Agriculture, 1100 Robert E. Lee Blvd., New Orleans, Louisiana 70124, USA

[6], starch/glycogen branching enzymes [7,8], and starch/glycogen debranching enzymes [9].

Triacylglycerols (TAGs) are the major molecules of energy storage in eukaryotes. They also serve as a reservoir of fatty acids for membrane biogenesis and lead to obesity due to excessive accumulation in adipose tissues. Diacylglycerol acyltransferase families (DGATs) are integral microsomal membrane proteins that catalyze the last and rate-limiting step of TAG biosynthesis in eukaryotic organisms. DGATs esterify sn-1,2-diacylglycerol with a long-chain fatty acyl-CoA. DGAT genes have been isolated from many organisms. At least two forms of DGATs are present in mammals [10,11] and plants [12,13] with additional forms reported in burning bush (Euonymus alatus) [14], peanut [15] and Arabidopsis [16]. Plants and animals deficient in DGATs accumulate less TAG [17-19]. Animals with reduced DGAT activity are resistant to diet-induced obesity [18,20] and lack milk production [18]. Over-expression of DGAT enzymes increases TAG content in plants [14,21-26], animals [27-30] and yeast [31]. DGATs have nonredundant functions in TAG biosynthesis in species such as mice [19] and tung tree (Vernicia fordii) [13]. Mice deficient in DGAT1 are viable, have modest decreases in TAG, and are resistant to diet-induced obesity [18,32]. In contrast, mice deficient in DGAT2 have severe reduction of TAG and die shortly after birth [19]. The fact that DGAT1 is unable to compensate for the deficiency in DGAT2 knockout mice indicates the nonredundant functions of each DGAT isoform in TAG biosynthesis during mammal development. Therefore, understanding the roles of DGATs in plants and animals will have tremendous potential in creating new oilseed crops with value-added properties and providing information for therapeutic intervention for obesity and related diseases.

Limited numbers of DGAT amino acid sequences were analyzed previously [13,33]. However, there is a lack of comprehensive analysis of amino acid sequences of DGATs among diverse organisms. The objective of this analysis was to identify conserved sequence motifs and amino acid residues in 117 DGATs from 70 organisms to provide a better understanding of the structurefunction relationship of these important enzymes.

#### Results

#### Phylogenetic analysis and classification of DGATs

A database search using tung tree (*Vernicia fordii*) DGAT1 and DGAT2 protein sequences [13] has identified 117 DGAT sequences from 70 organisms including plants, animals, fungi and human (Table 1). More than two forms of DGATs are present in a number of species. For example, *Bos taurus* (cow) and *Brassica napus* (rape) have four forms of DGATs, whereas *Homo sapiens* (human) and *Danio rerio* (zebrafish) have three forms of DGATs (Table 1). Phylogenetic analysis indicates that all 117 DGAT protein sequences are grouped in the same phylogenetic tree (data not shown) and clearly separated into DGAT1 and DGAT2 subfamilies (Figure 1). DGAT1s are more conserved than DGAT2s. DGATs from plants, animals and fungi are also distinctly separated from each other with a few exceptions (Figure 1).

Six DGATs previously classified as DGAT2s in the GenBank databases [GenBank:EGC41804.1, GenBank: EEQ31683.1, GenBank:AAY40785.1, GenBank: AAD40881.1, GenBank:EFY86774.1 and GenBank: EFY97444.1] are clustered within the DGAT1 subfamily in this phylogenetic analysis (Figure 1). More detailed sequence analysis within the multiple sequence alignment also indicates that their sequences are more similar to DGAT1s (see below). Therefore, these six DGAT2s are reclassified as DGAT1s (AcDGAT1-EGC41804.1, AoDGAT1-EEQ31683.1, BjDGAT1b-AAY40785.1, BnDGAT1b-AAD40881.1, MaDGAT1a-EFY86774.1, and MaDGAT1b-EFY97444.1, respectively) (Table 1).

#### DGAT properties and amino acid composition

It is generally observed that DGAT1s are much bigger proteins with more transmembrane domains than DGAT2s. The properties and amino acid composition of 109 full-length DGATs were analyzed and calculated according to the classifications of DGAT1 and DGAT2 subfamilies by phylogenetic analysis. The eight partial sequences listed in Table 1 were excluded from this analysis (AaDGAT1-XP\_001658299, BnDGAT1b-AAD40881.1, BtDGAT2c-XP\_002683800.1, ChDGAT1-ABD59375.1, CvDGAT2-EFN51306.1, OaDGAT2-XP\_001518899.1, PtDGAT1-XP\_002177753.1 and PtDGAT2-XP\_002317635.1).

DGAT1s have an average of 515 amino acid residues with a standard deviation of 44 amino acid residues among the 55 sequences (Table 2). The average residue number of DGAT1 is 171 amino acids greater than the average DGAT2 length, which has an average of 344 residues with a standard deviation of 29 residues among the 54 sequences (Table 2). This corresponds to approximately 20 kDa difference in the molecular mass. The isoelectric points of DGAT1s and DGAT2s are similar with an average value of 9.17 and 9.28, respectively, however DGAT1s have approximately 4 more charges at pH 7 than DGAT2s (Table 2).

The frequency of functional amino acid residue groups between DGAT1 and DGAT2 subfamilies is remarkably similar (Table 2). DGAT1s and DGAT2s have charged residues (RKHYCDE) of approximately 26%, which includes approximately 7% of acidic residues (DE) and 10% of basic residues (KR). The frequency of polar

Table 1 DGATs sequence information

Ne	DGAT name	Organism	DCAT	ConPont #	No		Organism	DCAT	ConPont #
.00			JUGAI		110.			JUGAI	
1	AaDGAT1- XP_001658299 *	Aedes aegypti (A) ***	1	XP_001658299	60	ntDGA11- AAF19345.1	<i>Nicotiana tabacum</i> (P, tobacco)	1	AAF19345.1
2	AcDGAT1- EGC41804.1 **	Ajellomyces capsulatus (F)	1	EGC41804.1	61	NvDGAT2a- XP_001630435.1	Nematostella vectensis (A, worm)	2a	XP_001630435.1
3	AcDGAT2- XP_003225477.1	Anolis carolinensis (A)	2	XP_003225477.1	62	NvDGAT2b- XP_001633322.1	Nematostella vectensis (A, worm)	2b	XP_001633322.1
4	AgDGAT2- NP_983542.1	Ashbya gossypii (F)	2	NP_983542.1	63	NvDGAT2c- XP_001635548.1	Nematostella vectensis (A, worm)	2c	XP_001635548.1
5	AoDGAT1- EEQ31683.1 **	Arthroderma otae (F)	1	EEQ31683.1	64	OaDGAT1- NP_001103634.1	Ovis aries (A, sheep)	1	NP_001103634.1
6	AtDGAT1- NP_179535.1	Arabidopsis thaliana (P)	1	NP_179535.1	65	OaDGAT2- XP_001518899.1 *	Ovis aries (A, sheep)	2	XP_001518899.1
7	AtDGAT2- NP_566952	Arabidopsis thaliana (P)	2	NP_566952	66	OcDGAT1- XP_002724427.1	<i>Oryctolagus cuniculus</i> (A, rabbit)	1	XP_002724427.1
8	BbDGAT1- AAZ22403.1	<i>Bubalus bubalis</i> (A, buffalo)	1	AAZ22403.1	67	OeDGAT1- AAS01606.1	<i>Olea europaea</i> (P, tree)	1	AAS01606.1
9	BjDGAT1a- AAY40784.1	Brassica juncea (P)	1a	AAY40784.1	68	OeDGAT2- ADG22608.1	<i>Olea europaea</i> (P, tree)	2	ADG22608.1
10	BjDGAT1b- AAY40785.1 **	Brassica juncea (P)	1b	AAY40785.1	69	OsDGAT1- NP_001054869.2	<i>Oryza sativa</i> (P, rice)	1	NP_001054869.2
11	BnDGAT1a- AAD45536.1	Brassica napus (P)	1a	AAD45536.1	70	OsDGAT2a- NP_001047917	<i>Oryza sativa</i> (P, rice)	2a	NP_001047917
12	BnDGAT1b- AAD40881.1 *, **	Brassica napus (P)	1b	AAD40881.1	71	OsDGAT2b- NP_001057530	<i>Oryza sativa</i> (P, rice)	2b	NP_001057530
13	BnDGAT2a- ACO90187	Brassica napus (P)	2	ACO90187	72	OtDGAT2- XP_003083539.1	<i>Ostreococcus tauri</i> (Algae)	2	XP_003083539.1
14	BnDGAT2b- ACO90188	Brassica napus (P)	2	ACO90188	73	PaDGAT2- XP_002822304.1	Pongo abelii (A, Sumatran orangutan)	2	XP_002822304.1
15	BtDGAT1- NP_777118.2	Bos taurus (A, cattle)	1	NP_777118.2	74	PbDGAT1- EEH17170.1	Paracoccidioides brasiliensis (F)	1	EEH17170.1
16	BtDGAT2a- DAA21853.1	Bos taurus (A, cattle)	2a	DAA21853.1	75	PfDGAT1- AAG23696.1	Perilla frutescens (P)	1	AAG23696.1
17	BtDGAT2b- XP_875499.3	<i>Bos taurus</i> (A, cattle)	2b	XP_875499.3	76	PpDGAT1- EFA85004.1	Polysphondylium pallidum (F, mold)	1	EFA85004.1
18	BtDGAT2c- XP_002683800.1 *	<i>Bos taurus</i> (A, cattle)	2c	XP_002683800.1	77	PpDGAT2- EFA83646.1	Polysphondylium pallidum (F, mold)	2	EFA83646.1
19	CeDGAT2a- NP_505413.1	<i>Caenorhabditis elegans</i> (A, worm)	2a	NP_505413.1	78	PpDGAT1- XP_001770929.1	Physcomitrella patens (P, moss)	1	XP_001770929.1
20	CeDGAT2b- NP_872180.1	<i>Caenorhabditis elegans</i> (A, worm)	2b	NP_872180.1	79	PpDGAT2a- XP_001758758.1	Physcomitrella patens (P, moss)	1	XP_001758758.1
21	CfDGAT1b- XP_849176.1	Canis familiaris (A, dog)	1b	XP_849176.1	80	PpDGAT2b- XP_001777726.1	Physcomitrella patens (P, moss)	2b	XP_001777726.1
22	CfDGAT1c- XP_858062.1	Canis familiaris (A, dog)	1c	XP_858062.1	81	PsDGAT2- ABK26256.1	Picea sitchensis (P, tree)	2	ABK26256.1
23	ChDGAT1- ABD59375.1 *	<i>Capra hircus</i> (A, sheep)	1	ABD59375.1	82	PtDGAT1- XP_520014.2	<i>Pan troglodytes</i> (A, chimpanzee)	1	XP_520014.2
24	CiDGAT2- XP_002120879.1	Ciona intestinalis (A)	2	XP_002120879.1	83	PtDGAT2- XP_527842.2	<i>Pan troglodytes</i> (A, chimpanzee)	2	XP_527842.2
25	CrDGAT2a- XP_001694904.1	Chlamydomonas reinhardtii (Algae)	2a	XP_001694904.1	84	PtDGAT1- XP_002177753.1 *	Phaeodactylum tricornutum (F)	1	XP_002177753.1
26	CrDGAT2b- XP_001693189.1	Chlamydomonas reinhardtii (Algae)	2b	XP_001693189.1	85	PtDGAT1a- XP_002308278.1	<i>Populus trichocarpa</i> (P, tree)	1a	XP_002308278.1
27	CvDGAT1- EFN50697.1	Chlorella variabilis (Algae)	1	EFN50697.1	86	PtDGAT1b- XP_002330510.1	<i>Populus trichocarpa</i> (P, tree)	1b	XP_002330510.1
28	CvDGAT2- EFN51306.1 *	Chlorella variabilis (Algae)	2	EFN51306.1	87	PtDGAT2- XP_002317635.1 *	<i>Populus trichocarpa</i> (P, tree)	2	XP_002317635.1

### Table 1 DGATs sequence information (Continued)

29	DdDGAT1- XP_645633.2	Dictyostelium discoideum (mold)	1	XP_645633.2	88	RcDGAT1- XP_002514132.1	<i>Ricinus communis</i> (P, castor bean)	1	XP_002514132.1
30	DdDGAT2- XP_635762.1	Dictyostelium discoideum (mold)	2	XP_635762.1	89	RcDGAT2- XP_002528531.1	<i>Ricinus communis</i> (P, castor bean)	1	XP_002528531.1
31	DmDGAT1a- NP_609813.1	Drosophila melanogaster (A, fly)	1a	NP_609813.1	90	RnDGAT1- NP_445889.1	<i>Rattus norvegicus</i> (A, rat)	1	NP_445889.1
32	DmDGAT1d- NP_995724.1	Drosophila melanogaster (A, fly)	1d	NP_995724.1	91	RnDGAT2- NP_001012345.1	<i>Rattus norvegicus</i> (A, rat)	2	NP_001012345.1
33	DrDGAT1a- NP_956024.1	<i>Danio rerio</i> (A, zebrafish)	1a	NP_956024.1	92	SbDGAT1a- XP_002437165.1	Sorghum bicolor (P, sorghum)	1a	XP_002437165.1
34	DrDGAT1b- NP_001002458.1	<i>Danio rerio</i> (A, zebrafish)	1b	NP_001002458.1	93	SbDGAT1b- XP_002439419.1	Sorghum bicolor (P, sorghum)	1b	XP_002439419.1
35	DrDGAT2- NP_001025367.1	<i>Danio rerio</i> (A, zebrafish)	2	NP_001025367.1	94	SbDGAT2- XP_002452652.1	Sorghum bicolor (P, sorghum)	2	XP_002452652.1
36	EaDGAT1- AAV31083.1	Euonymus alatus (P)	1	AAV31083.1	95	ScDGAT2- NP_014888.1	Saccharomyces cerevisiae (F, yeast)	2	NP_014888.1
37	EaDGAT2- ADF57328.1	Euonymus alatus (P)	2	ADF57328.1	96	SkDGAT1- XP_002736160.1	Saccoglossus kowalevskii (A, worm)	1	XP_002736160.1
38	EoDGAT2- ACO35365.1	Elaeis oleifera (P)	2	ACO35365.1	97	SmDGAT1- XP_002964165.1	Selaginella moellendorffii (P)	1	XP_002964165.1
39	EpDGAT1- ACO55635.1	Echium pitardii (P)	1	ACO55635.1	98	SmDGAT2- XP_002972054.1	Selaginella moellendorffii (P)	2	XP_002972054.1
40	GmDGAT1a- AAS78662.1	<i>Glycine max</i> (P, soybean)	1a	AAS78662.1	99	SpDGAT2- AAQ89590.1	Spirodela polyrhiza (P)	2	AAQ89590.1
41	GmDGAT1b- BAE93461.1	<i>Glycine max</i> (P, soybean)	1b	BAE93461.1	100	SpDGAT2- XP_001713160.1	Schizosaccharomyces pombe (F, yeast)	2	XP_001713160.1
42	GmDGAT2- ACU20344.1	<i>Glycine max</i> (P, soybean)	2	ACU20344.1	101	SsDGAT1- NP_999216.1	Sus scrofa (A, pig)	1	NP_999216.1
43	HaDGAT2- ABU50328.1	Helianthus annuus (P)	2	ABU50328.1	102	TcDGAT1- XP_975142.1	Tribolium castaneum (A)	1	XP_975142.1
44	HsDGAT1- NP_036211.2	<i>Homo sapiens</i> (human)	1	NP_036211.2	103	TcDGAT2- XP_975146.1	Tribolium castaneum (A)	2	XP_975146.1
45	HsDGAT2a- AAQ88896.1	<i>Homo sapiens</i> (human)	2a	AAQ88896.1	104	TgDGAT1- AAP94209.1	Toxoplasma gondii (A, worm)	1	AAP94209.1
46	HsDGAT2b- NP_835470.1	<i>Homo sapiens</i> (human)	2b	NP_835470.1	105	TgDGAT2- XP_002187643.1	<i>Taeniopygia guttata</i> (A, bird)	2	XP_002187643.1
47	HvDGAT2- BAJ85730.1	<i>Hordeum vulgare</i> (P, barley)	2	BAJ85730.1	106	TmDGAT1- AAM03340.2	Tropaeolum majus (P)	1	AAM03340.2
48	IpDGAT2b- NP_001188005.1	<i>lctalurus punctatus</i> (A, catfish)	2b	NP_001188005.1	107	UrDGAT2A- AAK84179.1	Umbelopsis ramanniana (F)	2a	AAK84179.1
49	JcDGAT1- ABB84383.1	Jatropha curcas (P)	1	ABB84383.1	108	UrDGAT2B- AAK84180.1	Umbelopsis ramanniana (F)	2b	AAK84180.1
50	LjDGAT1- AAW51456.1	Lotus japonicas (P)	1	AAW51456.1	109	VfDGAT1- DQ356680.1	<i>Vernicia fordii</i> (P, tung tree)	1	DQ356680.1
51	MaDGAT1a- EFY86774.1 **	Metarhizium acridum (F)	1a	EFY86774.1	110	VfDGAT2- DQ356682.1	<i>Vernicia fordii</i> (P, tung tree)	2	DQ356682
52	MaDGAT1b- EFY97444.1 **	Metarhizium anisopliae (F)	1b	EFY97444.1	111	VgDGAT1- ABV21945.1	Vernonia galamensis (P)	1	ABV21945.1
53	MdDGAT1- XP_001371565.1	Monodelphis domestica (A, opossum)	1	XP_001371565.1	112	VgDGAT2- ACV40232.1	Vernonia galamensis (P)	2	ACV40232.1
54	MdDGAT2- XP_001365685.1	Monodelphis domestica (A, opossum)	2	XP_001365685.1	113	VvDGAT1- XP_002279345.1	<i>Vitis vinifera</i> (P, grape)	1	XP_002279345.1
55	MmDGAT1- NP_034176.1	<i>Mus musculus</i> (A, mouse)	1	NP_034176.1	114	VvDGAT2- XP_002263626	<i>Vitis vinifera</i> (P, grape)	2	XP_002263626
56	MmDGAT2- NP_080660.1	<i>Mus musculus</i> (A, mouse)	2	NP_080660.1	115	XtDGAT2- NP_989372.1	<i>Xenopus tropicalis</i> (A, frog)	2	NP_989372.1
57	MmDGAT1- XP_001090134.1	<i>Macaca mulatta</i> (A, monkey)	1	XP_001090134.1	116	ZmDGAT1b- EU039830	Zea mays (P, corn)	1b	EU039830

Table T DGATS sequence information (Continue	Table 1 DG/	's sequence	information	(Continued
--	-------------	-------------	-------------	------------

58	MtDGAT1- ABN09107.1	<i>Medicago truncatula</i> (P)	1	ABN09107.1	117 ZmDGAT2- NP_001150174.1	Zea mays (P, corn)	2	NP_001150174.1
59	MtDGAT2- ACJ84867.1	<i>Medicago truncatula</i> (P)	2	ACJ84867.1				

\*8 partial sequences were not used for statistics. \*\*6 DGATs named as DGAT2s in the GenBank Databases should be DGAT1s as indicated in the table because they are more similar to DGAT1s. \*\*\*A: animal, F: fungus, P: plant.

resides (NCQSTY) is also similar with an average of 25% vs. 22% in polar residues for DGAT1s and DGAT2s, respectively. DGAT1s and DGAT2s are integral membrane proteins with 42% and 41% of the total residues being hydrophobic (AILFWV), respectively (Table 2).

### Identification of amino acid residues and sequence motifs conserved among all DGATs

The sequences between DGAT1s and DGAT2s are very divergent. No common features between these two subfamilies were reported previously using 7 DGAT1s and 8 DGAT2s from 10 organisms [13]. However, the biochemical function of the two DGAT isoforms is essentially identical in enzymatic assays, suggesting that certain common sequence conservations are probably present between them. Multiple sequence alignment was performed to analyze all 117 DGATs including 59 DGAT1s from 48 organisms and 58 DGAT2s from 44 organisms. Sequence alignment has identified only two completely conserved amino acid residues among all DGATs. One perfectly conserved proline residue corresponds to P248 in AaDGAT1-XP\_001658299 and P151 in VfDGAT2-DQ356682.1 (Figure 2A); the other residue, a perfectly conserved phenylalanine residue corresponds to F344 in AaDGAT1-XP\_001658299 and F225 in VfDGAT2-DQ356682.1 (Figure 2B). The conserved phenylalanine residue is followed by a glycine residue conserved in all except DGAT1s of Aedes aegypti, Drosophila melanogaster and Tribolium castaneum (Figure 2B). The other proline residue in Figure 2C is almost completely conserved except DGAT2 of Ostreococcus tauri, which corresponds to P411 in AaDGAT1-XP\_001658299 and P276 in VfDGAT2-DQ356682.1. Based on the sequence conservation patterns with these conserved residues as anchors, the conserved sequence motifs among all DGATs are named as Motif 1 (P Block), Motif 2 (FG Block) and Motif 3 (P-1 Block). The highly conserved residues among diverse organisms may be located at the active sites of the enzymes and play important roles in structure, substrate binding and/or catalysis.

### Identification of amino acid residues and sequence motifs conserved in DGAT1s

Multiple sequence alignment was performed to identify conserved amino acid residues and sequence motifs within the DGAT1 subfamily. Among the 55 full-length DGAT1s, 41 amino acid residues are completely conserved, which correspond to 8.0% of the total 515 residues of DGAT1s. Table 3 shows the positions of the 41 completely conserved residues in DGAT1s from representatives of animal group (mouse), plant group (tung tree) and fungus group (Dictyostelium discoideum). These completely conserved residues are located in seven sequence motifs of DGAT1s. Based on the sequence conservation patterns with the completely conserved residues as anchors, the conserved sequence motifs of DGAT1s are named as Motif 1 (GL Block), Motif 2 (KSR Block), Motif 3 (PTR Block) (Figure 3A-C), Motif 4 (QP Block), Motif 5 (LWLFFEFDR-FYWWNWWNPPFSHP Block) (Figure 4A-B), Motif 6 (FQL Block) and Motif 7 (NGQPY Block) (Figure 5A-B). The great majority of the conserved residues are located within the carboxyl terminal halves of DGAT1s. Among the completely conserved residues, 33 of them are located within the last 200 residues from the carboxyl termini in Motifs 4-7 (Figures 4, 5), and 23 of them are concentrated in the most conserved region with approximately 100 residues in Motif 6 (Figure 4B). The first two conserved residues (G, L) in Motif 1 start approximately 100 residues from the amino termini (Figure 3A). The next three conserved residues (K, S, R) in Motif 2 start until approximately 200 residues from the amino termini (Figure 3B). The last conserved residue (Y) in Motif 7 ends within the last 20 residues from the carboxyl termini of DGAT1s (Figure 5B).

### Identification of amino acid residues and sequence motifs conserved in DGAT2s

Multiple sequence alignment was also performed using all DGAT2s to identify conserved amino acid residues and sequence motifs within this subfamily. Sixteen residues are completely conserved in the 54 full-length DGAT2s, corresponding to 4.7% of the total 344 residues. Table 3 shows the positions of the 16 completely conserved residues in DGAT2s from representatives of animal group (mouse), plant group (tung tree) and fungus group (*Dictyostelium discoideum*). Based on the sequence conservation patterns with the completely conserved residues as anchors, the conserved sequence motifs of DGAT2s are named as Motif 1 (PH Block), Motif 2 (PR Block), Motif 3 (GGE Block) (Figure 6A-C), Motif 4 (RGFA Block), Motif 5 (VPFG Block) and Motif 6 (G Block) (Figure 7A-C).

**Figure 1 Phylogenetic analysis of DGAT1s and DGAT2s**. The presumed evolutionary relationships among the 117 DGATs from 70 organisms (listed in Table 1) were analyzed by phylogenetic analysis based on the Neighbor-Joining method of Saitou and Nei. The numbers in the parenthesis following DGAT names are the calculated distance values which reflect the degree of divergence between all pairs of DGAT sequences analyzed. The sequences above the red line are from animals, whereas the sequences below the green line are from plants. A red "star" before the sequence indicates the exceptional sequence from the grouping.

Cao *BMC Research Notes* 2011, **4**:249 http://www.biomedcentral.com/1756-0500/4/249



Table 2 DGATs properties and amino acid composition(% by frequency)

DGAT1 (n = 55)	DGAT2 (n = 54)
515 ± 44	344 ± 29
58796 ± 4871	38920 ± 3330
9.17 ± 0.46	9.28 ± 0.42
15.11 ± 5.98	11.34 ± 4.25
26.03 ± 1.27	26.35 ± 2.02
7.20 ± 0.90	7.38 ± 0.87
9.97 ± 0.78	10.47 ± 1.15
24.59 ± 2.49	22.25 ± 2.93
42.18 ± 2.25	$40.92 \pm 2.70$
	DGAT1 (n = 55) 515 ± 44 58796 ± 4871 9.17 ± 0.46 15.11 ± 5.98 26.03 ± 1.27 7.20 ± 0.90 9.97 ± 0.78 24.59 ± 2.49 42.18 ± 2.25

Similar to DGAT1s, these conserved residues are located at the carboxyl termini of DGAT2s. Eight of them are concentrated within 25 residues in the highly conserved Motifs 4 and 5 of DGAT2s (Figure 7A-B). The first two conserved residues (PH) in Motif 1 start until approximately 100 residues from the amino termini but the last residue (G) ends within the last 50 residues from the carboxyl termini of DGAT1s (Figure 6A and 7C).

## Sequence analysis of important motifs in less conservative regions of DGATs

Several studies have reported functional motifs in DGATs. However, the conserved sequence motifs identified by our extensive sequence analysis as discussed above do not contain any of the reported putative neutral lipid-binding domain [34], mitochondrial targeting signal [35] or ER retrieval motif [13]. It was reported that mouse DGAT2 contains a consensus sequence (FLXLXXX<sup>n</sup>) for a putative neutral lipid-binding domain [34] which was shown to be present in proteins that either bind to or metabolize neutral lipids [36]. However, this putative motif is only modestly conserved in animal DGAT2s and not present in any plant DGAT2 (Figure 8A). Mouse DGAT2 was also reported to contain a putative mitochondrial targeting signal with positively charged residues (RXKXXK) targeting proteins to mitochondria [35]. This motif is only found in a few animal DGATs but not conserved in any of the plant or fungi DGAT2s (Figure 8B). DGATs are ER-localized enzymes with an ER retrieval motif (LKLEI) at the extreme carboxyl terminus of tung DGAT2 [13]. This sequence analysis shows that this pentapeptide ER-retrieval motif is only modestly conserved in plant DGAT2s but not in animal or fungal DGAT2s, although animal and fugal DGATs are also located to ER (Figure 8C).

# Sequence analysis of important amino acid residues in less conservative regions of DGATs

The importance of some less conserved residues of DGATs has been demonstrated by site-directed mutants

(Table 4). Mutagenesis of a putative SnRK1 target site S197 in *Tropaeolum majus* DGAT1 results in a 38%-80% increase in DGAT1 activity, and over-expression of the mutated TmDGAT1 in *Arabidopsis* results in a 20%-50% increase in oil content on a per seed basis (Table 4) [25]. Figure 9A shows that this serine residue is conserved in most of the plants and some animals except that the same position in DGAT1s alignment is replaced with proline, glycine, threonine and lysine residue in DGAT1s from other organisms. A similar serine residue is not found in any of DGAT2s. Mutagenesis at P216 in *Tropaeolum majus* DGAT1 eliminates almost all of the activity (Table 4) [25]. The P216 residue is completely conserved in plant DGAT1s but is missing in mammalian DGAT1s (Figure 9B).

A highly conserved region with a consensus sequence of "YFP" in DGAT2s is essential for enzymatic activity of DGAT2 from Saccharomyces cerevisiae (Table 4) [33]. These three residues are highly conserved and located before Motif 1 but none of the three amino acid residues is completely conserved among all DGAT2s in our analysis using 54 full-length DGAT2s (Figure 10A). The consensus "YFP" is replaced with "YYP" in DGAT2s of human, chimpanzee (Pan troglodytes) and Ashbya gossypii, "FFP" in Helianthus annuus and Nematostella vectensis, and "HFP" in caster bean (Ricinus communis), Vernonia galamensisand, and Selaginella moellendorffii. It was reported that a unique region is present in DGAT2 of Saccharomyces cerevisiae [33], but in our expanded analysis, a similar region is also found in Ashbya gossypii (data not shown).

Mutations at F80/L81/L83 in mouse DGAT2 [34] and F71/L73 in baker's yeast DGAT2 [33] result in partial loss of the activity (Table 4). This region is only relatively conserved in some animal DGAT2s (Figure 10B). Finally, ScDGAT2 has a unique cysteine residue (C314) which is not involved in catalysis but may be located near the active site or related to proper folding of the protein [37]. However, this residue is only found in DGAT2s from baker's yeast and the other fungi *Ashbya gossypii* and *Physcomitrella patens*, but is not present in the same position of the alignment in any of the other 51 DGAT2s or any of the 55 DGAT1s analyzed (Figure 10C).

# Sequence analysis of important amino acid residues of DGATs shown in natural mutants

The importance of some relatively conserved residues in TAG biosynthesis has been demonstrated by two well-known natural mutants in corn and cattle. A phenylalanine insertion (F469) in DGAT1-2 increases oil and oleic-acid contents in maize. Ectopic expression of the high-oil DGAT1-2 allele increases oil and oleic-acid contents by up to 41% and 107%, respectively (Table 4)

algorithm and 117 DGAT protein sequences from 70 organisms (listed in Table 1). DGAT sequence name is on the left of alignment followed by the start of the amino acid residue of each DGAT protein sequence. The completely conserved proline and phenylalanine residues are highlighted in red on yellow. Other color code and related information are described in "Methods" section. (A) Motif 1 (P Block): the completely conserved proline residue in all DGATs, (B) Motif 2 (FG Block): the completely conserved tryptophan residue followed by a highly conserved glycine residue in all DGATs, (C) Motif 3 (P-1 Block): the almost completely conserved proline residue except one in all DGATs.

Figure 2 Conservation of proline and phenylalanine residues in all DGATs. Multiple sequence alignment was performed using the ClustalW

Cao BMC Research Notes 2011, 4:249 http://www.biomedcentral.com/1756-0500/4/249



Y512

Organism	DGAT1	DGAT2		
Organism Mouse (animal)	DGAT1  1 MGDRGGAGSS RRRRTGSRVS VQGGSGPKVE EDEVRDAAVS PDLGAGGDAP 51 APAPAPAHTR DKDGRTSVGD GYWDLRCHRL QDSLFSSDSG FSNYRGILNW 101 CVVMLILSNA RLFLENLIKY GILVDPIQVV SLFLKDPYSW PAPCVIIASN 151 IFVVAAFQIE KRLAVGALTE QMGLLLHVVN LATIICFPAA VALLVESITP 201 VGSVFALASY SIMFLKLYSY RDVNLWCRQR RVKAKAVSTG KKVSGAAAQQ 251 AVSYPDNLTY RDLYYFIFAP TLCYELNFPR SPRIRKRFLL RRVLEMLFFT 301 QLQVGLIQQW MVPTIQNSMK PFKDMDYSRI IERLLKLAVP NHLINLIFFY 351 WFHSCLNAV AELLQFGDRE FYRDWWNAES VTYFWQNWNI PVHKWCIRHF 401 YKPMLRHGSS KWVARTGVFL TSAFFHEYLV SVPLRMFRLW AFTAMMAQVP 451 LAWIVGRFFQ GNYGNAAVWV TLIIGQPVAV LMYVHDYYVL NYDAPVGV 1) G96, 2) L105, 3) K216, 4) S219, 5) R228, 6) P270, 7) T271, 8) R285, 9) Q308, 10) P313, 11) L337, 12) W345, 13) L346, 14) F348, 15) F353, 16) E362, 17) F366, 18) D368, 19) R369, 20) F371, 21) Y372, 22) W375, 23) W376, 24) N377, 25) W385, 26 W388, 27) N389, 28) P391, 29) P403, 30) F419, 31) S422, 32) H426, 33) P433, 34) F442, 35) Q448, 36) L451, 37) N465, 38) G475, 39) Q476, 40) P477, 41) Y483 1 MTLETPDNS TDATTSGGAE SSSDLNLSLR PDPDEVENT	<pre>1 MKTLIAAYSG VLRGERRAEA ARSENKNKGS ALSREGSGRW GTGSSILSAL 51 QDIFSVTWLN RSKVEKQLQV ISVLQWVLSF LVLGVACSVI LMYTFCTDCW 101 LIAVLYFTWL AFDWNTPKKG GRRSQWVRNW AVWRYFRDYF PIQLVKTHNL 151 LTTRNYIFGY H<u>PHGIMGLGA FCNFSTEATE VSKKFPGIRP YLATLAGNFR</u> 201 MPVLREYLMS GGICPVNRDT IDYLLSKNGS GNAIIIVVGG AAESLSSMPG 251 KNAVTLKNRK GFVKLALRHG ADL<u>VP</u>TYSFG ENEVYKQVIF EEGSWGRWVQ 301 KKFQKYIGFA PCIFHGRGLF SSDTWGLVPY SKPITTVVGE PITVPKLEHP 351 TQKDIDLYHA MYMEALVKLF DNHKTKFGLP ETEVLEVN 1) P162, 2) H163, 3) <u>P202</u>, 4) R205, 5) G239, 6) G240, 7) E243, 8) R259, 9) G261, 10) F262, 11) A266, 12) V274, 13) P275, 14) <u>F279</u>, 15) G280, 16) E339</pre>		
(piant)	<pre>RRTASNSDG AVAELASKID 51 ELESDAGGGQ VIKDPGAEMD SGTLKSNGKD CGTVKDRIEN RENRGGSDVK 101 FTYRPSVPAH RALKESPLSS DNIFKQSHAG LFNLCIVVLV AVNSRLIIEN 151 IMKYGWLIKT GFWFSSRSLR DWPLLMCCLT LPIFSLAAYL VEKLACRKYI 201 SAPTVVFLHI LFSSTAVLYP VSVILSCESA VLSGVALMLF ACIVWLKLVS 251 YAHTNFDMRA IANSVDKGDA LSNASSAESS HDVSFKSLVY FMVAPTLCYQ 301 PSYPRTASIR KGWVVRQFVK LIIFTGFMGF IIEOYINPIV QNSQHPLKGD 351 LLYAIERVLK LSVPNLYVWL CMFYCFFHLW LNILAELLFF GDREFYKDWW 401 NARTVEEYWR MWNMPVHKWM VRHIYFPCLR HKIPRGVALL ITFFVSAVFH 451 ELCIAVPCHI FKLWAFIGIM FQIPLVGITN YLQNKFRSSM VGNMIFWFIF 501 CILGOPMCLL LYYHDLMNRK GTTESR 1) G130, 2) L139, 3) K247, 4) S250, 5) R259, 6) <u>P295</u>, 7) T296, 8) R310, 9) Q334, 10) P338, 11) L361, 12) W369, 13) L370, 14) F373, 15) F377, 16) E386, 17) F390, 18) D392, 19) R393, 20) F395, 21) Y396, 22) W399, 23) W400, 24) N401, 25) W409, 26) W412, 27) N413, 28) P415, 29) P427, 30) F443, 31) S446, 32) H450, 33) P457, 34) F466, 35) Q472, 36) L475, 37) N493, 38) G504, 39) Q505, 40) P506, 41)</pre>	ALAIWLGSFH FILFLVSSS1 51 FLPFSKFLLV IGLLLFFMVI PINDRSKLGQ CLFSYISRHV CSYFPITLHV 101 EDINAFRSDR AYVFGYEPHS VFPIGVMILS LGLIPLPNIK FLASSAVFYT 151 PFLRHIWSWC GLTPATRKNF VSLLSSGYSC ILVPGGVQET FYMKQDSEIA 201 FLKARRGFIR IAMQTGTPLV PVFCFGQMHT FKWWKPDGEL FMKIARAIKF 251 TPTIFWGVLG TPLPFKNPMH VVVGRPIEVK QNPQPTAEEV AEVQREFIAS 301 LKNLFERHKA RVGYSDLKLE IF 1) P118, 2) H119, 3) P151, 4) R154, 5) G185, 6) G186, 7) E189, 8) R205, 9) G207, 10) F208, 11) A212, 12) V220, 13) P221, 14) F225, 15) G226, 16) E274		

### Table 3 The completely conserved residues in DGATs using examples of mouse, tung tree and *Dictyostelium* discoideum

#### Table 3 The completely conserved residues in DGATs using examples of mouse, tung tree and *Dictyostelium discoideum* (*Continued*)

Dictyostelium	1 MEPIPPSNGN KNNSMDKQPQ QPQQPQQQQQ	1 MVRFVPWNVP LYRRLETMAV AIYAMVLPVC
discoideum	QQQQQRRDQR NSKLNELNET	LIMAFNLIVI PLFWGIAIPY
(fungus)	51 ERVRNRFISH EFHKLDRTKS RIDAPKISFS	51 LVWMFYFDTK HESGGRRVSL VRNSILWRYF
	DSESESDSEF FLAKRNTNNN	RDYFPISLII NSNYDPKKNY
	101 NQNNTSPTFS SANGKQSNLT QRKINTQIQS	101 IFAYH <u>PH</u> GII SIGAFCNFAT NANNIDEKLP
	KQPTNNNVQP LTDDEGTINH	GLKVHLLTLE SNFKIPFLRD
	151 SNHHHHHHNQ NNNGNNNNNN NNNNNNKIS	151 VLMSFGMSSV SKKSCENILN SGAGESICLV
	TPPKQEEKMT MNGLFTLRPS	VGGAEESLDA RPGLNEITLK
	201 ILSSESNGSS YRGFLNLLLI LLITASFRLV	201 KRKGFIKLAL VNGASLVPVY SFGENDIYDQ
	ILNHLLYGIR INLDLYKISE	VPNPRGSLVR KIQTKIKDLT
	251 YHRWPGVMIS LMINLFIIAA YLIEKAAAKQ	251 GIAPPLFMGR GIFNYDFGLL PVRHKIVTVV
	LLPDRICYLL RIINCAAVII	GEPIDIPKIK SPTDQVIEHY
	301 VPSGSIIAFS PNPASGIIVM ILICTFSMKI	301 HQIYVEALQN LFDKHKNSCA DKETGNLKIN
	ISYAYENSKQ RKLNPDNKKF	1) P106, 2) H107, 3) <u>P146</u> , 4) R149, 5) G182,
	351 VIDPTNTSIY PNNLSLRSTY WFMLVPTLVY	6) G183,
	QLSYPRSPKI <u>R</u> KGYLLRRIV	7) E186, 8) R202, 9) G204, 10) F205, 11) A209,
	401 EALSLSLLIL WMVNQYMLPL VQNSIEPLEK	12) V217,
	IDIVLIVERI MKLSLPNLYV	13) P218, 14) <u>F222</u> , 15) G223, 16) E281
	451 WLLGFYVFFH LYLNIVAEIT RFGDREFYRD	
	WWNSTGLDYF WRTWNMPVHH	
	501 WMVVLIYTPM RRRGFSKNMG YFMCFFVSAI	
	FHELVISIPF HSLKLWGFFG	
	551 IMSQMVLIAL TKNLMNGRNL GNVIFWISIV	
	LGQPLVVLLY YRNFVLENPE	
	601 WYRNVEPPTS PPVMPFY	
	1) G213, 2) L222, 3) K329, 4) S332, 5) R341,	
	6) <u>P376</u> ,	
	7) T377, 8) R391, 9) Q415, 10) P419, 11) L443,	
	12) W451,	
	13) L452, 14) F455, 15) F459, 16) E468, 17)	
	<u>F472</u> , 18) D474,	
	19) R3475, 20) F477, 21) Y3478, 22) W481, 23)	
	W482,	
	24) N483, 25) W491, 26) W494, 27) N495, 28)	
	P497, 29) P509,	
	30) F525, 31) S528, 32) H532, 33) P539, 34)	
	F548, 35) Q554,	
	36) L557, 37) N572, 38) G582, 39) Q583, 40)	
	P584, 41) Y590	

The completely conserved residues in DGAT1s and DGAT2s are underlined and listed below the sequences. The underlined P and F residues are conserved in all DGATs.

[38]. This phenylalanine residue is conserved in all plants except *Brassica napus* [GenBank:AAD45536.1] and conserved in all fungi except mold (*Dictyostelium discoideum* and *Polysphondylium pallidum*) (Figure 11A). Rape DGAT1 is the only sequence in the plant group with a serine residue in the place of phenylalanine in the sequence alignment (Figure 11A). Since this gene is isolated from suspension cultures of *Brassica napus* [39] and the native form is not available, it is not known if this replacement was caused by mutations due to cell culture conditions, considering that cell culture could cause significant changes in gene expression [40]. It is interesting to note that a similar phenylalanine residue is not present in any of the animal DGAT1s or any of DGAT2s.

In cattle, a nonconservative substitution of lysine by alanine (K232A) in DGAT1 (Table 4) is directly responsible for the quantitative trait loci (QTL) variation with the lysine-encoding allele being associated with higher milk fat content [41]. Figure 11B shows that the lysine residue is conserved in mammalian DGAT1s but not in plants, or fungi, or other animals (fly, frog, insect and worm) except one of the two forms from dog and zebra-fish. The wild-type cattle DGAT1 shows the normal lysine at 232 position [GenBank:AAL49962.1] (Table 4). A similar lysine residue is not found in any of DGAT2s.

#### Discussion

#### **DGAT** classification

The nomenclature of proteins derived from DNA sequences in GenBank databases can lead to confusion in some cases. One of the utilities of this extensive sequence analysis is to use the completely conserved amino acid residues in respective sequence blocks of DGAT1 and DGAT2 subfamilies as signatures of DGAT proteins for classification. It is generally accepted that DGATs are divided into DGAT1 and DGAT2 subfamilies. However, more than two forms of DGATs are

present in a number of species (Table 1). Phylogenetic analysis and multiple sequence alignment classify all 117 DGAT protein sequences into DGAT1 and DGAT2 subfamilies (Figures 1, 3, 4, 5, 6, 7). Furthermore, six DGATs currently designated as DGAT2s in the databases are reclassified here as DGAT1s due to their close alignment with DGAT1s and these sequences containing all 41 completely conserved amino acid residues found in DGAT1s. These six reclassified sequences are [Gen-Bank:EGC41804.1, GenBank:EEQ31683.1, GenBank: AAY40785.1, GenBank:AAD40881.1, GenBank: EFY86774.1 and GenBank:EFY97444.1] (Table 1).

DGATs which diverge from the DGAT1 and DGAT2 subfamilies have been recently reported from *Arabidopsis thaliana* [GenBank:AAN31909.1] [16], Burning Bush (*Euonymus alatus*) [GenBank:GU594061.1] [14], peanut (*Arachis hypogaea*), [GenBank:AY875644.1] [15], caster bean (*Ricinus communis*) [GenBank:XP\_002519339.1] and yeast (*Rhodotorula glutinis*) [GenBank:DG315417.1]. Phylogenetic analysis and multiple sequence alignment indicate that these sequences are completely different from DGAT1s and DGAT2s. In fact, none of the completely conserved residues in DGAT1s (41 residues) and DGAT2s (16 residues) is present in these new DGATs in the multiple sequences alignment (data not shown). This sequence divergence is in contrast to the general belief that the active sites of enzymes should be conserved during the evolution because all catalyze the same/similar biochemical reaction. Therefore, this sequence divergence raises an important question how completely different proteins could perform the same biochemical reaction.

#### DGAT properties and amino acid composition

This study has analyzed the properties and amino acid composition of 109 full-length DGATs from 70 organisms. The average DGAT1s are 171 amino acid residues longer than DGAT2s resulting in approximately 20 kDa difference in the molecular mass. Other DGAT properties are similar: both are basic proteins under neutral pH with

Α			В					
(580) 580 599	(625)	625 630	640	650 6	60 670	680 690	700	710 721
DmDGAT1-NP_609813.1(359) VNVVMALFQQWIIPSVRN	DmDGAT1-NP_609813.1(396)	LALPNHLCWI	LCF <mark>FYLMF</mark> HSFLNA	VG <mark>ELL</mark> N <mark>FADR</mark> NFY	C <mark>DWWNA</mark> NNIDTF <mark>WR</mark> TWI	NMEVHRWCVRHLYIEVV(	MCYSSRQASTINE	LF <mark>SAVEHE</mark> YLVSVP
TcDGAT1-NP_995724.1(409)VNVVMALFQUMITESVRN TcDGAT1-XP_975142.1(286)TOTTLOTFOOYMTESVRN	DmDGAT1-NP_995724.1(446) TcDGAT1-XP_975142.1(323)	LALPNHLOW	LOPEYLMEHSFLNA	GETTHEADRNEY	GOWWNANNIDTFWRTW	MPVHRWCVRHLYIPVVÇ	UGYGKOVAGITVE	FTSAFFHEYMVSVP
BbDGAT1-AAZ22403.1(288) FLTOLQVGLTOCWMVPAION	BbDGAT1-AAZ22403.1(327)	LAVPNHLI <mark>WI</mark>	LIF <mark>FYWLFH</mark> SCPNA	VAELMQFODREFY	RDWWNSESITYF <mark>W</mark> LNWI	IPVHKWCIRHFYK <mark>P</mark> MLF	RGSSKWAARTAVE	la <mark>safeh</mark> gylvsi <mark>p</mark>
BtDGAT1-NP_777118.2(288) FLTQLQVGLIQQWMVPAION	BtDGAT1-NP_777118.2(327)	LAVPNHLIWI	LIFFYWLFHSCLNA	VAELMQFGDREFY	R DWWNSESITYFWQNWI	IPVHKWCIRHFYKPML	RCSSKWAARTAVE	LASAFEHEYLVSIP
SsDGAT1-NP_001103634.1(288) FLICEOVGLICOVMVPAID SsDGAT1-NP 999216.1(288) FLICEOVGLICOVMVPTION	SsDGAT1-NP_001103634.1(327) SsDGAT1-NP 999216.1(327)	LAVPNHLIWI LAVPNHLIWI	LIFEYWLEHSCLNA	VAELMOFGDREFY	RDWWNSESTTYFWQNWI	IPVHKWCLRHFYKPMLF	RGSSKWAARTGVE	ASAFFHEYLVSIP
CfDGAT1-XP_849176.1(297) FLTQLQVGLTQQWMVFTIQN	CfDGAT1-XP_849176.1(336)	LAVPNHLI <mark>WI</mark>	LIF <mark>FYWEF</mark> HSCMNA	VAELMQFGDREFY	rd <mark>wwn</mark> sesvtyf <mark>w</mark> Qn <mark>wi</mark>	IPVHKWCLRHFYK <mark>P</mark> MLF	RGSSKWVARTGVE	FA <mark>SAFEHEYLVSIP</mark>
CfDGAT1-XP_858062.1(304) FLTQLQVGLTQQWMVPTION	CfDGAT1-XP_858062.1(343)	LAVPNHLIW	LIFFYWFFHSCMNA	VAELMQFGDREFY	R DWWNSESVTYFWQNWI	IPVHKWCLRHFYKPMLF	RGSSKWVARTGVE	FASAFFHEYLVSIP
MmDGAT1-XP 001090134.1(290) FFTCLOVGLICOWMVETION	MmDGAT1-XP_001090134.1(329)	LAVPNHLIW	LIFFYWLFHSCLNA	VAELMOFGDREFY	RDWWNSESVTIFWQNWI	IPVHKWCIRHFYKPML	RCSSRWMARTGVE	LASAFFHEYLVSVP
PtDGAT1-XP_520014.2(287) FFTQLQVGLIQQWMVFTIQN	PtDGAT1-XP_520014.2(326)	LAVPNHLI <mark>WI</mark>	LIFFYWLFHSCLNA	VAELMÕFGDREFY	rd <mark>wwn</mark> sesvtyf <mark>w</mark> Qn <mark>wi</mark>	IPVHKWCIRHFYK <mark>P</mark> MLF	RGSSKWMARTGVF	LA <mark>S</mark> AFE <mark>HE</mark> YLVSV <mark>P</mark>
MmDGAT1-NP_034176.1(298) FFT0_QVGLIQ0WMVFTION	MmDGAT1-NP_034176.1(337)	LAVPNHLIWI	LIFFYWFFHSCLNA	VAELLQFGDREFY	R. DWWNAESVTYFWQNWI	IPVHKWCIRHFYKPML	HCSSKWVARTG	TSAFEHEYLVSVP
OcDGAT1-XP 002724427.1(302) FFTOLOVGLIOOWMVFTION	OcDGAT1-XP 002724427.1(341)	LAVPNHLIWI LAVPNHLIWI	LIFEYWLEHSCLNA	VGEL MRFGDROFY	RDWWNAESVTYFWQNWI	IPVHKWCIRHFYRPMLF IPVHKWCSRHFYRPMOF	RCNSRWVARTGVE	ASAFFHEYLVSIP
MdDGAT1-XP_001371565.1(430) FFT01LVGLI00WMVFTI0N	MdDGAT1-XP_001371565.1(469)	LAVPNHFI <mark>WI</mark>	LIF <mark>FYWFF</mark> HSCLNA	LAELMQFGDREFY	rdwwnsesvtyfwQnwi	IPVHKWCLRHFYK <mark>P</mark> ALF	MCVSKWMAKTGV <mark>F</mark>	la <mark>s</mark> afe <mark>h</mark> eylvs <mark>vp</mark>
DrDGAT1-NP_001002458.1(303) LLTQLLVGLTQQWMVFIIRS	DrDGAT1-NP_001002458.1(342)	LAVPNHLL <mark>WI</mark>	LIF <mark>FYSEF</mark> YSSMNF	MAELL REG DREFY	r d <mark>wwn</mark> setityf <mark>w</mark> QN <mark>WI</mark>	IPVHKWCLRHFYK <mark>P</mark> LLF	RGAGKLLSQSAV <mark>F</mark>	FA <mark>SAFFHEYLVSVP</mark>
SkDGAT1-XP_002736160_1(280)_FLT0TVLGLT00WTVFT10N	DrDGAT1-NP_956024.1(335) SkDGAT1-XP_002736160_1(319)	LAVENHEIW	LIFEYWYEHSSMNE	VAEIMQFGDREFY	RUWWNSETIPIFWSNWI	TPVHKWCLRHFYKPML	RGANRLAAQIAVE	ILSAFFHEYLVSVP
DdDGAT1-XP_645633.2(404) SLSLIILWMVNQYMLFLVQN	DdDGAT1-XP_645633.2(443)	LSLPNLYVWI	LLG <mark>FYVFFHL</mark> YLNI	VAEITREGDREFY	R DWWNSTGLDYFWRTWI	MPVHHWMVVLIYTPMR	RGFSKNMGYFMCF	TVSAITHELVISIP
PpDGAT1-EFA85004.1(391) FLSALIFWMVEQYMVFLVEN	PpDGAT1-EFA85004.1(430)	LSLPNLYV <mark>W</mark> I	LLG <mark>FYVEF</mark> HLYLNI	CAELTREG DREFY	R. <mark>DWWN</mark> STGLDYF <mark>WR</mark> TWI	MPVHHWMVVMIYT <mark>P</mark> MRF	RGYSKNAGYFMC <mark>F</mark>	FV <mark>SAIFHE</mark> LVIS <mark>VP</mark>
AcDGAT2-EGC41804.1(305) GLSVFIWLTSACYAAPVLRN PhDGAT1-EEH17170 1(298) GLSVFIWLTSACYAAPVLRN	AcDGAT2-EGC41804.1(344) PbDGAT1-EEH17170_1(337)	LSTISLIIWI	LAGEFALEQSELNA	LAEVMHFGDREFY	TEWWNSPSVGTYWRTWI	KPVYHFMRRHIFSPLVC	RGWSPFAASVMVE	FFSGIIHELLVGIP
AoDGAT2-EEQ31683.1(299) GLGVFIWI SAOYAAEVLRN	AoDGAT2-EEQ31683.1(338)	LSTISLVIW	LAGFFALFQSFLNA	LAEVMRFGDREFY	NEWWNSPSVGSYWRSWI	RPVYHFMKRHIFSPLV0	RGWSPFAASLVVF	TFSAVIHEVLVGIP
MaDGAT2-EFY86774.1(270) CLSAFIWEASF <mark>QY</mark> AA <mark>E</mark> VLQN	MaDGAT2-EFY86774.1(309)	LSTISLVI <mark>W</mark> I	lag <mark>f</mark> fal <mark>f</mark> Qsflna	LA <mark>EVLRFGDR</mark> SFY	d <mark>dwwn</mark> seslga <mark>ywr</mark> twi	NRPVYTYFKRHVYV <mark>P</mark> MIC	RGWSPWAASCAV <mark>F</mark>	FV <mark>SAVIHEVLVGVP</mark>
MaDGAT2-EFY97444.1(299) CLSAFIWEASFOYAAEVLON	MaDGAT2-EFY97444.1(338)	LSTISLVIW	LAGFFALFQSFLNA	LAEVEREGORSEY	DDWWNSESLGAYWRTWI	RPVYTYFKRHVYVPMIC	RCWSPWAASCAVE	TOAT THEY MITTYP
CVDGAT1-EFN50697.1(222) GGLGIMLFSIEQY10FTIDN	CvDGAT1-EFN50697.1(261)	LSIPTLYWW	LAMFYTLFDLWLNI	IAELIRFODREFY	KEWWNATTVGEYWRLWI	OPVHKWMLRHVYFPLIF	HCVPKFHAGLMVF	TVSAVEHEVLVGVP
PpDGAT1-XP_001770929.1(247) VFLGLGGGFIIEQYINPTVKN	PpDGAT1-XP_001770929.1(285)	LSLPVLYV <mark>WI</mark>	l <mark>Clfy</mark> Clfhlwlni	VA <mark>ELLRFGDREF</mark> Y	K <mark>owwna</mark> qtveey <mark>wrm</mark> wi	M <sup>MP</sup> VHKWMVRHIYF <mark>P</mark> SIF	aglskkaavllv <mark>e</mark>	AI <mark>SALEHE</mark> VIIG <mark>VP</mark>
SmDGAT1-XP_002964165.1(219) VENCEMGFIIGQYMNPIIRN 0-DCAT1 ND_001054860 2(237) J DOC 00000000000000000000000000000000000	SmDGAT1-XP_002964165.1(257)	LSIPTLYVWI	LGFFYCFFHLWLNI	VAEILCFGDREFY	KOWWNAKSVDEYWRLWI	MPVHRWLVRHVYFPCL	LCLHKQFAILV	VI <mark>SGIEHLICIAVP</mark>
SbDGAT1-NP_001054869.2(337) IFTGLOGFILEOTINEIVVN SbDGAT1-XP 002439419.1(314) IFTGLOGFILEOTINEIVVN	SbDGAT1-XP_001054869.2(375) SbDGAT1-XP_002439419.1(352)	LSLPNVILWI	LCMFYRLFHLWLSI LCMFYCLFHLWLNI	LAEILREGDREFY	KDWWNAKTIDETWRKWI	MPVHKWVVRHIIFPCM MPVHKWMLRHIYFPCIF	NGISKEVAVLISE	TVSAVEHELCVAVE
SbDGAT1-XP_002437165.1(288) VFTGLMGFIIEQYINFIVKN	SbDGAT1-XP_002437165.1(326)	LSVPTLYV <mark>W</mark> I	LCMFYCFFHLWLNI	LAELLCFGDREFY	KD <mark>WWNA</mark> KTVEEY <mark>WRM</mark> WI	MPVHKWIIRHIYF <mark>P</mark> CIF	KGFSRGVAILVSF	V <mark>SAVEHEICIAVP</mark>
ZmDGAT1-2-EU039830(292) VFTGLMGFIIEQYINFIVKN	ZmDGAT1-2-EU039830(330)	LSVPTLYV <mark>W</mark> I	LCMFYCFFHLWLNI	VAELLCFGDREFY	K.DWWNAKTVEEYWRMWI	MPVHKWIIRHIYF <mark>P</mark> CIF	KGFSRGVAILIS <mark>F</mark>	V <mark>SAVEHEICIAVP</mark>
VgDGAT1-ABV21945.1(322) IFTGFMGF11EQYINF1VEN FnDGAT1-ACO55635.1(272) IFTGFMGF11EQYINF1VEN	VgDGAT1-ABV21945.1(360) FpDGAT1-ACO55635.1(310)	LSVPNLYVWI LSVPNTYVWI	LOMFYCEFHLWLNI	LAELLOPEDREFY	KOWWNAQTIEEYWRLWI	MPVHKWIVRHLYFPCLE	NGIPKGAAILVAP	TSATEHRICIAVE
OeDGAT1-AAS01606.1(331) IFTGFMGFIVEQYINFIVEN	OeDGAT1-AAS01606.1(369)	LSVPNLYVW	LCMFYCFFHLWLNI	LAELLCFGDREFY	KOWWNAKTVEEYWRMWI	MPVHKWMVRHIYF <mark>P</mark> CLF	NCMPRGGAILIAF	IISAIFHELCIAVP
PfDGAT1-AAG23696.1(333) IFTGLMGFIIEQYINFIVQN	PfDGAT1-AAG23696.1(371)	LSVPNLYV <mark>W</mark> I	LCM <mark>F</mark> YCFFHLWLNI	LAELLOFGDREFY	K.D <mark>WWNA</mark> RTVEEY <mark>WR</mark> MWI	MPVHKWMVRHIYC <mark>P</mark> CLQ	NGIPKIVAVLIAF.	V <mark>SAIFHELCVAVP</mark>
AtDGATI-NP 179535 1(320) TETGEMORTIROVINETURN	NEDGATI-NP 179535 1(358)	LSVPNLYV <mark>W</mark> I LSVPNLYV <mark>W</mark> I	LOMFYCEFHLWLNI	LAELLOPEDREFY	KOWWNAKTIDEY WRMWI	MPVHKWMVRHIYFPCLE	SKIPKTLATIAP	VSAVEHELCIAVE
BjDGAT1-AAY40784.1(303) IFTGIMGFIIEQYINFIVRN	BjDGAT1-AAY40784.1(341)	LSVPNLYVW	LCMFYCFFHLWLNI	LAELLCFGDREFY	KDWWNAKSVGDYWRMWI	MPVHKWMVRHVYFPCLE	IKIPKVPAIIIAF	VSAVEHELCIAVP
BnDGAT1-AAD45536.1(303) IFTGLMGFIIEQYINFIVRN	BnDGAT1-AAD45536.1(341)	LSVPNLYV <mark>W</mark> I	LCM <mark>F</mark> YCF <mark>F</mark> HLWLNI	LAELLCFGDREFY	KD <mark>WWNA</mark> KSVGDY <mark>WRMWI</mark>	MPVHKWMVRHVYF <mark>P</mark> CLF	IKIPKVPAIIIA <mark>F</mark>	LV <mark>SAVEHELCIAVP</mark>
BJDGAT2-AAY40/85.1(303) IFTGLMGFILEOYINFIVEN TmDGAT1-AAM03340.2(319) VETCLMGFILEOYINFIVEN	BJDGAT2-AAY40785.1(341) TmDGAT1-AAM03340.2(357)	LSVPNLYV <mark>W</mark> I	LCMFYCFFHLWLNI	LAELLCFGDREFY LAELLRFGDREFY	KOWWNAKSVGDYWRMWI KOWWNAKTVAEYWKMWI	MPVHKWMVRHVYFPCLE	IKIPKVPAIIIAP.	VSAVEHELCIAVP
EaDGAT1-AAV31083.1(306) IFTGLMGFIIEQYINFIVQN	EaDGAT1-AAV31083.1(344)	LSVPNLYVW	LCMFYCLFHLWLNI	LAELLCFGDREFY	KOWWNAKTVEEYWRMWI	MPVHKWMVRHIYFPCLE	NGIPKGVAFVISF	VSAVEHELCIAVE
GmDGAT1-AAS78662.1(297) IFTGVMGFIIEQYINFIVQN	GmDGAT1-AAS78662.1(335)	LSVPNLYV <mark>W</mark> I	LCMFYCFFHLWLNI	LAELLREGDREFY	QD <mark>WWNA</mark> KTVEDY <mark>WRMW</mark> I	MPVHKWMIRHLYF <mark>P</mark> CLF	HGIPKAVALLIA <mark>F</mark>	LV <mark>SALEHELCIAVP</mark>
IDGAT1-AAW51456 1(308) TETCVMGETTEOVINETVON	GMDGAT1-BAE93461.1(341)	LSVPNLYV <mark>W</mark> I LSVPNLYV <mark>W</mark> I	LOMFYCEFHLWLNI	LAELLREGDREFY	KOWWNAKTVEDYWRMWI KOWWNAKTFEEYWRMWI	MPVHKWMIRHLYYPCLE	HGLPKAAALLIAP	VSALPHELCIAVP
MtDGAT1-ABN09107.1(338) IFTGVMGFIIEOYMNFIVON	MtDGAT1-ABN09107.1(376)	LSVPNVYVWI	LCMFYCFFHLWLNI	LAELLREGDREFY	KOWWNAOTVEE YWRMWI	MPVHKWMVRHVYFPCI	FGIPKGAAALTAF	V <mark>SAVEHELCIAVP</mark>
VvDGAT1-XP_002279345.1(314) IFTGVMGFIIEQYINFIVQN	VvDGAT1-XP_002279345.1(352)	LSVPNLYV <mark>W</mark> I	LCMFYCFFHLWLNI	LAELLRFGDREFY	KD <mark>WWNA</mark> KTVEEY <mark>WR</mark> MWI	MPVHKWMVRHLYF <mark>P</mark> CLF	NGISKGVSVVIA <mark>F</mark>	AI <mark>SAIFHE</mark> LCIAV <mark>P</mark>
JCDGA11-ABB84383.1(318) IFTGFMGF1IEQYINFIVQN VfDGAT1-D0356680.1(323) IFTGFMGF1FEQYINFIVON	JcDGAT1-ABB84383.1(356) VfDGAT1-DO356680.1(361)	LSVPNLYV <mark>W</mark> I	LCMFYCFFHLWLNI LCMFYCFFHLWLNI	LAELLREGDREFY	KOWWNARTVEEYWRMWI KOWWNARTVEEYWRMWI	MPVHKWMVRHIYFPCLE	HKIPRGVALLIAF	FVSAVEHELCIAVP
RcDGAT1-XP_002514132.1(322) IFTGFMGFIIEQYINFIVQN	RcDGAT1-XP_002514132.1(360)	LSVPNLYVW	LCLFYCFFHLWLNI	VAELLREGDREFY	KOWWNAKTVEEYWRMWI	MPVHKWMVRHIYFPCLE	RKIPRGVAIVIAF	TVSAVEHELCIAVE
PtDGAT1-XP_002308278.1(249) IFTGFMGFIIEQYINFIVKN	PtDGAT1-XP_002308278.1(287)	LSVPNLYV <mark>W</mark> I	lcmfycffhlwlni	LAELLCFGDREFY	KO <mark>WWNA</mark> RTVEEY <mark>WRMW</mark> I	MMPVHKWMVRHIYF <mark>P</mark> CLF	NKIPKGLAILIA <mark>F</mark>	V <mark>SAVEHELCIAVP</mark>
PtDGAT1-XP_002330510.1(290) IFTGFMGFIIEQYINPIVON Consensus(580) FTGL FILEOXI R VON	PtDGAT1-XP_002330510.1(328)	LSVPNLYV <mark>W</mark> I SVDNI VUMI	LOMFYCEFHLWLNI	LAELIRE DREFT	KOWWNART EEYWRMWI	MPVHKWMVRHIMFPOLI	NKIPKWAALLIAE	TVSAVEHELCIAVP
<b>Figure 4 Identification of completely conserved amino acid residues in sequence motifs 4-5 of DGAT1s.</b> (A) DGAT1-Motif 4 (QP Block), (B) DGAT1-Motif 5 (LWLFFEFDRFYWWNWNPPFSHP Block; The first boxed F residues are also conserved in DGAT2s; The boxed Y, W and the second F residues are mutated in TmDGAT1 and the boxed H residue is mutated in MmDGAT1). Refer to Figure 3 legend for additional								
information.								

high isoelectric points (Table 2). The frequency of functional amino acid residue groups between DGAT1 and DGAT2 subfamilies is also very similar in terms of charged residues, acidic residues, basic residues, polar residues and hydrophobic residues (Table 2). The remarkable feature of DGAT1s and DGAT2s is that both subfamilies of proteins contain more than 40% of hydrophobic residues (Table 2). These high amounts of hydrophobic residues in DGATs are in agreement with them being integral membrane proteins [33,34] with multiple transmembrane domains [13,33,34], localized to endoplasmic reticulum of plant and animal cells [13,34], and associated with mitochondria in COS-7 cells [35,42] and lipid bodies in 3T3-L1 adipocytes [42]. The membrane association of the proteins presents extra huddle to purification of recombinant DGATs from any source [43,44].

#### Catalytic and regulatory domains of DGATs

Generally speaking, critical amino acid residues of proteins are conserved during the evolution because

they are essential for enzymatic activity. The conserved amino acid residues are clustered at the active centers of the enzymes. Multiple sequence alignment has shown that DGAT1s and DGAT2s have 41 and 16 completely conserved amino acid residues, respectively. Most of them are located at the carboxyl termini of DGATs (Table 3). This sequence analysis suggests that the catalytic domains of DGATs are located at the carboxyl termini of the proteins. This is supported by mutations of some completely conserved amino acid residues in the C-termini of these proteins resulted in complete loss of the enzymatic activity of DGATs (see below). This suggestion is in line with our previous assignment of the catalytic domains of ADPGlc-dependent  $\alpha$ -1,4-glucosyltransferases and  $\alpha$ -1,6-glucan hydrolases from plants and prokaryotes at the carboxyl termini of the enzymes because of the presence of the conserved amino acid residues and sequence motifs in the different isoforms from diverse organisms [6,9].

A		В			
(747)	747 761	(777)	777 ,79	90 801	
DmDGAT1-NP_609813.1(499)	WA <mark>F</mark> MGMMG <mark>Q</mark> IP <mark>L</mark> SAI	DmDGAT1-NP_609813.1(525)	MG <mark>N</mark> IIVWAS-IIL <mark>G(</mark>	2PLCIMA <mark>Y</mark> YHD	
DmDGAT1-NP_995724.1(549)	WA <mark>F</mark> MGMMG <mark>Q</mark> IP <mark>L</mark> SAI	DmDGAT1-NP_995724.1(575)	M <mark>GN</mark> IIVWAS-IIL <mark>G</mark>	<mark>QP</mark> LCIMA <mark>Y</mark> YHD	
TcDGAT1-XP_975142.1(426)	WA <mark>F</mark> MGMMG <mark>Q</mark> IPLSNI	TcDGAT1-XP_975142.1(452)	MG <mark>N</mark> IVVWAS-LII <mark>G</mark>	2PLCIMM <mark>Y</mark> YHD	
BbDGAT1-AAZ22403.1(430)	WA <mark>F</mark> TGMMA <mark>Q</mark> IP <mark>L</mark> AWI	BbDGAT1-AAZ22403.1(453)	Y <mark>GN</mark> AAVWLS-LII <mark>G</mark>	<mark>2P</mark> VAVLM <mark>Y</mark> VHD	
BtDGAT1-NP_777118.2(430)	WA <mark>F</mark> TGMMA <mark>Q</mark> IP <mark>L</mark> AWI	BtDGAT1-NP_777118.2(453)	Y <mark>GN</mark> AAVWLS-LII <mark>G</mark>	2PVAVLM <mark>Y</mark> VHD	
OaDGAT1-NP_001103634.1(430)	WA <mark>F</mark> TGMMA <mark>Q</mark> IP <mark>L</mark> AWI	OaDGAT1-NP_001103634.1(453)	Y <mark>GN</mark> AAVWLS-LII <mark>G</mark>	2PVAVLM <mark>Y</mark> VHD	
SsDGAT1-NP_999216.1(430)	WAFTGMMAQIPLAWI	SsDGAT1-NP_999216.1(453)	Y <mark>GN</mark> AAVWLS-LII <mark>G</mark>	2PVAVLM <mark>Y</mark> VHD	
CfDGAT1-XP_849176.1(439)	WAFTGMMAQIPLAWI	CfDGAT1-XP_849176.1(462)	Y <mark>GN</mark> AAVWLT-LII <mark>G</mark>	2PVAVLM <mark>Y</mark> VHD	
CfDGAT1-XP_858062.1(446)	WAFTGMMAQIPLAWI	CfDGAT1-XP_858062.1(469)	Y <mark>GN</mark> AAVWLT-LII <mark>G</mark>	2PVAVLM <mark>Y</mark> VHD	
HsDGAT1-NP_036211.2(429)	WAFTGMMAQIPLAWF	HsDGAT1-NP_036211.2(452)	Y <mark>GN</mark> AAVWLS-LII <mark>G</mark>	2PIAVLM <mark>Y</mark> VHD	
MmDGAT1-XP_001090134.1(432)	WAFTGMMAQIPLAWF	MmDGAT1-XP_001090134.1(455)	Y <mark>GN</mark> AAVWLT-LII <mark>G</mark>	2PIAVLM <mark>Y</mark> VHD	
PtDGAT1-XP_520014.2(429)	WAFTGMMAQIPLAWF	PtDGAT1-XP_520014.2(452)	Y <mark>GN</mark> AAVWLS-LII <mark>G</mark>	2 <mark>P</mark> IAVLM <mark>Y</mark> VHD	
MmDGAT1-NP_034176.1(440)	WAFTAMMAQVPLAWI	MmDGAT1-NP_034176.1(463)	Y <mark>GN</mark> AAVWVT-LII <mark>G</mark>	2PVAVLM <mark>Y</mark> VHD	
RnDGAT1-NP_445889.1(440)	WAFTAMMAQVPLAWI	RnDGAT1-NP_445889.1(463)	Y <mark>GN</mark> AAVWVT-LII <mark>G</mark>	2PVAVLM <mark>Y</mark> VHD	
UCDGA I 1-XP_002724427.1(444)	WAFTGMMAQVPLAWI	UCDGAT1-XP_002724427.1(467)	YG <mark>N</mark> AAVWLT-LII <mark>G</mark>	2PVAVLM <mark>Y</mark> VHD	
MdDGAT1-XP_001371565.1(572)	WSFTGMMAQIPLAWI	MdDGAT1-XP_001371565.1(595)	Y <mark>GN</mark> AAVWLS-LII <mark>G</mark>	2PVAVLM <mark>Y</mark> VHD	
DrDGAT1-NP_001002458.1(445)	WA <mark>F</mark> MGMMAQIPLAWL	DrDGAT1-NP_001002458.1(468)	Y <mark>GN</mark> AAVWLS-LII <mark>G</mark>	2PIAVLM <mark>Y</mark> VHD	
DrDGAT1-NP_956024.1(438)	WAFLGMMLQVPLAIL	DrDGAT1-NP_956024.1(461)	Y <mark>GN</mark> AAVWMS-LIIG(	2PIAVLMYVHD	
SkDGAT1-XP_002/36160.1(422)	WAFMGMMTQVPLALV	SkDGAT1-XP_002736160.1(447)	YS <mark>NMMVW</mark> MS-VIA <mark>G</mark>	2PVAIMM <mark>Y</mark> VHD	
DdDGAT1-XP_645633.2(546)	WGFFGIMSQMVLIAL	DdDGAT1-XP_645633.2(570)	LG <mark>NVIFWIS-IVLG</mark>	2PLVVLLYYRN	
PpDGAT1-EFA85004.1(533)	WAFFGIMSOMVLIAL	PpDGAT1-EFA85004.1(557)	LG <mark>NVIFW</mark> FS-IVL <mark>G(</mark>	2PMVVLLYYRN	
ACDGA I 2-EGC41804.1(447)	VAFAGMILQLPLIAV	AcDGAT2-EGC41804.1(476)	IG <mark>N</mark> CVFWVSFCLV <mark>G</mark>	2PLGALLYFFA	
PDDGAT1-EEH1/1/0.1(440)	VAFAGMVLQLPLIAA	PbDGAT1-EEH17170.1(469)	IG <mark>N</mark> CVFWVSFCLV <mark>G</mark>	2PLAALLYFFA	
AODGAT2-EEQ31683.1(441)	VAFAGMMLQLPLIAA	AoDGAT2-EEQ31683.1(470)	IG <mark>N</mark> SVFWLSFCLV <mark>G</mark>	2PLGALLYFFA	
MaDGAT2-EFY86/74.1(431)	VAFLGMFLQLPLIAL	MaDGAT2-EFY86774.1(461)	MG <mark>N</mark> VIFWVSFTIF <mark>G</mark> Ç	2PFAALMYFYA	
MaDGAT 2-EFT 97444.1(400)		MaDGA   2-EFY 9/444.1(490)	MG <mark>NVIEWVSETIE</mark> G	2PF'AALMYF'YA	
TyDGAT1-AAP94209.1(471)		I gDGA I 1-AAP94209.1(497)	VG <mark>N</mark> CFEWETECES <mark>G</mark>	2PLGILI <mark>Y</mark> WYL	
CVDGATI-EFN50097.1(304)		CVDGA I 1-EFN50697.1(391)	VG <mark>N</mark> IIEWVSECEV <mark>G</mark>	2PLAMILYYHD	
PDDGAT1-XP_001770929.1(388)		PpDGAT1-XP_001770929.1(415)	VG <mark>N</mark> MVEWE'EE'CLV <mark>G</mark>	2PMCLLLYYHD	
OCDCAT1 ND 0010E4960 2(478)		SmDGAT1-XP_002964165.1(387)	VG <mark>N</mark> MIEWE'EE'CIV <mark>G</mark> Ç	2PMCVLLYYHD	
SEDCAT1 VD 002420410 1(455)		OSDGAT1-NP_001054869.2(505)	VG <mark>N</mark> MIEWEEECIY <mark>G</mark>	2PMCLLLYYHD	
SUDGAT1-XP_002439419.1(435)		SDDGAT1-XP_002439419.1(482)	VG <mark>N</mark> MIEWEEECIY <mark>G</mark>	2PMCVLLYYHD	
ZmDCAT1-2-EU030830(433)		SDDGAT1-XP_002437165.1(456)	VG <mark>N</mark> MIEWE'EE'SIV <mark>G</mark> Q	2PMCVLLYYHD	
ZIIIDGATI-Z=L0059650(455)		ZmDGAT1-2-E0039830(460)	VG <mark>N</mark> MIEWE'EESIV <mark>G</mark> Q	2PMCVLLYYHD	
$F_{p}DGAT1-ACO55635 1(413)$		VGDGAT1-ABV21945.1(490)	VGNMIFWCFFSIFG	2PMCVLLYYHD	
OeDGAT1-AAS01606 1(472)	WARTCIMFOURIVII.	$ \begin{array}{c} EpDGAT1 ACO55635.1(440) \\ OpDGAT1 AASO1606 1(400) \\ \end{array} $	VGNMIFWCFFCILG	2PMSLLLYYHD	
PfDGAT1-AAG23696 1(474)	WAFSCIMLOVPLVIV	DEDCAT1 AAC22606 1(501)	VGNMIEWCEFSILG	2PMCLLLY YHD	
NtDGAT1-AAF19345 1(472)	WAFMGIMFOVPLVII.	PIDGATI-AAG23090.1(301)	VGNMMFWCFFCIFG	2PMCVLLIIHD	
AtDGAT1-NP 179535 1/461)	WAFLGIMFOUPLVFT	NLUGAT 1-AAF19345.1(499)	VGNMTEMETECTLG	2PMCVLLYYHD	
BiDGAT1-AAY40784 1(444)	WAFMGIMFOVPLVFT	ALDGATI-NP_1/9555.1(46/)	VGNMIEWEIFCIEG		
BnDGAT1-AAD45536.1(444)	WAFMGIMFOVPLVFT	DJDGAT 1-AAT 40764.1(470)	VGNMIEWESECTEG		
BiDGAT2-AAY40785.1(444)	WAFIGIMFOVPLVFT	DIDGAT 1-AAD45550.1(470)	VGNMIFGSASCIFG		
TmDGAT1-AAM03340.2(460)	WAFIGIMFOVPLVLT	$D_{D}DGAT2-AAT40765.1(470)$	VGNMIEWEJECIEG		
FaDGAT1-AAV31083.1(447)	WAFFGTMLOVPLVLT	$F_{a}$ $CAT1 AAV21082 1(474)$	VGNMIEWEIFCILG		
GmDGAT1-AAS78662.1(438)	WAFGGTMFOVPLVFT	EdDGAT1-AAV31003.1(474)	VGNMEWESECTEG	2PMCLLLI I HD	
GmDGAT1-BAE93461.1(444)	WAEGGTMFOVPLVLT	GmDGAT1-RA578002.1(403)	VGNMIEWEIESILG		
LiDGAT1-AAW51456.1(449)	WAEGGTMEOVPLIT	LIDCAT1 AAWE14E6 1(476)	VGNMIEWEIESILG		
MtDGAT1-ABN09107.1(479)	WAFIGIMFOVPLVI.T	LJUGA   1-AAW 51450.1(470) MtDGAT1_ARN00107 1(506)	VOIMTEWEIESILG	DEMONT TANK	
VvDGAT1-XP 002279345.1(455)	WAFIGIMFOVPLVLV	MUDGAT1_YD 000070024E 1(400)	VGINTEWFIECTLG	2 FMCVLL <sup>Y</sup> YHD	
JcDGAT1-ABB84383.1(459)	WAFIGIMFOTPLVGT	VVDAII-APP042031(482)	VGIMILEWLEESTLG	2 FMCVLL <sup>Y</sup> YHD	
VfDGAT1-D0356680.1(464)	WAFTGIMFOIPLVGT	JCDGAT1 PO256680 1(401)	VGNMIEWEIECILG	2PMCVLLYYHD	
RcDGAT1-XP_002514132_1(463)	WAFFGIMFOIPLVVI	VIDGATI-DQ35080.1(491)	VGINITEWFIFCILG	2PMCLLLYYHD	
PtDGAT1-XP_002308278.1(390)	WAFTGTMLOVPLVVT	RCDGATI VD 002200270 1(417)	VGINITEWFFFCTLG	2PMCVLLYYHD	
PtDGAT1-XP 002330510.1(431)	WAFIGIMFOVPLVVT	PEDCAT1_YD 002220510 1(450)	VOINTENLEESTLO	2 FMCVLLI YHD	
Consensus(747)	WAF GIM O PL T	FLUGATI-AF_002300310.1(458)	VOINTEWLFFOILG	OMONT I AAND	
	·	CONSENSUS(///)	VGINMITEM ET C(	JENCVILIIHD	

Several lines of evidence suggest that the regulatory domains of DGATs are located at the amino termini of the proteins. First, a recent study showed that the amino terminal domain of DGAT1 of mouse is not required for the catalytic activity of DGAT1 but may be involved in regulating enzyme activity and dimer/tetramer formation [45]. Second, the N-terminal region of mouse DGAT2 or yeast DGAT2 is not essential for DGAT activity *in vitro* [33,35]. Finally, mutagenesis of a putative protein kinase SnRK1 (SNF1-related kinase 1) target site at S197 to alanine in TmDGAT1 results in a 38%-80% increase in DGAT1 activity, and over-expression of the mutated TmDGAT1 in *Arabidopsis* results in a 20%-50% increase in oil content on a per seed basis [25]. This serine residue is conserved in most of the plants and located at the N-termini of DGAT1s (Figure



9A). All of the above mentioned sequence analysis and experimental evidence support the concept that the catalytic and regulatory domains of DGATs are located at the C- and N-termini of the enzymes, respectively.

## Functional significance of less conserved motifs in DGAT2s

Recent studies have reported functional motifs in DGATs including putative neutral lipid-binding domain (FLXLXXX<sup>n</sup> in mouse DGAT2) [34], mitochondrial targeting signal (RXKXXK in mouse DGAT2) [35] and ER retrieval motif (LKLEI in tung DGAT2) [13]. However, the conserved sequence motifs identified by our extensive sequence analysis do not contain any of these reported motifs. In our analysis, the putative neutral lipid-binding domain [34] which was shown to be presented in proteins that either bind to or metabolize neutral lipids [36], is only modestly conserved in animal DGAT2s and not present in any plant DGAT2 (Figure

8A). Similarly, the putative mitochondrial targeting signal is only found in a few animal DGAT2s but not conserved in any plant or fungi DGAT2 (Figure 8B). This sequence analysis also shows that the pentapeptide (LKLEI) ER-retrieval motif identified at the extreme carboxyl terminus of tung DGAT2 [13] is only modestly conserved in plant DGAT2s but not in animal or fungus DGAT2s (Figure 8C). All these studies point out that less conserved regions in a subset of DGATs may play specific roles in TAG biosynthesis in that particular subset of organisms.

# Functional significance of the completely conserved residues

Multiple sequence alignment has shown that 55 DGAT1s and 54 DGAT2s have 41 and 16 completely conserved amino acid residues, respectively, although only two residues are completely conserved among all DGATs (Table 3). It is likely that these completely

А		В	}	С			
(242) 34	42 358	(261) 36	.1 376	(419) 418 436			
(342) 3-3 ACDGAT2-YP 003225477 1(247) 5		(301) 30 ACDGAT2-YP 003225477 1(264) 31					
BLDCAT2-DAA21853 1(227) V	TI PNPKCEVKT AT PUC	B+DCAT2-DAA21853 1(244)		B+DCAT2-DAA21853 1(301) VPUSKETTTVVGEFTTTP			
MmDGAT2-NP_080660.1(254) V	TLKNRKGEVKLALBHG	MmDGAT2-NP_080660.1(271) Ar	UVPTYSEGENEVYK	MmDGAT2-NP 080660.1(328) VPYSKPTTTVVGEPTTVPK			
RnDGAT2-NP_001012345.1(254) V	TLENEKGEVKLALEHG	RnDGAT2-NP 001012345.1(271) AC	UVPTYSEGENEVYK	RnDGAT2-NP 001012345.1(328) VPYSKPTTTVVGEPTTVPK			
HsDGAT2-AAO88896.1(254)	TLENRKGEVKLALRHG	HsDGAT2-AAO88896.1(271) AD	LVPIYSEGENEVYK	HsDGAT2-AAO88896.1(328) VPYSKPITTVVGEPITIPK			
PaDGAT2-XP 002822304.1(254)	TLRNRKGFVKLALRHG	PaDGAT2-XP 002822304.1(271) AD	OLVPIYSEGENEVYK	PaDGAT2-XP 002822304.1(328) VPYSKPITTVVGEPITIPK			
XtDGAT2-NP_989372.1(227)	TLKORKGFVKVALOHG	XtDGAT2-NP_989372.1(244) AD	OL <mark>VPVYSEGENE</mark> AYK	XtDGAT2-NP_989372.1(301) VPYANPITTVVGEPITVPK			
DrDGAT2-NP_001025367.1(227)	MLKK <mark>RKGFVKLAL</mark> KQG	DrDGAT2-NP_001025367.1(244) AD	DL <mark>VPVYSEGENEV</mark> YK	DrDGAT2-NP_001025367.1(301) VPYCKPITTVVGEPITVPK			
IpDGAT2-NP_001188005.1(227)	TLRN <mark>RKGFIKLAL</mark> QKG	IpDGAT2-NP_001188005.1(244) AD	DL <mark>VPVYSEGENE</mark> SYK	IpDGAT2-NP_001188005.1(301) VPYGKAINTVVGEPITVPK			
BtDGAT2-XP_875499.3(272) L	TLRN <mark>RKGFVRLAL</mark> RHG	BtDGAT2-XP_875499.3(289) AS	SL <mark>VPVYS<mark>EG</mark>ENDVFR</mark>	BtDGAT2-XP_875499.3(346) MPLARPITTVVGRPIPVPQ			
HsDGAT2-NP_835470.1(207)	TLQK <mark>RKGFVRLAL</mark> RHG	HsDGAT2-NP_835470.1(224) AS	SL <mark>VPVYS<mark>EG</mark>ENDIFR</mark>	HsDGAT2-NP_835470.1(281) LPEAVPITTVVGRPIPVPQ			
PtDGAT2-XP_527842.2(207)	TLQK <mark>RKGFVRLAL</mark> RHG	PtDGAT2-XP_527842.2(224) 🔼 S	SL <mark>VPVYS<mark>EG</mark>ENDIFR</mark>	PtDGAT2-XP_527842.2(281) LPEAVPITTVVGRPIPVPQ			
DdDGAT2-XP_635762.1(197) I	TLKK <mark>RKGFIKLAL</mark> VNG	DdDGAT2-XP_635762.1(214) 🗛 S	SL <mark>VPVYSEGENDIY</mark> D	DdDGAT2-XP_635762.1(270) LPVRHKIVTVVGEPIDIPK			
PpDGAT2-EFA83646.1(197)	LLKK <mark>RKGF</mark> VK <mark>LA</mark> FAHG	PpDGAT2-EFA83646.1(214) AS	SL <mark>VPIYS<mark>EGE</mark>NDIFD</mark>	PpDGAT2-EFA83646.1(270) LPHRRKIVSVVGEPIDVPK			
NvDGAT2-XP_001630435.1(204)	TLKA <mark>RYGFIKLAI</mark> RNG	NvDGAT2-XP_001630435.1(221) AS	SL <mark>VPVYA<mark>FG</mark>ENDVF</mark> N	NvDGAT2-XP_001630435.1(277) LPHRKPIHVVVGSPIEVEK			
NvDGAT2-XP_001633322.1(205)	TLKE <mark>RKGF</mark> VK <mark>IAM</mark> KTG	NvDGAT2-XP_001633322.1(222) AS	SL <mark>VPVFSEGENEL</mark> YS	NvDGAT2-XP_001633322.1(278) LPHRRPINTVVGAPITVER			
NvDGAT2-XP_001635548.1(180)	TLKN <mark>RKGFVKY</mark> ALKHG	NvDGAT2-XP_001635548.1(197)	SL <mark>VPVYSEGENELY</mark> C	NvDGAT2-XP_001635548.1(253) LPHKRSVCTVVGAPVIVKR			
TgDGAT2-XP_002187643.1(202)	L <mark>LKNRKGFVRLAI</mark> QHG	TgDGAT2-XP_002187643.1(219) TF	PLVPAFSEGENDVFD	TgDGAT2-XP_002187643.1(275) MPYRRPICTVVGKPIPVQK			
MdDGAT2-XP_001365685.1(202)	CITQ <mark>RKGFIKMAL</mark> THG	MdDGAT2-XP_001365685.1(219)	LVPVFSFGENDLFY	MdDGAT2-XP_001365685.1(275) MPYRKPIHTVVGQPIPVQQ			
CiDGAT2-XP_002120879.1(202)	LLKQ <mark>RLGFLKLAI</mark> TNG	CiDGAT2-XP_002120879.1(219) VP	PLVPVFSFGDHALWE	CiDGAT2-XP_002120879.1(275) IPYRRSVHTVVGEPIEVPQ			
CeDGAT2-NP_505413.1(216) L	TLKK <mark>RKGFVKIAL</mark> QTG	CeDGAT2-NP_505413.1(233) AC	2LVPCYSEGENDIEN	CeDGAT2-NP_505413.1(289) LPFRKPINTVLGAPISVTK			
CeDGAT2-NP_872180.1(206) L1	TLKK <mark>RKGF</mark> VKIALQTG	CeDGAT2-NP_872180.1(223) AC	<u>DIVPCYSEGENDIFN</u>	CeDGAT2-NP_872180.1(279) LPERKPINTVLGAPISVTK			
TcDGAT2-XP_975146.1(207) I	ILKN <mark>RKGFVKLAL</mark> RNG	TcDGAT2-XP_975146.1(224)	PLVPVISEGEPELFD	TcDGAT2-XP_975146.1(277) IPQRRPITTVVGHPIEVTK			
AgDGAT2-NP_983542.1(326)	VLNK <mark>RKGFIKLAL</mark> ETG	AgDGAT2-NP_983542.1(344) VN	ILVPIYARGETDCFN	AgDGAT2-NP_983542.1(400) LPFRNPINVVGKPVYDK			
ScDGAT2-NP_014888.1(283)	ILNKRKGFIKLAIQTG	ScDGAT2-NP_014888.1(301) IN	ILVPVFAHGEVDCYN	ScDGAT2-NP_014888.1(357) LPFRAPINVVGRPIYVEK			
SPDGAT2-XP_001/13160.1(211)	VLKKRFGFVKLAFLTG	SPDGAT2-XP_001/13160.1(228) SS	SLVPCHARGESDIFE	SPDGA12-XP_001/13160.1(284) MPWRKPINIVVGEPIDVPK			
CrDGAT2-XP_001694904.1(200)		CrDGAT2-XP_001694904.1(219)	IVPVYHHGNSQVLD	CDGAT2-XP_001694904.1(262) LPRRRNTYMVCGRPVPVTR			
CIDGAT 2-XP_001693189.1(197) A	YLSSRIGF VRLAVQHG	CrDGAT2-XP_001693189.1(214) AP	PLVPVWARGQTRATS	CDGAT2-XP_001693189.1(265) MPHREPLITVVGRPTPVPE			
DCAT2-XP_003063539.1(198) 13	YI KKNYCEVKIAI ETC	PDCAT2-XP_001758758 1(208) 01	TUDERCHCOTINCYK	PDCAT2-XP_003083339.1(200) TPTATPTATVGELTETTO			
PDGAT2-XP_001777726 1(187) VI	FLKORVCEVRIAMEAC	PDDGAT2-XP_001730730.1(200) B1 PDDGAT2-XP_001777726 1(204) SE	TVPTFCFCORNAVK	PDDGAT2-XP_001737726 1(250) VPVPTPMYYVV/CKPTPVPK			
SmDGAT2-XP_002972054 1(218) VI	FLKKREGEVRVATETC	SmDGAT2-XP_002972054 1(235) AF	T VPVFCFCOTFAVK	SmDGAT2-XP_002972054 1(281) VPCOOFMYWAVCNPTEVTK			
PsDGAT2-ABK26256 1(206) AT	FIRRRHCEVRVATETC	PsDGAT2-ABK26256 1(223) CE	V V V V C V C V C V V V V V V V V V V V	PSDGAT2-ABK26256 1(269) TEXTREPTOWWCKPTEWKP			
SpDGAT2-AA089590 1(200) A3	YLKKRHGEVRVATETG	SpDGAT2-AA089590 1(217) 5F	LVPVECHGONEARS	SpDGAT2-AA089590 1(263) TPYRSPTDWWCKPMEWKO			
RcDGAT2-XP 002528531.1(218) AT	FLKARRGEVRVAMEMG	RcDGAT2-XP 002528531.1(235) KP	TVPVFCFGOSNVYK	RCDGAT2-XP 002528531.1(281) LPLORPMHVVGKPTEVKO			
VgDGAT2-ACV40232.1(216)	FISRRKGFIKVAIOTV	VgDGAT2-ACV40232.1(233) TE	<b>LVPVFFEGOAHTYK</b>	VaDGAT2-ACV40232.1(279) LPCKNPTVVVGRPITVEK			
VfDGAT2-DG356682(200) AI	FLKARRGFIRIAMOTG	VfDGAT2-DG356682(217) TF	PLVPVFCEGOMHTEK	VfDGAT2-DG356682(263) LPFKNPMHVVVGRPIEVKO			
EoDGAT2-ACO35365.1(205) AI	FLKARKGFVRIAMETG	EoDGAT2-ACO35365.1(222) RF	IVPVFCEGOSYVYK	EoDGAT2-ACO35365.1(268) IPFOHPMOVVVGKPIGLKR			
HvDGAT2-BAJ85730.1(211) AI	FLKS <mark>RKGFVKIAM</mark> QSG	HvDGAT2-BAJ85730.1(228) CF	LVPVFC <mark>FG</mark> QSKAYK	HvDGAT2-BAJ85730.1(274) IAFSSPMHVVVGRPIELKK			
OsDGAT2-NP_001057530(218) AI	FLKS <mark>RKGF</mark> VK <mark>IAM</mark> QSG	OsDGAT2-NP_001057530(235) CF	PL <mark>VPVFC<mark>EG</mark>QSYAYK</mark>	OsDGAT2-NP_001057530(281) IPFPTPMHVVVGRPIEVEK			
OsDGAT2-NP_001047917(216) AI	FLKS <mark>R</mark> KGFVKIAMETG	OsDGAT2-NP_001047917(233) SF	PL <mark>VPVF<mark>AFG</mark>QSY<mark>V</mark>YK</mark>	OsDGAT2-NP_001047917(279) IPFATPMHVVVGRPIEVKK			
SbDGAT2-XP_002452652.1(212) AH	FLKS <mark>R</mark> KGFVK <mark>IAI</mark> EMG	SbDGAT2-XP_002452652.1(229) CP	VVPVF <mark>AFG</mark> QSY <mark>V</mark> YK	SbDGAT2-XP_002452652.1(275) IPFATPMHVVVGRPIEVVK			
ZmDGAT2-NP_001150174.1(211) A	FLKP <mark>RKGFVKIAI</mark> EMG	ZmDGAT2-NP_001150174.1(228) CF	P <mark>VVPVF<mark>AFG</mark>QSY<mark>V</mark>YK</mark>	ZmDGAT2-NP_001150174.1(274) IPEATPMHVIVGRPIEVVK			
AtDGAT2-NP_566952(193) VI	FLSR <mark>RRGFVRIAM</mark> EQ <mark>G</mark>	AtDGAT2-NP_566952(210) SF	PL <mark>VPVFC</mark> EGQAR <mark>V</mark> YK	AtDGAT2-NP_566952(256) LPCRQPMHVVVGKPIEVTK			
BnDGAT2a-ACO90187(195) V	FLSS <mark>RRGF</mark> VRIAMEQG	BnDGAT2a-ACO90187(212) AF	PL <mark>VPVFC<mark>FG</mark>QSRAYK</mark>	BnDGAT2a-ACO90187(258) IPYRHPIHVVVGKPIQVAK			
BnDGAT2b-ACO90188(195) 🔽	FLSS <mark>RRGFVRIAI</mark> EQG	BnDGAT2b-ACO90188(212) 🖪 🛛	PL <mark>VPVFC<mark>EG</mark>QSRAYK</mark>	BnDGAT2b-ACO90188(258) IPYRHPIHVVVGKPIQVTK			
VvDGAT2-XP_002263626(215) AF	FLKS <mark>RKGFVRIAM</mark> EMG	VvDGAT2-XP_002263626(232) RF	PL <mark>VPVFCEG</mark> QSR <b>V</b> YK	VvDGAT2-XP_002263626(278) LPYRHPMDVVVGRPIEVKK			
EaDGAT2-ADF57328.1(209) AI	FLKS <mark>RRGFVRIAM</mark> EMG	EaDGAT2-ADF57328.1(226) TP	PLVPVFC <mark>EG</mark> QSR <mark>V</mark> YK	EaDGAT2-ADF57328.1(272)			
GmDGAT2-ACU20344.1(195) AI	FIKQ <mark>RRGFVRIAL</mark> QMG	GmDGAT2-ACU20344.1(212) LF	LVPVFC <mark>EG</mark> QTKAYK	GmDGA12-ACU20344.1(258) IPFKNPLYTVVGRPIELEK			
MEDGA I 2-ACJ84867.1(199) AS	YLKARRGFVRIALEKG	MtDGA12-ACJ84867.1(216) HP	LVPVFCEGQSDIYK	MEDGATZ ADUS0220 1(277)			
HaDGA I 2-ABU50328.1(214) AI	FLRTRKGFVRLAMETN	HaDGAT2-ABU50328.1(231)	LVFVFGEGQSYVYK	HaDGA I 2-ABU50328.1(2//) IPEROPMHVVVGRPIHFKK			
	ILNIRGEVRIAMEMG	UEDGATZA AAK84170 1(230) KE	LVPVECEGQTKVYK	UEDGATZA AAK84170 1(202) TPUENDAVOGRPTVPKR			
	VERREGETRIAVQ'I'G	ULUGATZR AAK841/9.1(236) AS	LVPTISHGENELYE	UDOAT2R AAK041/9.1(292) TERRET YTTVGKPIPVPS			
	IN DECEMENT	UIDGATZD-AAK84180.1(228)	LUDUESTCE FUUR	Consensus(418) LDVD DI MUNCODDI N.K.			
Consensus(342) L	IN REGEVELAL G	Consensus(301) A	LVEVESEGE EVIK	CONSCISUA(TIO) LFIK FI VVVGKPI V K			
Figure 7 Identification of completely conserved amino acid residues in sequence motifs 4-6 of DGAT2s. (A) DGAT2-Motif 4 (RGFA Block),							



conserved amino acid residues are critical for DGAT enzymatic activities. These residues may be involved in substrate binding, direct catalysis, and/or maintenance of protein structure including oligomer formation. The importance of some conserved residues in DGAT1s has been demonstrated by site-directed mutagenesis (Table 4). Mutagenesis at H426 in mouse DGAT1 to alanine impairs the ability of DGAT1 to synthesize triacylglycerols, retinyl and wax esters in an *"in vitro"* acyltransferase assay [45]. This histidine residue is completely conserved in Motif 5 of all DGAT1s (Figure 4B). Similarly, mutagenesis at Y392, W395 and F439 in *Tropaeolum majus* DGAT1 eliminates nearly all activity [25]. These three residues are also completely conserved in Motif 5 of all DGAT1s (Figure 4B). All four residues are located in the most conserved region of DGAT1s in which 23 completely conserved residues are located in Motif 5 of the multiple sequence alignment (Figure 4B).

The importance of the completely conserved residues in DGAT2s is also supported by site-directed mutagenesis. Mutagenesis at H161, P162 and H163 sites, and the triple mutant in mouse DGAT2 results in a substantial loss of activity (Table 4) [34]. Mutation at the corresponding sites at H193 and H195 in DGAT2 of baker's yeast results in complete loss of the activity (Table 4) [33]. These three resides are located in the highly conserved Motif 1 (PH Block) of DGAT2s (Figure 6A). These results suggest that they may be located at the active center of DGAT1s, but the precise roles of these residues either involved in substrate binding or catalysis

Α	В	С
(110) 112 110	<b>(20)</b> 90 100 110 121	(176 192
(113) (12) (113) (12) (113) (12) (113) (12) (12) (12) (12) (12) (12) (12) (12	AcDGAT2-XP 003225477.1 (51) -WLSKSKVEKC OV SV OWV SFLIMGWACT	(476) 476 485 AcDGAT2-XP 003225477.1(374) AEVIEVN
BtDGAT2-DAA21853.1 (53) FLVLGV	BtDGAT2-DAA21853.1 (30) TWLNRSKVEKCLQVISVLQWVLSFLVLGVACS	BtDGAT2-DAA21853.1(354) ETEVLEVN
MmDGAT2-NP_080660.1 (80) FLVLGV2	MmDGAT2-NP_080660.1 (57) TWLNRSKVEKCLQVISVLQWVLSFLVLGVACS	MmDGAT2-NP_080660.1(381) ETEVLEVN
RnDGAT2-NP_001012345.1 (80) FIVLGV2.	RnDGAT2-NP_001012345.1 (57) TWLNRSKVEKFLQVISVLQWVLSFLVLGVACS	RnDGAT2-NP_001012345.1(381) ETEVLEVN
HSDGAT2-AAQ88896.1 (80) FLVLGVA	HSDGAT2-AAQ88896.1 (57) TWLNRSKVEKCLQVISVLQWVLSFIVLGVACS	HSDGAT2-XP 002822204 1(381) HTEVLEVN
YEDGAT2-NP 989372 1 (53) FT TT CV	YEDGAT2-NP 989372 1 (30) DWLSKSSVERCLOVISVLOWVLSFLVLGWACS	XtDGAT2-NP 989372 1(354) DSEMIETV
DrDGAT2-NP 001025367.1 (53) FT TMGT2	DrDGAT2-NP 001025367.1 (30) PWLTRSKMVKHLOVISVLOFIMTFLIMGTACS	DrDGAT2-NP 001025367.1(354) ESDTLIIH
IpDGAT2-NP_001188005.1 (53) FLALGI	IpDGAT2-NP_001188005.1 (30) PWLSRSKLVKCLQAISVLQFVLSFLALGTAAT	IpDGAT2-NP_001188005.1(354) ESDVLLID
BtDGAT2-XP_875499.3 (98) FLFLGPI	BtDGAT2-XP_875499.3 (75) TPSMKTLKKQWLEVLSTYQYVLCFLFLGPFFS	BtDGAT2-XP_875499.3(399) ASTHLTFI
HsDGAT2-NP_835470.1 (33) FLFMGP1	HSDGAT2-NP_835470.1 (10) PTTSKTLQKQFLEAVGAYQYVLTFLFMGPFFS	HsDGAT2-NP_835470.1(334) ASTCLTFI
PtDGAT2-XP_527842.2 (33) F1 FMGP1	PtDGAT2-XP_527842.2 (10) PTTS KTLQKQF LEAVGAYQYVLTF1 FMGPFFS	PtDGAT2-XP_527842.2(334) ASTCLTFI
PDDGAT2-FEA83646 1 (27) VENCHIM	PDGAT2-FFA83646 1 (4) WIDASLELHERTETTAVIEVIMUPOCHEMAR	PDGAT2-FFA83646 1(322) FSDDTKTH
NvDGAT2-XP 001630435.1 (29) FFFGALL	NvDGAT2-XP 001630435.1 (6) FAPLRIPLHRRLETAAVALYYYSFFFGALLGF	NvDGAT2-XP 001630435.1(330) KEDKLEIL
NvDGAT2-XP_001633322.1 (29) FLFGHIL	NvDGAT2-XP_001633322.1 (6) WAPVKLPFKRRLETLTVILIVHSFIFGHILGC	NvDGAT2-XP_001633322.1(331) KDVHLEIC
NvDGAT2-XP_001635548.1 (5) FFLGHTL	NvDGAT2-XP_001635548.1 (1)MCTIFFLGHTLGT	NvDGAT2-XP_001635548.1(307) KETRLIIQ
TgDGAT2-XP_002187643.1 (28) FLCLALC	TgDGAT2-XP_002187643.1 (5) FAPLSVPLQRRLQTASVAQWIFSFLCLALCCT	TgDGAT2-XP_002187643.1(328) HDSHLEFI
MdDGAT2-XP_001365685.1 (28) FIMLAKL	MDDGAT2-XP_001365685.1 (5) FAPLNIPLERRLQTGAV_QWIFSFIMLAKLWF	MdDGAT2-XP_001365685.1(328) EQETLVFK
CeDGAT2-NP 505413 1 (42) TEALAAT	CeDGAT2-NP_505413.1 (19) WAPLNIPLARECTECTECTECTEFT	CeDGAT2-XP_002120079.1(327) $m_{\rm DCRBTTN}$
CeDGAT2-NP 872180.1 (32) TLFTPVL	CeDGAT2-NP_872180.1 (9) WAPLNIPLARRIOTLGAIHFFFITLFTPVLVL	CeDGAT2-NP_872180.1(332) PTTOLVIN
TcDGAT2-XP_975146.1 (31) LAFGTFI	TcDGAT2-XP_975146.1 (8) FAPLHIPLERRLQTLAAGCWFTTLAFGTFIGT	TcDGAT2-XP_975146.1(331) KDIHLEFE
AgDGAT2-NP_983542.1(113) FIYMTVL	AgDGAT2-NP_983542.1 (90) VAPLRIPARRRLQTLVVAWHTSSFIYMTVLVL	AgDGAT2-NP_983542.1(454) GKELKIVE
ScDGAT2-NP_014888.1 (71) FVLFSIF	ScDGAT2-NP_014888.1 (48) CCPLATPFERRIQTLAVAWHTSSFVLFSTFT	ScDGAT2-NP_014888.1(411) DAELKIVG
SpDGAT2-XP_001713160.1 (37) HSVSLTL	SPDGAT2-XP_001/13160.1 (14) TPPISKDSRRNVSHWLQALAVFLHSVSLTLTA	SPDGAT2-XP_001/13160.1(337) RISELKLS
CrDGAT2-XP_001694904.1 (38) 10GWMHV	CrDGAT2-XP_001693189.1 (1)MPLAKLENVULEYAATATYVSATYTSVUL	$CrDGAT2-XP_001094904.1(318) DXPLV13-$
OtDGAT2-XP_003083539.1 (17) LFHALVV	OtDGAT2-XP_003083539.1 (1)MSRSIVDHGVLLVWLGLFHALVVVV	OtDGAT2-XP_003083539.1(319) DVDLVVC-
PpDGAT2-XP_001758758.1 (26) GLHLNVL	PpDGAT2-XP_001758758.1 (14)VLTFLAIVVWMGGLHLNVLVV	PpDGAT2-XP_001758758.1(307) DSPLFVY-
PpDGAT2-XP_001777726.1 (22) GFHINFI	<pre>PpDGAT2-XP_001777726.1 (10)LWSIIAMVMWLGGFHINFIVG</pre>	PpDGAT2-XP_001777726.1(303) DDTLLVY-
SmDGAT2-XP_002972054.1 (53) AIHFNTI	SmDGAT2-XP_002972054.1 (41)VLSWIAVLIWLGATHFNTIVG	SmDGAT2-XP_002972054.1(334) DTMLHVY-
PSDGAT2-ABK26256.1 (41) TIHFDAI	PSDGAT2-ABK26256.1 (28)GFRSLVALIMWLGT1HFDAILV	PSDGAT2-AAO80500 1(322) DTHLRVL-
RCDGAT2-XP 002528531.1 (53) STHENTE	RCDGAT2-XP 002528531.1 (39)NIFHALLATSTWIGSTHEN FLL	RcDGAT2-XP 002528531.1(334)
VgDGAT2-ACV40232.1 (49) SIHFILF	VgDGAT2-ACV40232.1 (35)SIPRALIALSLWIGSIHFILFLL	VgDGAT2-ACV40232.1(332) (LELKIL-
VfDGAT2-DG356682 (38) SFHFILF	VfDGAT2-DG356682 (24)NIFQSVLALAIWLGSFHFILFLV	VfDGAT2-DG356682(316) ILKLEIF-
EoDGAT2-ACO35365.1 (40) LIHFNVA	EoDGAT2-ACO35365.1 (27)VVRTTVALALWIGLIHFNVALV	EoDGAT2-ACO35365.1(321) IFRLRVL-
HvDGAT2-BAJ85730.1 (46) GIHFNVL	HvDGAT2-BAJ85730.1 (33)LPRTIAALALWIGGTHFNVLLI	HvDGAT2-BAJ85730.1(327) (LHLRVL-
OsDGAT2-NP_001057530 (53) GIHFNVF	OSDGAT2-NP_001057530 (40)LPRITVALALWIGGIHFNMFLV	OSDGAT2-NP_001057530(334) (LHLRVL-
SbDGAT2-XP_001047917 (31) ATHENAF	SbDGAT2-XP 002452652.1 (34)PLRTTVALALWIGATHFNAFILL	SbDGAT2-XP_002452652.1(328) ILBLBVI-
ZmDGAT2-NP 001150174.1 (46) AIHFNAF	ZmDGAT2-NP_001150174.1 (33)PIRTTVALALWIGAIHFNAFLV	ZmDGAT2-NP_001150174.1(327) ILQLRVL-
AtDGAT2-NP_566952 (28) AIHFNVA	AtDGAT2-NP_566952 (14)NQFHSIIAMAIWLGAIHFNWALV	AtDGAT2-NP_566952(308) ILELKIL-
BnDGAT2a-ACO90187 (30) AIYLNLA	BnDGAT2a-ACO90187 (16)NIFHAVTAISICLSAIYLNLALV	BnDGAT2a-ACO90187(311) ILQLNIL-
BnDGAT2b-ACO90188 (30) AIYLNLA	BnDGAT2b-ACO90188 (16)NILHAVTAISICLSAIYLNIALV	BnDGAT2b-ACO90188(311) ILQINIL-
VVDGA12-XP_002263626 (50) AIHEVVA	FaDGAT2-ADE57328 1 (32)MDTVVANTINT CATHENVALV	VVDGAT2-XP_002203020(331) Π_QLKLL- FaDGΔT2-ΔDF57328 1(325) ΤΤ ΗΤΕΤΤ-
GmDGAT2-ACU20344.1 (30) ATHENTA	GmDGAT2-ACU20344.1 (17)VFKTVPALVLYLGATHFNLALI	GmDGAT2-ACU20344.1(311) NLELKTF-
MtDGAT2-ACJ84867.1 (34) SIHFNIA	MtDGAT2-ACJ84867.1 (21)TFRSILALSLWLGSTHFNIAII	MtDGAT2-ACJ84867.1(315) FLELKIV-
HaDGAT2-ABU50328.1 (49) SIHLNVF	HaDGAT2-ABU50328.1 (35)PLIHTALAMIAWLGSIHLNVFIV	HaDGAT2-ABU50328.1(330) NLQLRIM-
OeDGAT2-ADG22608.1 (48) SVHFNII	OeDGAT2-ADG22608.1 (34)SIFHTILALVLWLGSWHFNIIVV	OeDGAT2-ADG22608.1(329)
UrDGAT2A-AAK84179.1 (47) MSICMFI	UrDGAT2A-AAK84179.1 (24) YAPLRVPLRRRLOTLAV LWCSMMSICMFIFF	
UTUGAIZB-AAK84180.1 (42) LPICLII	Consensus (90) T TT AT MIT T V TV	
Figure 8 Sequence analysis of imp	portant motifs in less conservative regions of DGAT2s. Multiple	e sequence alignment was performed
Using 54 Iuli-length DGAT2 protein se	quences from 44 organisms (listed in Table T). Color code and relate	a mornation are described briefly in
Figure 2 legend and with details in "N	vietnoas section. The motifs are boxed within the sequence alignme	nt. (A) Putative neutral lipid-binding
domain (FLXLXXX in mouse DGAT2) (	B) Mitochondrial targeting signal (RXKXXK in mouse DGAT2), (C) ER re	etrieval motif (LKLEI in tung DGAT2).

are not clear. Further experiments are required to assess the contribution of the other completely conserved residues to the enzymatic activity of DGATs.

## Functional significance of the less-well conserved residues in site-directed mutants

The importance of some less conserved residues in DGATs has also been demonstrated by site-directed mutagenesis (Table 4). As described above, mutation at

S197 (a putative SnRK1 target site) in TmDGAT1 results in a 38%-80% increase in DGAT1 activity. This serine residue is conserved in most of the plants (Figure 9A). In addition, mutagenesis at E145 in Motif 1 of *Tropaeolum majus* DGAT1 results in the loss of almost half of the activity [25]. This glutamate residue is conserved in all plant DGAT1s and most other DGAT1s except bird, chimpanzee, *Dictyostelium discoideum, Polysphondylium pallidum* and *Metarhizium acridum* 

### Table 4 Site-directed and natural mutants of DGATs and their effects on enzymatic activity and TAG accumulation.

DGAT	Amino acid sequence with altered residues underlined	Mutation	Activity	Reference
MmDGAT1- NP_034176.1 [ <i>Mus musculus</i> ]	1 MGDRGGAGSS RRRRTGSRVS VQGGSGPKVE EDEVRDAAVS PDLGAGGDAP APAPAPAHTR DKDGRTSVGD 71 GYWDLRCHRL QDSLFSSDSG FSNYRGILNW CVVMLILSNA RLFLENLIKY GILVDPIQVV SLFLKDPYSW 141 PAPCVIIASN IFVVAAFQIE KRLAVGALTE QMGLLLHVVN LATIICFPAA VALLVESITP VGSVFALASY 211 SIMFLKLYSY RDVNLWCRQR RVKAKAVSTG KKVSGAAAQQ AVSYPDNLTY RDLYYFIFAP TLCYELNFPR 281 SPRIRKRFLL RRVLEMLFFT QLQVGLIQQW MVPTIQNSMK PFKDMDYSRI IERLLKLAVP NHLIWLIFFY 351 WFFHSCLNAV AELLQFGDRE FYRDWNAES VTYFWQNWNI PVHKWCIRHF YKPMLRHGSS KWVARTGVFL 421 TSAFFHEYLV SVPLRMFRLW AFTAMMAQVP LAWIVGRFFQ GNYGNAAVWV TLIIGQPVAV LMYVHDYYVL 491 NYDAPVGV	H426A	-100%	[45]
TmDGAT1- AAM03340.2 [Tropaeolum majus]	1 MAVAESSQNT TTMSGHGDSD LNNFRRRKPS SSVIEPSSSG FTSTNGVPAT GHVAENRDQD RVGAMENATG 71 SVNLIGNGGG VVIGNEEKQV GETDIRFTYR PSFPAHRRVR ESPLSSDAIF KQSHAGLFNL CIVVLIAVNS 141 RLIIENLMKY GWLIDTGFWF SSRSLGDWSI FMCCLTLPIF PLAAFIVEKL VQRNHISELV AVLLHVIVST 211 AAVLYPVIVI LTCDSVYMSG VVLMLFGCIM WLKLVSYAHT SSDIRTLAKS GYKGDAHPNS TIVSCSYDVS 281 LKSLAYFMVA PTLCYQPSYP RSSCIRKGWV VRQFVKLIVF IGLMGFIIEQ YINPIVRNSK HPLKGDFLYA 351 IERVLKLSVP NLYWLCMFY SFFHLWLNIL AELLRFGDRE FYKDWWNAKT VAEYWKMWM PVHRWMVRHL 421 YFPCLRNGIP KEGAIIIAFL VSGAFHELCI AVPCHVFKLW AFIGIMFQVP LVLITNYLQE KFSNSMVGNM 491 IFWFIFCILG QPMCVLLYYH DLINLKEK	E145V S197A P216R Y392A Y392A/ W395G F439R	-43% +38-80% -100% -80% -100% -100%	[25]
MmDGAT2- NP_080660.1 [ <i>Mus musculus</i> ]	1 MKTLIAAYSG VLRGERRAEA ARSENKNKGS ALSREGSGRW GTGSSILSAL QDIFSVTWLN RSKVEKQLQV 71 ISVLQWVLSFLVLGVACSVI LMYTFCTDCW LIAVLYFTWL AFDWNTPKKG GRRSQWVRNW AVWRYFRDYF 141 PIQLVKTHNL LTTRNYIFGY <u>HPH</u> GIMGLGA FCNFSTEATE VSKKFPGIRP YLATLAGNFR MPVLREVLMS 211 GGICPVNRDT IDYLLSKNGS GNAIIIVVGG AAESLSSMPG KNAVTLKNRK GFVKLALRHG ADLVPTYSFG 281 ENEVYKQVIF EEGSWGRWVQ KKFQKYIGFA PCIFHGRGLF SSDTWGLVPY SKPITTVVGE PITVPKLEHP 351 TQKDIDLYHA MYMEALVKLF DNHKTKFGLP ETEVLEVN	F80A L81A L83A H161A P162G H163A H161A/ P162G/ H163A	-66% -85% -100% ~50-60% ~50-60% < 20% < 20%	[34]
ScDGAT2 (Dga1p)- NP_014888.1 [Saccharomyces cerevisiae]	1 MSGTFNDIRR RKKEEGSPTA GITERHENKS LSSIDKREQT LKPQLESCCP LATPFERRLQ TLAVAWHTSS 71 FVLFSIFTLF AISTPALWVL AIPYMIYFFF DRSPATGEVV NRYSLRFRSL PIWKWYCDYF PISLIKTVNL 141 KPTFTLSKNK RVNEKNYKIR LWPTKYSINL KSNSTIDYRN QECTGPTYLF GYHPHGIGAL GAFGAFATEG 211 CNYSKIFPGI PISLMTLVTQ FHIPLYRDYL LALGISSVSR KNALRTLSKN QSICIVVGGA RESLLSSTNG 281 TQLILNKRKG FIKLAIQTGN INLVPVFAFG EVDCYNVLST KKDSVLGKMQ LWFKENFGFT IPIFYARGLF 351 NYDFGLLPFR APINVVVGRP IYVEKKITNP PDDVVNHFHD LYIAELKRLY YENREKYGVP DAELKIVG	CC48/49AA C127A C183A C211S C264A C314A F71A L73A Y129A/ F130A/ P131A H193A H195A	-30% -40% -10% -20% -15% -60% -40% -100% -100%	[33,37]
ZmDGAT1-2- EU039830 [Zea mays]	1 MAPPPSMPAA SDRAGPGRDA GDSSSLRLRR APSADAGDLA GDSSGGLREN GEPQSPTNPP PQEQQQHEML 71 YYRASAPAHR RVKESPLSSD AIFRQSHAGL LNLCIVVLIA VNSRLIIENL MKYGLLIRAG FWFSARSLGD 141 WPLLMCCLTL PVFPLVALMA EKLITRKLIG EHVVILLHII ITTSAIVYPV VVTLKCDSAV LSGFVLMFLA 211 SIMWMKLVSY AHTNYDIRVL SKSTEKGAAY GNYVDPENMK DPTFKSLVYF MLAPTLCYQP TYPQTTCIRK 281 GWVTQQLIKC VVFTGLMGFI IEQYINPIVK NSKHPLKGNF LNAIERVLKL SVPTLYVWLC MFYCFFHLWL 351 NIVAELLCFG DREFYKDWWN AKTVEEYWRM WNMPVHKWII RHIYFPCIRK GFSRGVAILI SFLVSAVFHE 421 ICIAVPCHIF KFWAFSGIMF QIPLVFLTRY LHATFKHVMV GNMIFWFFFS IVGQPMCVLL YYHDVMNRQA	F469 insertion	+41-107% oil contents	[38]

**Table 4 Site-directed and natural mutants of DGATs and their effects on enzymatic activity and TAG accumulation.** *(Continued)* 

BtDGAT1-	1 MGDRGGAGGS RRRRTGSRPS IQGGSGPAAA EEEVRDVGAG	K232A	Devoid of milk	[41]
[Bos taurus]	71 SLFSSDSGFS NYRGILNWCV VMLILSNARL FLENLIKYGI		Sceletion	
	LVDPIQVVSL FLKDPYSWPA LCLVIVANIF			
	141 AVAAFQVEKR LAVGALTEQA GLLLHGVNLA TILCFPAAVA			
	FLLESITPVG SVLALMVYTI LFLKLFSYRD			
	211 VNLWCRERRA GAKAKAALAG K <b>K</b> ANGGAAQR TVSYPDNLTY			
	RDLYYFLFAP TLCYELNFPR SPRIRKRFLL			
	281 RRLLEMLFLT QLQVGLIQQW MVPAIQNSMK PFKDMDYSRI			
	VERLLKLAVP NHLIWLIFFY WLFHSCLNAV			
	351 AELMQFGDRE FYRDWWNSES ITYFWQNWNI PVHKWCIRHF			
	YKPMLRRGSS KWAARTAVFL ASAFFHEYLV			
	421 SIPLRMFRLW AFTGMMAQIP LAWIVGRFFR GNYGNAAVWL			
	SLIIGQPVAV LMYVHDYYVL NREAPAAGT			
			,	

(Figure 3A). Mutagenesis at P216 in *Tropaeolum majus* DGAT1 eliminates almost all of the activity [25]. P216 is completely conserved in plant DGAT1s but is missing in all mammalian DGAT1s (Figure 9B). Mutation at Y129/F130/P131 in DGAT2 of baker's yeast results in a complete loss of the activity [33]. These three residues are highly conserved but none of them is completely conserved among all DGAT2s in our analysis using 54 full-length DGATs (Figure 10A). Mutations at F80/L81/ L83 in mouse DGAT2 [34] and F71/L73 in baker's yeast DGAT2 [33] result in partial loss of the DGAT activity (Figure 10B). Finally, ScDGAT2 has a unique cysteine residue (C314) which is not involved in catalysis but may be located near the active site or related to proper folding of the protein [37]. However, this residue is only found in DGAT2s from baker's yeast and the other two fungi Ashbya gossypii and Physcomitrella patens, but is not present in any of the other 51 DGAT2s or any of the 55 DGAT1s analyzed (Figure 10C). Nonetheless, site-directed mutagenesis indicates that these less conserved residues, although not essential, contribute to the full activity of DGATs.

### Functional significance of the relatively conserved residues in natural mutants

Two well-known natural mutants in corn and cattle demonstrate the importance of some relatively conserved residues in TAG biosynthesis (Table 4). Genetic mapping has identified a high-oil QTL (qHO6) that affects maize seed oil and oleic-acid contents associated with DGAT1-2 [38]. A phenylalanine insertion (F469) in DGAT1-2 is responsible for the increased oil and oleic-acid contents. Ectopic expression of the high-oil DGAT1-2 allele increases oil and oleic-acid contents by up to 41% and 107%, respectively [38]. This phenylalanine residue is conserved in all plants except *Brassica napus* (rape, AAD45536.1) and conserved in all fungi except mold (*Dictyostelium discoideum* and *Polysphon-dylium pallidum*) (Figure 11A). It is not present in any of the animal DGAT1s or any of DGAT2s. This case suggests that oil content can be potentially improved in transgenic plants by introducing site-specific amino acid substitutions/changes in DGATs.

DGAT1 knockout mice are completely devoid of milk secretion, most likely because of deficient triglyceride synthesis in the mammary gland [18]. DGAT1 sequences from pooled DNA show significant frequency shifts at several residue positions between groups of animals with high and low breeding values for milk fat content in different breeds [41]. Substitution of lysine by alanine (K232A) is directly responsible for the QTL variation with the lysine-encoding allele being associated with higher milk fat content [41]. Both DGAT1 alleles are expressed in Sf9 cells, an insect expression system, and characterized the expressed proteins. The K allele, causing an increase in milk fat percentage in the live animal, is characterized by a higher Vmax in producing triglycerides than the A allele [46]. This lysine residue is conserved in mammals but not in plants, fungi or other animals except one of the two forms from dog and zebrafish (Figure 11B). This case also suggests that lipid content can be improved in transgenic animals by bioengineering specific amino acid residues of DGATs.

#### Conclusions

Understanding the precise roles of DGATs may help to create transgenic plants with value-added properties and provide information for therapeutic intervention for obesity and related diseases because DGATs catalyze the final and rate-limiting step of TAG biosynthesis in eukaryotic organisms. This report analyzed 117 DGAT sequences from 70 organisms ranging from plants, animals and fungi to aid our understanding of the structure-function relationship of these important enzymes. The report identified conserved sequence motifs and amino acid residues in all 117 DGATs and DGAT1 and DGAT2 subfamilies, reassigned some DGAT subfamily members based on the phylogenetic analysis and

(41 residues) and DGAT2s (16 residues) in the multiple sequences alignment. This sequence divergence is in contrast to the general belief that the active sites of enzymes should be conserved during the evolution because all catalyze the same/similar biochemical



<b>A</b>						
A	В	С				
(168) 168	188 (101) 101 110 121	(361) 361 379				
AcDGAT2-XP_003225477.1(125) VWRYFRD	TFPWRLWKTHNLMT AcDGAT2-XP_003225477.1 (61) LQVLSVLQWVLSFLIMEWACT	AcDGAT2-XP_003225477.1(264) ADLVPIYSFGENEVYEQVI				
BtDGAT2-DAA21853.1(105) VWRYFRD	TEP QT WRTHNT.LT BtDGAT2-DAA21853.1 (41) TQVTSVT QWVT SFTUULEWACS	BtDGAT2-DAA21853.1(244) ADT.VPTYSFGENEVYKQVT				
MmDGAT2-NP_080660.1(132) VWRYFRD	(FP.QLVKTHNLLT MmDGAT2-NP_080660.1 (68) LQVLSVLQWVLSFLVLGVACS	MmDGAT2-NP_080660.1(271) ADLVPTYSFGENEVYKQVI				
RnDGAT2-NP_001012345.1(132) VWRYFRD	(FP:QLVKTHNLLT RnDGAT2-NP_001012345.1 (68) LQVISVLQWVLSFLVLGVACS	RnDGAT2-NP_001012345.1(271) ADLVPIYSFGENEVYKQVI				
HsDGAT2-AAQ88896.1(132) VWRYFRD	(FPLQLWKTHNLLT HsDGAT2-AAQ88896.1 (68) LQVLSVLQWVLSFLVLGWACS	HsDGAT2-AAQ88896.1(271) ADLVPIYSFGENEVYKQVI				
PaDGA12-XP_002822304.1(132) GWRYFRE	TP QLVKTHNLLT PADGAT2-XP_002822304.1 (68) LQVLSVLQWVLSFLVLGVACS	PaDGAI2-XP_002822304.1(2/1) ADLVPIYSFGENEVYKQVI				
XtDGAT2-NP_989372.1(105) VWRYFRD	TEPIKLWKTHNLLP XtDGAT2-NP_989372.1 (41) LQLLSVLQWVLSELLLGVACT	XtDGAT2-NP_989372.1(244) ADLVPVYSFGENBAYKQVV				
DrDGAT2-NP_001025367.1(105) VWK1MRD		INDEAT 2-NP_001025367.1(244) ADEVEVISEGENEVINGEL				
BEDGAT2-VD 975400 3(150) VWTTERD		BEDGAT2-VP 875499 3(289) ACT VP VI ST GENESI ROVI				
HSDGAT2-NP 835470 1 (85) TWROLED	WPWKTWKTAELPP HSDGAT2-NP 835470 1 (21) LEAVGAYOYU TELEMEPERS	HSDGAT2-NP 835470 1(224) ASLVEVISEGENDIERLKA				
PtDGAT2-XP 527842.2 (85) IWROLRD	YPWKLWKTAELPP PtDGAT2-XP 527842.2 (21) LEAWGAYOYVLTFLEMGPFFS	PtDGAT2-XP 527842.2(224) ASLVPVYSFGENDIFRLKA				
DdDGAT2-XP 635762.1 (76) LWRYFRD	CFPISITINSNYDP DdDGAT2-XP 635762.1 (15) LETMAVAIYAMVLPVCLIMAF	DdDGAT2-XP 635762.1(214) ASLVPVYSFGENDIYDOVP				
PpDGAT2-EFA83646.1 (76) TWRYFRD	YFPISLVTSTKLDP PpDGAT2-EFA83646.1 (15) LETTAVLFYLMIVPMCMLFAF	PpDGAT2-EFA83646.1(214) ASLVPIYSFGENDIF				
NvDGAT2-XP_001630435.1 (82) VWKYFRD	VFPIOLVKTTELDP NvDGAT2-XP_001630435.1 (17) LETAAVALYYYSFFFGALLGF	NvDGAT2-XP_001630435.1(221) ASLVPVYAFGENDVFNQVS				
NvDGAT2-XP_001633322.1 (83) MWKYFCE	FP.SLIKTADLDP NvDGAT2-XP_001633322.1 (17) LETHTVILIVHSFIFGHILGC	NvDGAT2-XP_001633322.1(222) ASLVPVFSFGENELYSQVD				
NvDGAT2-XP_001635548.1 (58) TWHFFRD	(FP.KLIRTKRLDP NvDGAT2-XP_001635548.1 (1)MCTIFFLGHTLGT	NvDGAT2-XP_001635548.1(197) ASLVPVYSFGENELYCQMD				
TgDGAT2-XP_002187643.1 (80) VWKYMRD	(FPISLVKTAELDP TgDGAT2-XP_002187643.1 (16) LQTASVAQWIFSFLCLALCCT	TgDGAT2-XP_002187643.1(219) TPLVPAFSFGENDVFDQVK				
MdDGAT2-XP_001365685.1 (80) IWKHFRN	(FP1HLWKTVDLDP MdDGAT2-XP_001365685.1 (16) LQUGAVLQWIFSFIMLAKLWF	MdDGAT2-XP_001365685.1(219) AYLVPVFSFGENDLFYQVN				
CiDGAT2-XP_002120879.1 (80) LWKWMSD	(FPGTLHKTVDLDP CDGAT2-XP_002120879.1 (15) LCTLAVSHFVFCFLALAATCT	CIDGAT2-XP_002120879.1(219) VPLVPVFSFGDHALWEQKP				
CeDGAT2-NP_505413.1 (94) THSWYAN	/FPVKLHTTSDMPE CeDGAT2-NP_505413.1 (30) LOTLGALHFFFTTLFTPVLVL	CeDGAT2-NP_505413.1(233) AQLVPCYSFGENDIFNQAE				
CeDGAT2-NP_872180.1 (84) HSWYAN	TEPEKLETTSDMPE CEDGAT 2-NP_8/2180.1 (20) LOTIGALEFFTTEFTPULVL	TEDCATE VD. 075146 1(223) AQLVPCY SFGENDIFNQAE				
ACCAT2 NR 082542 1(164) THRAVER	ADCAT2-NP_9/5140.1 (19) LOTTARGEWETTLAFGEFIGT					
SCDCAT2-NP_983542.1(104) LWK11CB		SCDGAT2-NP_903542.1(344) VNLVP11AFGE11CCNVLD SCDGAT2-NP_014888 1(301) TNLVPVFAFGEVUCVNVLS				
SpDGAT2-XP_001713160 1 (90) PYRWECH	VEPTRUHKTELDS SDBGAT2-XP_001713160.1 (25) VSHWLOALAVELHOVELTUTA	SpDGAT2-XP_001713160.1(228) SST VPCFAEGES TEEOVD				
CrDGAT2-XP_001694904.1 (86) TROTWRE	CREATE AND A CODE AND A COMPANY AND A COMPAN	CrDGAT2-XP_001694904.1(219) GGTVPVYHEGNSOVLDEGP				
CrDGAT2-XP_001693189.1 (77) SVARAAA	(FPTRVVVTDPEAF CrDGAT2-XP 001693189.1 (10) VLEYAATAIYVSATYTSVVLL	CrDGAT2-XP 001693189.1(214) APL VPVWAFGOTRAMSWFR				
OtDGAT2-XP 003083539.1 (66) ITRTAKS	(FPCALTFENEEAY OtDGAT2-XP_003083539.1 (5) IVDHGVLLVWLGLFHALVVVV	OtDGAT2-XP_003083539.1(215) AALVPAYTFGQTRAYKYWR				
PpDGAT2-XP_001758758.1 (75) ICKYAPA	YFPIKWVFDDENAF PpDGAT2-XP_001758758.1 (14) VLTFLATVVWMGGLHLNVLVV	PpDGAT2-XP_001758758.1(208) SLIVPTFCFGQTNCKKWWK				
PpDGAT2-XP_001777726.1 (71) IVTHAKN	<pre>/FP1KVIFEDESAF PpDGAT2-XP_001777726.1 (10) IWSIIAMVMWLGGFHINFIVG</pre>	PpDGAT2-XP_001777726.1(204) SPLVPTFCFGQRNAYKWWK				
SmDGAT2-XP_002972054.1(102) ICKYAPK	FP_RLHVDDIKSF SmDGAT2-XP_002972054.1 (41) VLSWIAVLIWLGAIHFNTIVG	SmDGAT2-XP_002972054.1(235)				
PsDGAT2-ABK26256.1 (90) ICYHATS	(FPVTLIVEDMKAF PsDGAT2-ABK26256.1 (29) FRSIVALIMALGTIHFDAILV	PsDGAT2-ABK26256.1(223) CPLVPVFCFGQTEAYFWWR				
SpDGAT2-AAQ89590.1 (84) VGRYSPG	(FHWNVHFEDVKAF SpDGAT2-AAQ89590.1 (24) IRSIVATMLWVCPIALTAFLC	SpDGAT2-AAQ89590.1(217) SPLVPVFCFGQNBAFSWWK				
RcDGAT2-XP_002528531.1(102) VCRHAC6	IFPVTLHVEDMNAF RcDGAT2-XP_002528531.1 (41) FHALLALSIWIGSIHFNLFLL	RcDGAT2-XP_002528531.1(235) KPLVPVFCFGQSNVYKWWK				
VgDGAT2-ACV40232.1 (98) VSKYVMG	AFPVTLYVEDMKCF VgDGAT2-ACV40232.1 (3/) PRALIALSLWIGSTHFILFLL	VgDGAT2-ACV40232.1(233) TPLVPVFFFGQAHTYKWWR				
VTDGA12-DG356682 (86) ISRHVCS	(FPITIHVEDINAF VIDGAT2-DG550082 (20) FQSVLALAINLGSFHFILFLV	VIDGAT2-DG356682(217) TPLVPVFCFGQMHTFKWWK				
E0DGAT2-AC035365.1 (89) ICKYAVG	TEPYTLYVEDIKVE EODGAT2-ACO35305.1 (28) WEITVALALWEGEIHENVALV					
OCDGAT2-NP 001057530(102) TCRYAMS		OCDGAT2-NP 001057530(235) CPLVPVFCFGOSVAVKWWR				
OSDGAT2-NP_001037330(102) TEXTAINS	VEDWITHVEDYKAF OSDGAT2-NP 001047917 (39) VETVALALWI GATHENAFLL	OSDGAT2-NP_001047917(233) SPLVPVFAEGOSYMYKWWR				
ShDGAT2-XP_002452652_1_(96)SKYVIG	SbDGAT2-XP 002452652.1 (35) TRUTVALATAL GATHENAFLY	SbDGAT2-XP 002452652.1(229) CPWVPVFAFGOSYWYKWWR				
ZmDGAT2-NP 001150174.1 (95) ISKYVIG	TENTLHVEDYGAF ZmDGAT2-NP 001150174.1 (34) IRTTVALALWIGAIHFNAFLV	ZmDGAT2-NP 001150174.1(228) CPVVPVFAFGOSYVYKWWR				
AtDGAT2-NP_566952 (77) ICKHACN	YFPVSLYVEDYEAF AtDGAT2-NP_566952 (16) FHSIIAMAIWLGAIHFNVALV	AtDGAT2-NP_566952(210) SPLVPVFCFGQARVYKWWK				
BnDGAT2a-ACO90187 (79) ICKHAAS	VFPVTLHVEDYEAF BnDGAT2a-ACO90187 (18) FHAVTAISICLSAIYLNLALV	BnDGAT2a-ACO90187(212) APLVPVFCFGQSRAYKWWK				
BnDGAT2b-ACO90188 (79) ICKHAAS	IFPVTLHVEDYEAF BnDGAT2b-ACO90188 (18) LHAVTAISICLSAIYLNLALV	BnDGAT2b-ACO90188(212) APLVPVFCFGQSRAYKWWK				
VvDGAT2-XP_002263626 (99) ICKHACG	<pre>(FPVTLYVEDIKAF VvDGAT2-XP_002263626 (38) IHSIAALTIWLGAIHFVVALV</pre>	VvDGAT2-XP_002263626(232) RPLVPVFCFGQSRVYKWWK				
EaDGAT2-ADF57328.1 (93) ICMHAFS	YFP.TLHVEDIDAF EaDGAT2-ADF57328.1 (32) MRTVVAMLIWLGAIHFNVALV	EaDGAT2-ADF57328.1(226) TPLVPVECFGQSRVYKWWK				
GmDGAT2-ACU20344.1 (79) ICKHICS	(FPITLHVEEAKAF GmDGAT2-ACU20344.1 (18) FKTVPALVLYIGAIHFNLALI	GmDGAT2-ACU20344.1(212) LPLVPVFCFGQTKAYKWWK				
MtDGAT2-ACJ84867.1 (83) ICKHACS	(FPITIHVEDIKAF MtDGAT2-ACJ84867.1 (22) FRSILALSIALGSIHFNIAII	MtDGAT2-ACJ84867.1(216) HPLVPVFCFGQSDIYKWWK				
HaDGA12-ABU50328.1 (98) ICKHAVG	TEPVILYVEDYKAF HADGATZ-ABUSU328.1 (37) IHIALAMIAWLGSIHLNVFIV	HADGATZ-ABUSU328.1(231) BPLVPVFGFGQSYVYKWWK				
OeDGAT2-ADG22608.1 (97) CKHAVG	(FPWN YVEDIKAF OEDGAT2-ADG22008.1 (30) FHTIALVLWLGSWHFNLIVV	UPDGAT 2-ADG22608.1(230) RELVEVE CEGOTINGENWIK				
UFDGAT2R AAK84179.1 (97) WWKLFAG						
Consensus(168) IWKY	TFPITLV E L Consensus(101) L TI AL WL I V LV	Consensus(361) A LVPVFSFGE EVYK				
Eigure 10 Seguence analysis	of important amino acid residues in less conservative regio	one of DGAT2E (A) VED motif (The bound V C				
and R residues are mutated in	by a minimum animum of the residues in less conservative region $h_{1,20}$ (P) The basis	ved E L and L residues are mutated in moure				
DGAT2 corresponding to E90.1	Jaker's yeast DGATZ corresponding to TTZ9/FTS0/PTST), (B) The DO	173 (C) The boxed C residue is mutated in				
baker's vesst DGAT2 correspond	dia to C314. Multiple sequence alignment was performed using 5	4 full-longth DGAT2 protein sequences from 44				
brancing state borners conceptioning to Contemporate sequence angiment was performed using of numericity DOAT2 conception and with detaile in "Mathede" rection						
The amine acid residues studie	nor code and related information are described priefly in Figure 2	legend and with details in internous section.				
i ine amino acid residues studie	u by mulagenesis and the corresponding conserved residues in oth	her organisms are boxed within the sequence				

The amino acid residues studied alignment.

reaction. Therefore, this sequence divergence raises an important question how proteins with completely different amino acid sequences could perform the same biochemical reaction, although some variations of the conserved sequence motifs and amino acid residues are expected when more sequences of DGATs are used in the multiple sequence alignment.

It has been well-documented that many of the enzymes in the oil biosynthesis pathway are not stable. Although the precise reasons are unknown, it is possible that plants develop a feedback mechanism to regulate the optimal amount of enzymes so that the biophysical properties of ER membranes are functionally intact without dramatic alterations by over-expressed enzymes in the host. If this is the case, it may be advantageous to introduce genes with low copy numbers but with high catalytic efficiency. This concept is supported by three studies with plants (S197 in *Arabidopsis* and F469 in corn) and animals (K232 in cattle) which demonstrate the potential to increase oil/fat production by altering a single amino acid residue of DGAT1. Therefore, the sequence analysis should facilitate studying the structure-function relationship of DGATs with the ultimate goal of identifying critical amino acid residues. This will guide the construction of superb enzymes for metabolic

А				В				
(777)	777 77	90 801	(486)	486 5	00 51	0	520	530
DmDGAT1-NP_609813.1(525)	MGNIIVWAS-IILG	QPLCIMA <mark>Y</mark> YHD	DmDGAT1-NP_609813.1(296)	GVLNGGEEDEDV	'SKLVQ <mark>YI</mark>	DNLTYKDI	LYFLC	A <mark>PT</mark> LCY
DmDGAT1-NP_995724.1(575)	MGNIIVWAS-IILG	<mark>QP</mark> LCIMA <mark>Y</mark> YHD	DmDGAT1-NP_995724.1(346)	GVLNGGEEDEDV	'SKLVQ <mark>YB</mark>	DNLTYKDI	LYFLC	A <mark>PT</mark> LCY
TcDGAT1-XP_975142.1(452)	MGNIVVWAS-LIIG	<mark>QP</mark> LCIMM <mark>Y</mark> YHD	TcDGAT1-XP_975142.1(216)	DEDSNKNHKDSKDDI	SNTERLLVQ <mark>YE</mark>	DNLNLRDI	FYFLC	A <mark>PT</mark> LCY
BbDGAT1-AAZ22403.1(453)	YGNAAVWLS-LIIG	<mark>QP</mark> VAVLM <mark>Y</mark> VHD	BbDGAT1-AAZ22403.1(231)	KKANGG	AAQRTVS <mark>YE</mark>	DNLTYRDI	YYFLF	A <mark>PT</mark> LCY
BtDGAT1-NP_777118.2(453)	YGNAAVWLS-LIIG	<mark>QP</mark> VAVLM <mark>Y</mark> VHD	BtDGAT1-NP_777118.2(231)	KAANGG	AAQRTVS <mark>YE</mark>	DNLTYRDI	YYFLF	A <mark>PT</mark> LCY
OaDGAT1-NP_001103634.1(453)	YGNAAVWLS-LIIG	<mark>QP</mark> VAVLM <mark>Y</mark> VHD	OaDGAT1-NP_001103634.1(231)	KKANGG	AAQRTVS <mark>YE</mark>	DNLTYRDI	YYFLF	A <mark>PT</mark> LCY
SsDGAT1-NP_999216.1(453)	YGNAAVWLS-LIIG	<mark>OP</mark> VAVLM <mark>Y</mark> VHD	SsDGAT1-NP_999216.1(231)	KKANGG	AAQHSVS <mark>YB</mark>	DNLTYRDI	YYFLL	A <mark>PT</mark> LCY
CfDGAT1-XP_849176.1(462)	YG <mark>N</mark> AAVWLT-LIIG	<mark>QP</mark> VA <mark>VLM</mark> YVHD	CfDGAT1-XP_849176.1(240)	KKANGG	TAQRMVS <mark>Y</mark> F	D <mark>N</mark> LTYRD <mark>I</mark>	YYFLF	A <mark>PT</mark> LCY
CfDGAT1-XP_858062.1(469)	YG <mark>N</mark> AAVWLT-LIIG	<mark>QP</mark> VA <mark>VLM</mark> YVHD	CfDGAT1-XP_858062.1(227)	+	AGWALASL	PSLPPADI	YYFLF	A <mark>PT</mark> LCY
HsDGAT1-NP_036211.2(452)	YGNAAVWLS-LIIG	<mark>QP</mark> IAVLM <mark>Y</mark> VHD	HsDGAT1-NP_036211.2(230)	KKASSA	AAPHTVS <mark>YE</mark>	DNLTYRDI	YYFLF	A <mark>PT</mark> LCY
MmDGAT1-XP_001090134.1(455)	YGNAAVWLT-LIIG	<mark>QP</mark> IA <mark>VLM</mark> YVHD	MmDGAT1-XP_001090134.1(233)	KKASSA	AAPHTVS <mark>YE</mark>	DNLTYRDI	YYFLF	A <mark>PT</mark> LCY
PtDGAT1-XP_520014.2(452)	YGNAAVWLS-LIIG	<mark>QP</mark> IAVLM <mark>Y</mark> VHD	PtDGAT1-XP_520014.2(230)	KKASSA	AAPHTVS <mark>YE</mark>	DNLTYRDI	YYFLF	A <mark>PT</mark> LCY
MmDGAT1-NP_034176.1(463)	YGNAAVWVT-LIIG	<mark>QP</mark> VA <mark>VLM</mark> YVHD	MmDGAT1-NP_034176.1(241)	KKVSGA	AAQQAVS <mark>YI</mark>	DNLTYRDI	YYFIF	A <mark>PT</mark> LCY
RnDGAT1-NP_445889.1(463)	YGNAAVWVT-LIIG	<mark>QP</mark> VA <mark>VLM</mark> YVHD	RnDGAT1-NP_445889.1(241)	KKVSGA	AAQNTVS <mark>YE</mark>	DNLTYRDI	YYFIF	A <mark>PT</mark> LCY
OcDGAT1-XP_002724427.1(467)	YGNAAVWLT-LIIG	<mark>QP</mark> VA <mark>VLM</mark> YVHD	OcDGAT1-XP_002724427.1(246)	-KAIGA	AGQCTVS <mark>YE</mark>	DNLTYRDI	YYFLF?	A <mark>PT</mark> LCY
MdDGAT1-XP_001371565.1(595)	YGNAAVWLS-LIIG	<mark>QP</mark> VA <mark>VLM</mark> YVHD	MdDGAT1-XP_001371565.1(374)	-KANGD	VVSGGVK <mark>YF</mark>	ANLTYQDI	YYFLF?	A <mark>PT</mark> LCY
DrDGAT1-NP_001002458.1(468)	YG <mark>N</mark> AAVWLS-LIIG	<mark>QP</mark> IAVLM <mark>Y</mark> VHD	DrDGAT1-NP_001002458.1(247)	-PSTSS	SMQSYVS <mark>YE</mark>	G <mark>N</mark> LSLRDI	Y <mark>YF</mark> VF	A <mark>PT</mark> LCY
DrDGAT1-NP_956024.1(461)	YGNAAVWMS-LIIG	<mark>QP</mark> IA <mark>VL</mark> M <mark>Y</mark> VHD	DrDGAT1-NP_956024.1(240)	-KANGT	AGYTHVT <mark>Y</mark> F	GNLTHRDI	YYFAF?	A <mark>PT</mark> LCY
SkDGAT1-XP_002736160.1(447)	YS <mark>NMMVW</mark> MS-VIA <mark>G</mark>	QPVAIMM <mark>Y</mark> VHD	SkDGAT1-XP_002736160.1(223)	TIPNGK	LSTHLVQ <mark>YB</mark>	DNLNLKDI	YYFLL?	A <mark>PT</mark> LCY
DdDGAT1-XP_645633.2(570)	LG <mark>N</mark> VIFWIS-IVL <mark>G</mark>	<mark>QP</mark> LVVLL <mark>Y</mark> YRN	DdDGAT1-XP_645633.2(355)		TNTSI <mark>YE</mark>	NNLSLRSI	YW <mark>f</mark> ML	/ <mark>PT</mark> LVY
PpDGAT1-EFA85004.1(557)	LG <mark>N</mark> VIFWFS-IVLG	QPMVVLL <mark>Y</mark> YRN	PpDGAT1-EFA85004.1(342)		GDTSV <mark>YE</mark>	NNLSLSNI	YWYML	E <mark>PT</mark> LVY
AcDGAT2-EGC41804.1(476)	IGNCVFWVSFCLVG	<mark>QP</mark> LGA <mark>LLY</mark> FFA	AcDGAT2-EGC41804.1(256)		YHSLP <mark>YE</mark>	RNITIKNI	TYFWL	A <mark>PT</mark> LVY
PbDGAT1-EEH17170.1(469)	IG <mark>N</mark> CVFWVSECLV <mark>G</mark>	<mark>QP</mark> LAA <mark>LL</mark> YFFA	PbDGAT1-EEH17170.1(249)		YQSCP <mark>YE</mark>	RNVTIGNI	TYFWL	A <mark>PT</mark> LVY
AoDGAT2-EEQ31683.1(470)	IG <mark>N</mark> SVFWLSECLV <mark>G</mark>	<mark>QP</mark> LGA <mark>LL</mark> YFFA	AoDGAT2-EEQ31683.1(250)		YKSCP <mark>YE</mark>	HNVTVSNI	SYFWL	A <mark>PT</mark> LVY
MaDGAT2-EFY86774.1(461)	MGNVIFWVSFTIFG	<mark>QP</mark> FAA <mark>L</mark> M <mark>Y</mark> FYA	MaDGAT2-EFY86774.1(221)		YSKCP <mark>YE</mark>	NNITFGNI	AYFWW	A <mark>PT</mark> LVY
MaDGAT2-EFY97444.1(490)	MGNVIFWVSFTIFG	<mark>QP</mark> FAALM <mark>Y</mark> FYA	MaDGAT2-EFY97444.1(250)		YSKCP <mark>YE</mark>	NNITFGNI	AYFWW	A <mark>PT</mark> LVY
TgDGAT1-AAP94209.1(497)	VGNCFFWFIFCFSG	QPLGI <mark>l</mark> I <mark>Y</mark> WYL	TgDGAT1-AAP94209.1(271)		EAEHVRR <mark>YE</mark>	YSITLRHI	YTEIWN	1 <mark>PT</mark> MCF
CvDGAT1-EFN50697.1(391)	VG <mark>NIIFWVSF</mark> CFV <mark>G</mark>	<mark>QP</mark> LAMI <mark>LY</mark> YHD	CvDGAT1-EFN50697.1(173)		EEELR <mark>Y</mark> F	ENIVAANI	AYFLL	A <mark>PT</mark> LCY
PpDGAT1-XP_001770929.1(415)	VGNMVFWFFFCIVG	QPMCLLL <mark>Y</mark> YHD	PpDGAT1-XP_001770929.1(198)		ADKIEV	DHLTIQNI	AYFML?	A <mark>PT</mark> LCY
SmDGAT1-XP_002964165.1(387)	VG <mark>N</mark> MIFWFFFCIV <mark>G</mark>	<mark>QP</mark> MCVLL <mark>Y</mark> YHD	SmDGAT1-XP_002964165.1(171)		P-AID <mark>YI</mark>	DNISLKNI	AYFMA.	A <mark>PT</mark> LCY
OsDGAT1-NP_001054869.2(505)	VG <mark>N</mark> MIFWFFFCIY <mark>G</mark>	QPMCLLL <mark>Y</mark> YHD	OsDGAT1-NP_001054869.2(288)		TVDMDNI	QPPTLGNI	IYFMM	A <mark>PT</mark> LCY
SbDGAT1-XP_002439419.1(482)	VG <mark>N</mark> MIFWFFFCIY <mark>G</mark>	<mark>QP</mark> MCVLL <mark>Y</mark> YHD	SbDGAT1-XP_002439419.1(265)		VADIDNI	QAPTLGS	TYFMM	A <mark>PT</mark> LCY
SbDGAT1-XP_002437165.1(456)	VG <mark>N</mark> MIFWFFFSIV <mark>G</mark>	<mark>QP</mark> MCVLL <mark>Y</mark> YHD	SbDGAT1-XP_002437165.1(240)		VDP-ESM	IKDPTFKS <mark>I</mark>	VYFML?	A <mark>PT</mark> LCY
ZmDGAT1-2-EU039830(460)	VG <mark>N</mark> MIFWFFFSIV <mark>G</mark>	<mark>QP</mark> MCVLL <mark>Y</mark> YHD	ZmDGAT1-2-EU039830(244)		VDP-ENM	IKDPTFKS <mark>I</mark>	VYFML?	A <mark>PT</mark> LCY
VgDGAT1-ABV21945.1(490)	VG <mark>N</mark> MIFWCFFSIF <mark>G</mark>	<mark>QP</mark> MCVLL <mark>Y</mark> YHD	VgDGAT1-ABV21945.1(274)		SN-MD <mark>Y</mark> F	'YDVNFKS <mark>I</mark>	VYFMV2	A <mark>PT</mark> LCY
EpDGAT1-ACO55635.1(440)	VG <mark>N</mark> MIFWCFFCIL <mark>G</mark>	QP <mark>MSLLL</mark> YYHD	EpDGAT1-ACO55635.1(224)		SME-DCS	FEVNFQA	VYFMV2	A <mark>PT</mark> LCY
OeDGAT1-AAS01606.1(499)	VG <mark>N</mark> MIFWCFFSIL <mark>G</mark>	QPMCLLL <mark>Y</mark> YHD	OeDGAT1-AAS01606.1(283)		WNS-DDS	YGASFQS <mark>I</mark>	AYFMV	A <mark>PT</mark> LCY
PfDGAT1-AAG23696.1(501)	VGNMMFWCFFCIFG	<mark>QP</mark> MCVLL <mark>Y</mark> YHD	PfDGAT1-AAG23696.1(285)		WNL-D <mark>Y</mark> A	YDVSFKS <mark>I</mark>	AYFMV	A <mark>PT</mark> LCY
NtDGAT1-AAF19345.1(499)	VGNMMFWCFFCILG	<mark>QP</mark> MCVLL <mark>Y</mark> YHD	NtDGAT1-AAF19345.1(286)		DFS	YDVSFKS <mark>I</mark>	AYFMV	A <mark>PT</mark> LCY
AtDGAT1-NP_179535.1(487)	VGNMIFWFIFCIFG	<mark>QP</mark> MCVLL <mark>Y</mark> YHD	AtDGAT1-NP_179535.1(275)		EVS	YYVSLKS <mark>I</mark>	AYFMV	A <mark>PT</mark> LCY
BjDGAT1-AAY40784.1(470)	VG <mark>N</mark> MIFWFSFCIF <mark>G</mark>	<mark>QP</mark> MCVLL <mark>Y</mark> YHD	BjDGAT1-AAY40784.1(258)		EIS	YYVSLKS	AYFMV	A <mark>PT</mark> LCY
BnDGAT1-AAD45536.1(470)	VG <mark>N</mark> MIFGSASCIF <mark>G</mark>	<mark>QP</mark> MCGLL <mark>Y</mark> YHD	BnDGAT1-AAD45536.1(258)		EIS	YYVSLKS	AYFMV	A <mark>PT</mark> LCY
BjDGAT2-AAY40785.1(470)	VG <mark>N</mark> MIFWFSFCIF <mark>G</mark>	<mark>QP</mark> MCVLL <mark>Y</mark> YHD	BjDGAT2-AAY40785.1(258)		EVS	YYVSLKS	AYFMV	A <mark>PT</mark> LCY
TmDGAT1-AAM03340.2(487)	VGNMIFWFIFCILG	<mark>QP</mark> MCVLL <mark>Y</mark> YHD	TmDGAT1-AAM03340.2(270)		STIVSCS	YDVSLKS	AYFMV	A <mark>PT</mark> LCY
EaDGAT1-AAV31083.1(474)	VG <mark>N</mark> MMFWFSFCIF <mark>G</mark>	QPMCLLL <mark>Y</mark> YHD	EaDGAT1-AAV31083.1(258)		QNMD- <mark>Y</mark> S	FDVNIKS	AYFMV	A <mark>PT</mark> LCY
GmDGAT1-AAS78662.1(465)	VG <mark>N</mark> MIFWFIFSIL <mark>G</mark>	<mark>QP</mark> MCVLL <mark>Y</mark> YHD	GmDGAT1-AAS78662.1(249)		LNMD-YE	YNVSFKSI	AYFLV	A <mark>PT</mark> LCY
GmDGAT1-BAE93461.1(471)	VGNMIFWFIFSILG	<mark>QP</mark> MCVLL <mark>Y</mark> YHD	GmDGAT1-BAE93461.1(255)		LNMD-YE	YNVSFKSI	AYFLV	A <mark>PT</mark> LCY
LjDGAT1-AAW51456.1(476)	VGNMIFWFIFSILG	<mark>QP</mark> MAVLL <mark>Y</mark> YHD	LjDGAT1-AAW51456.1(260)		LNMD- <mark>Y</mark> S	YDVSFKS	AYFMI?	A <mark>PT</mark> LCY
MtDGAT1-ABN09107.1(506)	VGNMIFWFIFCILG	<mark>QP</mark> MCVLL <mark>Y</mark> YHD	MtDGAT1-ABN09107.1(289)		FNMEE	HNVSFQSI	AYFMV	A <mark>PT</mark> LCY
VvDGAT1-XP_002279345.1(482)	VG <mark>N</mark> MIFWLFFSILG	<mark>QP</mark> MCVLL <mark>Y</mark> YHD	VvDGAT1-XP_002279345.1(266)		LNMD-YE	YDVNFKSI	AYFMV	A <mark>PT</mark> LCY
JcDGAT1-ABB84383.1(486)	VGNMIFWFIFCILG	QPMCVLL <mark>Y</mark> YHD	JcDGAT1-ABB84383.1(270)		SGAD-SS	RDVSFKS	VYFMV	A <mark>PT</mark> LCY
VfDGAT1-DQ356680.1(491)	VGNMIFWFIFCILG	QPMCLLLYYHD	VfDGAT1-DQ356680.1(275)		SSAE-SS	HDVSFKS	VYFMV	A <mark>PT</mark> LCY
RcDGAT1-XP_002514132.1(490)	VGNMIFWFFFCILG	QPMCVLL <mark>Y</mark> YHD	RcDGAT1-XP_002514132.1(274)		SSTE-YC	HDVSFKT	AYFMV	A <mark>PT</mark> LCY
PtDGAT1-XP_002308278.1(417)	VG <mark>N</mark> MIFWLFFSILG	<mark>QP</mark> MCVLL <mark>Y</mark> YHD	PtDGAT1-XP_002308278.1(201)		PYVG-N	YDTYFKSI	VYFMV	A <mark>PT</mark> LCY
PtDGAT1-XP_002330510.1(458)	VG <mark>N</mark> MIFWLFFSILG	<mark>QP</mark> MCVLL <mark>Y</mark> YHD	PtDGAT1-XP_002330510.1(242)		SKAD-NS	YDANFKS	VYFMV	A <mark>PT</mark> LCY
o (777)	VONMIEW E T C	ODMOUT T VVHD	Conconcud (196)		VE	NI T	VE 1	DULTON

**Figure 11 Sequence analysis of important amino acid residues of DGAT1s shown in natural mutants**. Multiple sequence alignment was performed using 55 full-length DGAT1 protein sequences from 45 organisms (listed in Table 1). The completely conserved amino acid residues are highlighted in red on yellow. Other color code and related information are described briefly in Figure 2 legend and with details in "Methods" section. The amino acid residues affected by natural mutation and the corresponding conserved residues in other organisms are boxed within the sequence alignment. (A) maize DGAT1-2 F468, (B) cattle DGAT1 K232A.

engineering and rational design of DGAT inhibitors to be used for obesity and related diseases.

#### Methods

### Database search of DGATs

DGAT sequences were obtained from Blastp search [47,48] using tung tree (*Vernicia fordii*) DGAT sequences [GenBank:DQ356680.1] (DGAT1) and [GenBank: DQ356682.1] (DGAT2) [13] against the National Center for Biotechnology Information (NCBI)'s non-redundant

protein sequence databases http://blast.ncbi.nlm.nih.gov/ Blast.cgi. Additional DGAT sequences were obtained using DGAT1 and DGAT2 search term in NCBI's Protein database. http://www.ncbi.nlm.nih.gov/protein. A total of 109 full-length and 8 near full-length DGATs were obtained from 70 organisms including plants (such as *Arabidopsis*, barley, caster bean, cauliflower, corn, rape, rice, sorghum, soybean, tobacco, tung tree), animals (such as bird, chimpanzee, cow, dog, fish, fly, frog, monkey, mosquito, mouse, pig, rabbit, rat, sheep, worm), fungi (such as yeast, mold, moss) and human. The names of DGATs used in the analysis and their corresponding organisms, classification of DGAT subfamily, and the GenBank accession numbers are presented in Table 1. The name of each protein sequence consists of the initials of the organism followed by the assigned subfamily of DGATs in the databases and the GenBank accession number.

#### **Protein analysis**

The properties and amino acid compositions of DGATs were analyzed using Vector NTI software (Invitrogen) [49]. Statistics was performed using Microsoft Excel.

#### **Phylogenetic analysis**

Phylogenetic analysis was used to study the presumed evolutionary relationships among the 117 DGATs from 70 organisms. This analysis was performed using the Vector NTI software (Invitrogen) based on the Neighbor-Joining method of Saitou and Nei [50]. The numbers in the parenthesis following DGAT names are the calculated distance values which reflect the degree of divergence between all pairs of DGAT sequences analyzed.

#### Multiple sequence alignment

Multiple sequence alignment was performed using the ClustalW algorithm [51,52] of the AlignX program of the Vector NTI software. This method is based on algorithms that assign scores to aligned residues and detect sequence similarities. Identical amino acid residues in alignment have higher scores than those not identical and less similar residues. Each DGAT sequence name is on the left of the alignment followed by the position of amino acid residue of DGAT protein sequence in the alignment. The numbers at the top of the alignment are the positions of the multiple sequence alignment. The letters at the bottom of the alignment are the consensus residues. Color codes for amino acid residues are as follows: 1) red on yellow: consensus residue derived from a completely conserved residue at a given position; 2) black on green: consensus residue derived from the occurrence of greater than 50% of a single residue at a given position; 3) blue on cyan: consensus residue derived from a block of similar residues at a given position; 4) green on white: residue weakly similar to consensus residue at a given position; 5) black on white: non-similar residues.

#### Lists of abbreviations

DGAT: diacylglycerol acyltransferase; QTL: quantitative trait loci; TAG: triacylglycerol

#### Acknowledgements

This work was supported by USDA-ARS Quality and Utilization of Agricultural Products Research Program. The author thanks Drs. Chris P. Mattison, Beiguan Mou and Dunhua Zhang for helpful comments on the manuscript.

#### Authors' contributions

HC carried out all aspects of the study including literature review, study design, database search, protein analysis, phylogenetic analysis and multiple sequence alignment. HC also wrote the draft, revised the manuscript after internal review, obtained submission permission from USDA-ARS and approved the final manuscript.

#### **Competing interests**

The author declares that they have no competing interests.

Received: 17 May 2011 Accepted: 21 July 2011 Published: 21 July 2011

#### References

- Cao H, Tuttle JS, Blackshear PJ: Immunological characterization of tristetraprolin as a low abundance, inducible, stable cytosolic protein. J Biol Chem 2004, 279:21489-21499.
- O'Quin JB, Bourassa L, Zhang D, Shockey JM, Gidda SK, Fosnot S, Chapman KD, Mullen RT, Dyer JM: Temperature-sensitive posttranslational regulation of plant omega-3 fatty-acid desaturases is mediated by the endoplasmic reticulum-associated degradation pathway. J Biol Chem 2010, 285:21781-21796.
- Cao H, Deterding LJ, Venable JD, Kennington EA, Yates JR, Tomer KB, Blackshear PJ: Identification of the anti-inflammatory protein tristetraprolin as a hyperphosphorylated protein by mass spectrometry and site-directed mutagenesis. *Biochem J* 2006, 394:285-297.
- Cao H, Deterding LJ, Blackshear PJ: Phosphorylation site analysis of the anti-inflammatory and mRNA-destabilizing protein tristetraprolin. Expert Rev Proteomics 2007, 4:711-726.
- Shannon JC, Pien FM, Cao H, Liu KC: Brittle-1, an adenylate translocator, facilitates transfer of extraplastidial synthesized ADP– glucose into amyloplasts of maize endosperms. *Plant Physiol* 1998, 117:1235-1252.
- Cao H, Imparl-Radosevich J, Guan H, Keeling PL, James MG, Myers AM: Identification of the soluble starch synthase activities of maize endosperm. *Plant Physiol* 1999, 120:205-216.
- Cao H, Preiss J: Evidence for essential arginine residues at the active sites of maize branching enzymes. J Protein Chem 1996, 15:291-304.
- Cao H, Preiss J: Site-directed mutagenesis evidence for arginine-384 residue at the active site of maize branching enzyme II. J Protein Chem 1999, 18:379-386.
- Beatty MK, Rahman A, Cao H, Woodman W, Lee M, Myers AM, James MG: Purification and molecular genetic characterization of ZPU1, a pullulanase-type starch-debranching enzyme from maize. *Plant Physiol* 1999, 119:255-266.
- Cases S, Smith SJ, Zheng YW, Myers HM, Lear SR, Sande E, Novak S, Collins C, Welch CB, Lusis AJ, Erickson SK, Farese RV Jr: Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis. Proc Natl Acad Sci USA 1998, 95:13018-13023.
- Cases S, Stone SJ, Zhou P, Yen E, Tow B, Lardizabal KD, Voelker T, Farese RV Jr: Cloning of DGAT2, a second mammalian diacylglycerol acyltransferase, and related family members. J Biol Chem 2001, 276:38870-38876.
- Lardizabal KD, Mai JT, Wagner NW, Wyrick A, Voelker T, Hawkins DJ: DGAT2 is a new diacylglycerol acyltransferase gene family: purification, cloning, and expression in insect cells of two polypeptides from Mortierella ramanniana with diacylglycerol acyltransferase activity. J Biol Chem 2001, 276:38862-38869.
- Shockey JM, Gidda SK, Chapital DC, Kuan JC, Dhanoa PK, Bland JM, Rothstein SJ, Mullen RT, Dyer JM: Tung tree DGAT1 and DGAT2 have nonredundant functions in triacylglycerol biosynthesis and are localized to different subdomains of the endoplasmic reticulum. *Plant Cell* 2006, 18:2294-2313.
- Durrett TP, McClosky DD, Tumaney AW, Elzinga DA, Ohlrogge J, Pollard M: A distinct DGAT with sn-3 acetyltransferase activity that synthesizes unusual, reduced-viscosity oils in Euonymus and transgenic seeds. Proc Natl Acad Sci USA 2010, 107:9464-9469.
- Saha S, Enugutti B, Rajakumari S, Rajasekharan R: Cytosolic triacylglycerol biosynthetic pathway in oilseeds. Molecular cloning and expression of peanut cytosolic diacylglycerol acyltransferase. *Plant Physiol* 2006, 141:1533-1543.

- Rani SH, Krishna TH, Saha S, Negi AS, Rajasekharan R: Defective in cuticular ridges (DCR) of Arabidopsis thaliana, a gene associated with surface cutin formation, encodes a soluble diacylglycerol acyltransferase. J Biol Chem 2010, 285:38337-38347.
- Zou J, Wei Y, Jako C, Kumar A, Selvaraj G, Taylor DC: The Arabidopsis thaliana TAG1 mutant has a mutation in a diacylglycerol acyltransferase gene. *Plant J* 1999, 19:645-653.
- Smith SJ, Cases S, Jensen DR, Chen HC, Sande E, Tow B, Sanan DA, Raber J, Eckel RH, Farese RV Jr: Obesity resistance and multiple mechanisms of triglyceride synthesis in mice lacking Dgat. Nat Genet 2000, 25:87-90.
- Stone SJ, Myers HM, Watkins SM, Brown BE, Feingold KR, Elias PM, Farese RV Jr: Lipopenia and skin barrier abnormalities in DGAT2-deficient mice. J Biol Chem 2004, 279:11767-11776.
- Chen HC, Rao M, Sajan MP, Standaert M, Kanoh Y, Miura A, Farese RV, Farese RV: Role of adipocyte-derived factors in enhancing insulin signaling in skeletal muscle and white adipose tissue of mice lacking Acyl CoA:diacylglycerol acyltransferase 1. *Diabetes* 2004, 53:1445-1451.
- Andrianov V, Borisjuk N, Pogrebnyak N, Brinker A, Dixon J, Spitsin S, Flynn J, Matyszczuk P, Andryszak K, Laurelli M, Golovkin M, Koprowski H: 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* 2010, 8:277-287.
- Burgal J, Shockey J, Lu C, Dyer J, Larson T, Graham I, Browse J: Metabolic engineering of hydroxy fatty acid production in plants: RcDGAT2 drives dramatic increases in ricinoleate levels in seed oil. *Plant Biotechnol J* 2008, 6:819-831.
- Jako C, Kumar A, Wei Y, Zou J, Barton DL, Giblin EM, Covello PS, Taylor DC: Seed-specific over-expression of an Arabidopsis cDNA encoding a diacylglycerol acyltransferase enhances seed oil content and seed weight. *Plant Physiol* 2001, 126:861-874.
- Lardizabal K, Effertz R, Levering C, Mai J, Pedroso MC, Jury T, Aasen E, Gruys K, Bennett K: Expression of Umbelopsis ramanniana DGAT2A in seed increases oil in soybean. *Plant Physiol* 2008, 148:89-96.
- Xu J, Francis T, Mietkiewska E, Giblin EM, Barton DL, Zhang Y, Zhang M, Taylor DC: 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 sitedirected mutagenesis to modify enzyme activity and oil content. *Plant Biotechnol J* 2008, 6:799-818.
- Bouvier-Nave P, Benveniste P, Oelkers P, Sturley SL, Schaller H: Expression in yeast and tobacco of plant cDNAs encoding acyl CoA:diacylglycerol acyltransferase. *Eur J Biochem* 2000, 267:85-96.
- Liu L, Zhang Y, Chen N, Shi X, Tsang B, Yu YH: Upregulation of myocellular DGAT1 augments triglyceride synthesis in skeletal muscle and protects against fat-induced insulin resistance. J Clin Invest 2007, 117:1679-1689.
- Roorda BD, Hesselink MK, Schaart G, Moonen-Kornips E, Martinez-Martinez P, Losen M, De Baets MH, Mensink RP, Schrauwen P: DGAT1 overexpression in muscle by in vivo DNA electroporation increases intramyocellular lipid content. J Lipid Res 2005, 46:230-236.
- Kamisaka Y, Kimura K, Uemura H, Shibakami M: Activation of diacylglycerol acyltransferase expressed in Saccharomyces cerevisiae: overexpression of Dga1p lacking the N-terminal region in the Deltasnf2 disruptant produces a significant increase in its enzyme activity. *Appl Microbiol Biotechnol* 2010, 88:105-115.
- Liu L, Shi X, Bharadwaj KG, Ikeda S, Yamashita H, Yagyu H, Schaffer JE, Yu YH, Goldberg IJ: DGAT1 expression increases heart triglyceride content but ameliorates lipotoxicity. J Biol Chem 2009, 284:36312-36323.
- Kamisaka Y, Tomita N, Kimura K, Kainou K, Uemura H: DGA1 (diacylglycerol acyltransferase gene) overexpression and leucine biosynthesis significantly increase lipid accumulation in the Deltasnf2 disruptant of Saccharomyces cerevisiae. *Biochem J* 2007, 408:61-68.
- Chen HC, Smith SJ, Ladha Z, Jensen DR, Ferreira LD, Pulawa LK, McGuire JG, Pitas RE, Eckel RH, Farese RV Jr: Increased insulin and leptin sensitivity in mice lacking acyl CoA:diacylglycerol acyltransferase 1. J Clin Invest 2002, 109:1049-1055.
- Liu Q, Siloto RM, Snyder CL, Weselake RJ: Functional and Topological Analysis of Yeast Acyl-CoA:Diacylglycerol Acyltransferase 2, an Endoplasmic Reticulum Enzyme Essential for Triacylglycerol Biosynthesis. J Biol Chem 2011, 286:13115-13126.

- Stone SJ, Levin MC, Farese RV Jr: Membrane topology and identification of key functional amino acid residues of murine acyl-CoA:diacylglycerol acyltransferase-2. J Biol Chem 2006, 281:40273-40282.
- Stone SJ, Levin MC, Zhou P, Han J, Walther TC, Farese RV Jr: The endoplasmic reticulum enzyme DGAT2 is found in mitochondriaassociated membranes and has a mitochondrial targeting signal that promotes its association with mitochondria. J Biol Chem 2009, 284:5352-5361.
- Alam M, Gilham D, Vance DE, Lehner R: Mutation of F417 but not of L418 or L420 in the lipid binding domain decreases the activity of triacylglycerol hydrolase. J Lipid Res 2006, 47:375-383.
- Liu Q, Siloto RM, Weselake RJ: Role of cysteine residues in thiol modification of acyl-CoA:diacylglycerol acyltransferase 2 from yeast. *Biochemistry* 2010, 49:3237-3245.
- Zheng P, Állen WB, Roesler K, Williams ME, Zhang S, Li J, Glassman K, Ranch J, Nubel D, Solawetz W, Bhattramakki D, Llaca V, Deschamps S, Zhong GY, Tarczynski MC, Shen B: A phenylalanine in DGAT is a key determinant of oil content and composition in maize. *Nat Genet* 2008, 40:367-372.
- Nykiforuk CL, Furukawa-Stoffer TL, Huff PW, Sarna M, Laroche A, Moloney MM, Weselake RJ: Characterization of cDNAs encoding diacylglycerol acyltransferase from cultures of Brassica napus and sucrose-mediated induction of enzyme biosynthesis. *Biochim Biophys Acta* 2002, 1580:95-109.
- Cao H, Shannon JC: BT1, a protein critical for in vivo starch accumulation in maize endosperm, is not detected in maize endosperm suspension cultures. *Physiol Plant* 1996, 97:665-673.
- 41. Winter A, Kramer W, Werner FA, Kollers S, Kata S, Durstewitz G, Buitkamp J, Womack JE, Thaller G, Fries R: Association of a lysine-232/alanine polymorphism in a bovine gene encoding acyl-CoA:diacylglycerol acyltransferase (DGAT1) with variation at a quantitative trait locus for milk fat content. Proc Natl Acad Sci USA 2002, 99:9300-9305.
- 42. Kuerschner L, Moessinger C, Thiele C: Imaging of lipid biosynthesis: how a neutral lipid enters lipid droplets. *Traffic* 2008, **9**:338-352.
- Cao H, Chapital DC, Shockey JM, Klasson KT: Expression of tung tree diacylglycerol acyltransferase 1 in *E. coli. BMC Biotechnology* 2011, 11:73.
- Cao H, Chapital DC, Howard OD, Jiang XN, Shockey JM, Klasson KT: Purification of recombinant tung tree diacylglycerol acyltransferases from E. coli. FASEB J 2011, 25, 765.8.
- McFie PJ, Stone SL, Banman SL, Stone SJ: 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 2010, 285:37377-37387.
- 46. Grisart B, Farnir F, Karim L, Cambisano N, Kim JJ, Kvasz A, Mni M, Simon P, Frere JM, Coppieters W, Georges M: Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition. *Proc Natl Acad Sci USA* 2004, 101:2398-2403.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: Basic local alignment search tool. J Mol Biol 1990, 215:403-410.
- Altschul SF, Lipman DJ: Protein database searches for multiple alignments. Proc Natl Acad Sci USA 1990, 87:5509-5513.
- Zimmerman JM, Eliezer N, Simha R: The characterization of amino acid sequences in proteins by statistical methods. J Theor Biol 1968, 21:170-201.
- 50. Saitou N, Nei M: The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 1987, 4:406-425.
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG: Clustal W and Clustal × version 2.0. *Bioinformatics* 2007, 23:2947-2948.
- Thompson JD, Higgins DG, Gibson TJ: CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994, 22:4673-4680.

#### doi:10.1186/1756-0500-4-249

Cite this article as: Cao: Structure-Function Analysis of Diacylglycerol Acyltransferase Sequences from 70 Organisms. *BMC Research Notes* 2011 4:249.