

Cloning and characterization of diacylglycerol acyltransferase (DGAT) cDNA sequence from *Brassica juncea* cv. Pusa Bold

N Ilaiyaraja, A P Rajarani and I M Santha*

Division of Biochemistry, Indian Agricultural Research Institute, New Delhi-110012

Received 30 May 2007; revised 3 January 2008

Diacylglycerol acyltransferase (DGAT: EC 2.3.1.20) is the only enzyme in the Kennedy pathway that is exclusively committed to the synthesis of storage oil in plants. In this study, cloning and characterization of *DGAT* gene sequence from *Brassica juncea* cv Pusa bold, an important oil seed crop of India is reported. A partial gene sequence of 2003 bp was PCR amplified and cloned from *B. juncea*. Sequence analysis showed that it has 10 exonic and 9 intronic sequences in the partial gene. Two cDNA sequences namely *BjDGAT* 1 and *BjDGAT* 2 (1.5 kb) encoding DGAT enzymes were amplified by RT-PCR from the developing seeds. The complete length of these two cDNAs as determined by RACE technique was 1768 bp, including 5' and 3'-UTR. Comparative analysis of the sequences showed that *BjDGAT*1 was 85.1% and 96% identical to *BjDGAT*2 across 1512 coding region and 503 overlapping deduced amino acids respectively. These proteins were alkaline in nature (pI, 8.5-8.6), having similar molecular size (56-57 kD), an N-terminal hydrophilic segment and 9 transmembrane segments. Diacylglycerol/phorbol ester-binding motif (HKWXXRHXYXP) and acyl CoA binding motif (FYXDWWN) required for binding of substrates remained conserved in these proteins. Expression of the two transcripts of DGAT and their role in oil biosynthesis can further be studied.

Keywords: Diacylglycerol acyltransferase (DGAT), Kennedy pathway, Triacylglycerol, RT-PCR, cDNA cloning.

Metabolic pathway engineering in oil seed crops is burgeoning and promising technique to obtain a desirable oil quality and yield for human consumption as well as industrial uses¹. Fatty acid biosynthesis and assembly into triacylglycerol (TAG) are highly regulated at the biochemical level². Thus, cloning and over-expression of the respective regulatory enzymes in this pathway is of major importance for the genetic manipulation. In plants, fatty acid biosynthesis has been found localized in the plastid and exported to the endoplasmic reticulum for synthesis of TAG through the enzymes of Kennedy pathway. Three acyltransferases — glycerol-3-phosphate acyltransferase (GPAT), lysophosphatidic acid acyltransferase (LPAT) and diacylglycerol acyltransferase (DGAT, EC 2.3.1.20) are involved in the storage lipid bioassembly and catalyze the stepwise acylation of glycerol backbone with the final step being the acylation of *sn*-1,2-diacylglycerol (DAG) by DGAT to form TAG. DGAT is the only enzyme that is exclusively committed to TAG biosynthesis and is less specific on acyl CoAs. The *in vivo* pool size of acyl CoA determines the acyl specificity in the *sn*-3

position of the TAG². DGAT also plays a major role in senescence by sequestering fatty acid from galactolipid into TAG³.

Two gene families of DGAT — DGAT 1 and 2 are reported to be mainly responsible for TAG synthesis in plants, fungi and mammals⁴. DGAT1 gene family has high sequence similarity with acylCoA cholesterol acyltransferase (ACAT), while *DGAT*2 family, identified first in the oleogenous fungus *Mortierella ramaniana* shows no sequence similarity to the members of *DGAT*1 gene family. Other TAG producing enzymes like bifunctional DGAT/wax synthase⁵ and acyl CoA independent phospholipid: diacylglycerol acyltransferase (PDAT) have also contribute to TAG synthesis significantly^{6,7}.

The cloning and functional expression of a *DGAT* gene in plants was first reported in *Arabidopsis thaliana*⁸. Seed-specific over-expression of *DGAT*1 cDNA in *Arabidopsis* enhances the oil deposition and its activity was increased by 10-70% in the seeds of transgenic crops, showing the potential role of regulatory enzyme in altering the oil content⁹. In an another study, silencing of *DGAT*1 in tobacco has been found to cause a reduction in seed oil content¹⁰. The cDNA of DGAT has been isolated and characterized from several plant species including

*Corresponding author
E-mail: ims_bio@yahoo.com
Tel.: 011-25841649

Brassica napus^{11,12}, *Ricinus communis*^{13,14}, *Euonymus alatus*¹⁵, *Olea europaea*¹⁶, *Nicotiana tabacum*¹⁷. However, no such reports are available in *B. juncea* (Mustard), an important oil seed crop family *Brassicaceae*. Here, we report, cloning and characterization of a partial *DGAT* genomic sequence and its transcripts from *B. juncea* cv Pusa bold which is a common cultivar grown in India.

Material and Methods

Seeds of *Brassica juncea* cv. Pusa Bold used in the investigation were obtained from the farm of Indian Agricultural Research Institute, New Delhi.

Isolation of genomic DNA and PCR

DNA was isolated from 7 days old etiolated seedlings using CTAB extraction method¹⁸. The nucleotide sequences of *DGAT* genes available in the GenBank database for plants viz cDNA sequences of *Brassica napus* (*DGAT1*), *Arabidopsis thaliana* (*DGAT1*) and *Nicotiana tabacum* were used to identify the conserved sequence in the gene with the help of CLUSTAL W multiple alignment tool (<http://www.genome.jp>). Based on this, a set of gene-specific primers were designed taking care that no intron interrupted the primers after observing their location on the *A. thaliana DGAT*. The primers were custom synthesized from Bangalore Genei. The sequences of the forward and reverse primers were: *DGATf1*: 5'-atggcgattttggattct-3', T_m 53°C, and *DGATr1*: 5'-accaatctttgtagaattc-3', T_m 51°C respectively.

PCR was carried out with a reaction mixture containing 50 ng template DNA, 2.5 µl of 10X PCR buffer, 2.5 µl 15 mM MgCl₂, 1.0 µl 2mM dNTP, 1 µl 100 pico moles of each primers, 2 units of enzyme *Taq* DNA polymerase and deionized water to make up to 25 µl. Partial amplification of *DGAT* gene by PCR reaction was carried out in PTC-150 Minicycler™ (MJ Research, USA), following the conditions: initial denaturation at 94°C for 4 min, followed by 35 cycles of denaturation at 94°C for 1 min, annealing at 51°C for 1 min and extension at 72°C for 2 min and a final extension at 72°C for 10 min after the completion of the cycles. An aliquot of 5 µl of the reaction mixture was run on a 1.0% (w/v) agarose gel containing 1 µg/ml ethidium bromide.

Southern analysis of genomic DNA

Ten µg of cesium chloride-purified plant genomic DNA was digested overnight to completion with *Sal* I, *Bam*HI, *Pst*I, *Eco*RI and *Hind*III. It was

electrophoresed on 0.8% agarose gel and then transferred on to a nylon membrane. The membrane was hybridized with [α -³²P] dCTP labeled PCR product obtained from genomic DNA amplification (Hexalabel™ DNA labeling kit, MBI Fermentas) as described¹⁹.

RNA isolation and RT-PCR

Total RNA from developing seeds (mid-stage) was isolated using TRI Reagent® (supplied by Molecular Research Center, USA) as the maximum activity of *DGAT* enzyme was reported to be at this stage²⁰. A set of gene specific primers (forward primer *DGATf* 1 and reverse primer *DGATr*2: 5'-aggacatggatggttgcg-3') were designed from conserved region of cDNA from *B. napus* and *A. thaliana* to amplify complete coding region of cDNA from *B. juncea*. First strand cDNA was prepared from 2 µg of total RNA using oligo-dT primers by M-MLV reverse transcriptase at 42°C for 1 h. As a result of this, all the mRNA was reverse transcribed into corresponding cDNAs and PCR was carried out using RETRiScript Kit (Ambion) at T_m , 52°C using gene-specific primers according to the manufacturer's protocol.

Rapid amplification of cDNA ends

5' and 3' ends of cDNA were obtained by RACE technique using Ambion's FirstChoice™ RLM-RACE kit using total RNA isolated from developing seeds (mid-stage) following the kit protocol. A forward primer (5'RACE: outer primer, 5'-gctgatggcgat-gaatgaactg-3') and a reverse primer, gene-specific (5'-gtttgaagatagcgtcgga-3') were used to amplify the 5'-untranslated region on the cDNA. 3'-RACE was carried out to clone the 3'-untranslated region using gene specific primer 5'-cgcaaaggatccatgtct-3' as forward and 3'-outer RACE primer 5'-gcgagcaca-gaattacgact-3' provided in the kit as reverse primer.

Cloning of PCR products

The amplified fragments were cloned in pGEMT-Easy T/A cloning vector (Promega, USA) and transformed into *E. coli* DH5 α and recombinants were selected by blue/white screening. Plasmids were isolated by alkaline lysis method.

Sequencing of plasmid clones

Sequencing of double-stranded plasmid clones were done by Automatic sequencer using the facility available at University of Delhi, South campus.

Results and Discussion

A partial genomic sequence of *DGAT* gene of about 2 kb was PCR amplified from *B. juncea* (Fig. 1). This amplified fragment was cloned in pGEMT-Easy T/A cloning vector and transformed into *E. coli* DH5 α . Restriction of recombinant plasmids with *NotI* enzyme confirmed the presence of ~2 kb insert in agarose gel electrophoresis. One of the recombinant clones was designated as p*DGAT* and was sequenced completely. Homology search²⁰ revealed p*DGAT* insert exhibited maximum sequence identity (97%) with *A. thaliana* *DGAT1*, *B. napus* *DGAT1* and *DGAT 2* genes and also from various other plant species like *Glycine max* (85%), *Olea europaea* (83%), *N. tabacum* (84%), *Ricinus communis* (84%) and *Perilla frutescens* (84%).

The nucleotide composition of partial gene obtained by sequencing showed 417 G, 483 A, 714 T and 389 C with 40.24% G + C content. As *B. napus* is a close relative of *B. juncea* cDNA of *BnDGAT1* (1512 bp) was used to find out the exonic and intronic regions of partial gene by pair-wise alignment. Results revealed that it contained 10 exonic (1138 bp) and 9 intronic sequences (865 bp) with precise GTX-XAG splicing boundaries, as expected for eukaryotic RNA polymerase II transcribed genes. It encoded for 379 amino acids from the start codon ATG. In database, *DGAT* genomic gene sequence is available only for *A. thaliana* (gb: Nm 127503), which has 16 exons interrupted by 15 introns in 3551 bp coding region. In comparison to this, *DGAT* of *B. juncea* matched towards the 5' end. Restriction sites of this sequence were analyzed using NEB CUTTER (<http://tools.neb.com/NEBcutter2>) and showed two

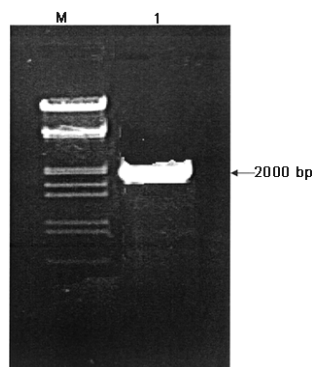


Fig. 1—PCR amplification product (lane 1) of *B. juncea* genomic DNA using gene specific primers designed for *DGAT* gene run on 1% agarose gel along with λ DNA cut with *EcoRI*+ *HindIII* marker (lane M)

sites each for *EcoRI*, *Hind III* and *Sac I*, along with sites for other uncommon enzymes. This partial genomic sequence from *B. juncea* was submitted to the NCBI nucleotide sequence database and has the accession number DQ016107.

Southern analysis of *B. juncea* genomic DNA was carried out to know the probable copy number. Genomic DNA was digested with *Sal I*, *BamH I*, and *Pst I*, owing to the fact that neither of them had a restriction site within the probe used. For comparison, DNA was also digested with *EcoRI* and *HindIII*, as their sites were present in the probe. When the specific probe was used, it revealed two and three bands in the *Sal I* and *EcoRI* digestion respectively (Fig. 2a, b). This suggested that the presence of more than one hybridizing fragments might be due to tetraploid nature of *B. juncea* genome (AABB), which derived from the hybridization of *B. campestris* (AA) and *B. nigra* (BB).

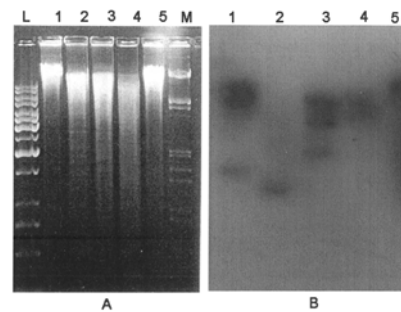


Fig. 2—(A): Restriction pattern of genomic DNA in *Brassica juncea*; and (B): Southern hybridization of restricted genomic DNA with p*DGAT* probe digested with *SalI* (lane 1), *BamHI* (lane 2), *EcoRI* (lane 3) *HindIII* (lane4) and *PstI* (lane 5) and run on 0.7% agarose gel

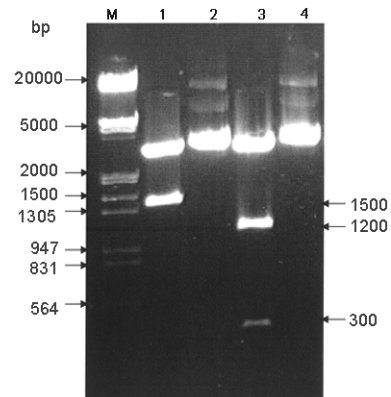


Fig. 3—Restriction pattern of the *BjDGAT1* (lane 1) and *BjDGAT2* (lane 3) clones cut with *EcoRI* enzyme run on EtBr stained 1% agarose gel along with respective uncut plasmids (lanes 2 and 4)

DGAT1 gaaaactttaagattccttagattcactaacaacccatccctctgcatgcggttcggttctttaaaccctaactctttgcaa 80
DGAT2 gaaaactttaagattccttagattcactaacaacccatccctctgcatgcggttcggttctttaaaccctaactctttgcaa
5'UTR
DGAT1 atggcgattttggattctggaggcgctgctgtaccgccgacggagaacggcgctcgcgatctcgacaggctccaccgtcg 160
M A I L D S G G V A V P P T E N G V A D L D R L H R R
DGAT2 atggcgattttggattctggaactgtcacgatggcgacggagaacgggtgctcgcgatctcgatgcttcgctcgctgtaa
M A I L D S G T V T M A T E N G V A D L D M L R R R K
DGAT1 taaatcgagatcggattcttccaacggactcctctccgatacttccccgctcgacgatggtggagctcgccggccgaaa 240
K S R S D S S N G L L S D T S P S D D V G A A A A E
DGAT2 atcgagatcggattcttccaacggacttctctccgagacttccccatcggatgatgctggagctccggccgacgtggagg
S R S D S S N G L L S E T S V P A H R R T R E S G G G
DGAT1 gggatcgggttattccgctgcccaggaggaggtcagggaaacagcgaatttagctggcgagatgccgaaactagggaa 320
R D R V D S A A E E E A Q G T A N L A G G D A E T R E
DGAT2 atcgggttattcagctgctcagggaaacagcgaatttagctggagatcgggaaactagggaaatcgggtggaggaggagga
D R V D S A A Q G T A N L A G D T E T R E S G G G G
DGAT1 tccgccggaggcgatgtaaggtttacgtatcgaccgtcggttccagctcatcggaggacgaggagagctcctctcagctc 400
S A G G D V R F T Y R P S V P A H R R T R E S P L S S
DGAT2 ggaggaacggcgaggtaaggtttacgtatcgaccgtcggttccagctcatcggaggacgaggagagctcactcagctc
G Y R E V R F T Y R P S V P A H R R T R E S G G G
DGAT1 tgacgctatcttcaaaacaaagccatgcaggattgttcaacctctgtgtagttgttctgttctgttgaacagtagacca 480
D A I F K Q S H A G L F N L C V V V L V A V N S R P
DGAT2 cgacgctatcttcaaaacaaagccatgcaggattgttcaacctctgtgtagttgttctgttctgttgaacagtagacca
D A I F K Q S H A G L F N L C V V V L V A V N S R L
DGAT1 tcacgaaaacctcatgaagtatggttgggtgatcagaactgattttgggttagttctacatccttaccgagactggccg 560
I I E N L M K Y G W L I R T D F W F S S T S L R D W P
DGAT2 tcacgaaaacctcatgaagtacgggttgggtgatcagaactgatttctgggttagttctacatccttccgagattggccg
I I E N L M K Y G W L I R T D F W F S S T S L R D W P
DGAT1 cttttcatgtgttctt 640
L F M C C L S L S V F P L A A F T V E K M V L Q K F I
DGAT2 cttttcatgtgttctt
L F M C C L S L S I F T L A A F T V E K L V L Q K F I
DGAT1 atctgagcctgttgcacatcattcttcatgtcattataaccttgacagaggtcttctgtatccagctcactgcactgaggt 720
S E P V A I I L H V I I T L T E V L Y P V Y V T L R
DGAT2 atctgaaacctgttgcacatcattcttcatattattaccactgactgaggtcttctgtatccagctcactgcaccctaaagt
S E P V I I L H I I I T M T E V L Y P V Y V T L R
DGAT1 gtgattctgccttctctgctcaggtgtcaggttgatgctgctcacttgcattgtgtggctgaagtgggttcttaccgctcat 800
C D S A F L S G V T L M L L T C I V W L K L V S Y A H
DGAT2 gtgattcggcttcttcttaccggtgtcacttgcattgtgtggctgaagtgggttcttaccgctcat
C D S A F L S G V T L M L L T C I V W L K L V S Y A H
DGAT1 actagctacgacataagaacctagctaatcagctgataaagtcgatcctgaaatctcctactatgtagcttgaagag 880
T S Y D I R T L A N S A D K V D P E I S Y Y V S L K S
DGAT2 actaattacgacataagaacctagctaatcagctgataaagtcgatcctgaaagctcctactatgtagcttgaagag
T N Y D I R T V A N S A D K V D P E V S Y Y V S L K S
DGAT1 cttggcgtatttcatggttctt 960
L A Y F M V A P T L C Y Q P S Y P R S P C I R K G W
DGAT2 cttggcgtatttcatggttctt
L A Y F M V A P T L C Y Q P S Y P R S P G I R K G W
DGAT1 tggctcgtcaatttgcgaaactgtcatattcactggactcgggtttataatagagcagatataataatcctattggtt 1040
V A R Q F A K L V I F T G L M G F I I E Q Y I N P I V
DGAT2 tggctcgtcaatttgcgaaactgtcatattcactggactcgggtttataatagagcaatataataatcctattggtg
V A R Q F A K L V I F T G L M G F I I E Q Y I N P I V
DGAT1 aggaactcaaacatcctctgaaagggaccttctatctgatttgaagagtggtgaagcttccagttccaatctata 1120
R N S K H P L K G D L L Y A I E R V L K L S V P N L Y
DGAT2 aggaactcaaacatccttgaaggggacttcttctatcagctatttgaagagtggtgaagcttccagttccaatctata
R N S K H P L K G D L L Y A I E R V L K L S V P N L Y
DGAT1 tgtgtgctctgcatgttctactgcttcttccaccttgggttaaacatattggcagagctgctctgctcggggaccgtg 1200
V W L C M F Y C F F H L W L N I L A E L L C F G D R
DGAT2 tgtgtgctctgcatgttctactgcttcttccaccttgggttaaatattggcagagctccttctgctcggggatcgtg
V W L C M F Y C F F H L W L N I L A E L L C F G D R
DGAT1 aattttacaagattggtggaatgcaaaaagcgttagagattatggagaatgtggaatagcctgttcacaaatggatg 1280
E F Y K D W W N A K S V G D Y W R M W N M P V H K W M
DGAT2 aattttacaagattggtggaatgcaaaaagcgttagagattatggagaatgtggaatagcctgttcacaaatggatg
E F Y K D W W N A K S V G D Y W R M W N M P V H K W M
DGAT1 gttcgacatgtatacttccctgctcgcataaagataccgaaagtagcaccgaccattatcattgcttcttcttctgctg 1360
V R H V Y F P C L R I K I P K V P A I I I A F L V S A
DGAT2 gttcgacatgtatacttccctgctcgcataaagataccgaaagtagcaccgaccattatcattgcttcttcttctgctg
V R H V Y F P C L R I K I P K V P A I I I A F L V S A
DGAT1 agtcttctcatgattatgcacatcgcagttccttccctctcttcaatctatgggcttcttcatgggaattatgcttcaggtcc 1440
V F H E L C I A V P C R L F N L W A F M G I M F Q V

Contd...

```

DGAT2 agttttccatgagctgtgctgattgcagttccttgcctgtctctttaattatgggctttcataggaatcatgtttcaggtgc
      V F H E L C I A V P C R L F N L W A F I G I M F Q V
DGAT1 ctttggtctttatcacaaaactttttacaagaagggttggctccatgggtgggaaacatgatcttctgggttcagcttctgc 1520
      P L V F I T N F L Q E R F G S M V G N M I F W F S F C
DGAT2 ctttggtctttatcacaaaactatttacaagaagggttggctccatgggtgggaaacatgatcttctgggttcagcttctgc
      P L V F I T N Y L Q E R F G S M V G N M I F W F S F C
DGAT1 attttcggacaaccgatgtgtgtgcttctttattaccatgacctcatgaaccgcaaaggatccatgtcctTGAgaaggact 1600
      I F G Q P M C V L L Y Y H D L M N R K G S M S      3'UTR
DGAT2 attttcggacaaccgatgtgtgtgcttctttattaccatgacctcatgaaccgcaaaggatccatgtcctTGAgaaggact
      I F G Q P M C V L L Y Y H D L M N R K G S M S
DGAT1 ttttacgccacagaaaaaattgggtcaattggaaagatgggagtttttgtatcccttggtagcgttataaaagacttttag 1680
DGAT2 ttttacgccacagaaaaaattgggtcaattggaaagatgggagtttttgtatcccttggtagcgttataaaagacttttag
DGAT1 agagacgaatccctttctgttctccttgtctgtcactcactggattttctatcttctcgtcttttaacaagccccaaaa 1760
DGAT2 agagacgaatccctttctgttctccttgtctgtcactcactggattttctatcttctcgtcttttaacaagccccaaaa
DGAT1 aaaaaaaaa 1768
DGAT2 aaaaaaaaa
  
```

Fig. 4—Complete sequence of *BjDGAT1* and *BjDGAT2* cDNAs with their deduced amino acids. [5'UTR and 3'UTR are underlined. A putative diacyl glycerol /phorbol ester-binding motif **HKWXXRHXYXP** and acyl CoA binding motif **FYXDWWN** are in bold. The stop codon **TGA** is at position 1590. The primers used are shown in italics]

Transcripts of *DGAT* genes

Agarose gel electrophoresis of RT-PCR product showed a prominent band of about 1.5 kb fragment. This fragment was cloned into the pGEMT-easy T/A cloning vector and transformed into *E. coli* DH5 α . Plasmid DNA was isolated from recombinant clones and restricted with *EcoRI*. Restriction pattern for recombinant clones analyzed in the gel gave interesting results. A single insert, sized about 1.5 kb was observed in one of the clone. Another clone contained a small (0.3 kb) and a large sized (1.2 kb) insert, due to presence of an internal *EcoRI* site in the insert (Fig. 3). Polymorphism for *EcoRI* restriction site indicated that two different transcripts of *DGAT* gene were present in *B. juncea*.

Sequencing results of the cDNA clones showed the presence of 1510 bp in both the cDNAs. As one of the cDNAs showed maximum sequence identity with *B. napus DGAT1*, it was designated as *BjDGAT1* (accession number: DQ016105). The second cDNA was designated as *BjDGAT2* (accession number: DQ016106). These two cDNAs were found to be members of the *DGAT1* gene family as they showed 34% and 32% sequence identity respectively with the *ACAT* gene of *Homo sapiens*. RACE was carried out to clone the untranslated region of cDNA using gene-specific primers and RACE outer primers. A 400 bp fragment was amplified by 5'-RACE technique and subsequently cloned into pGEMT T-Easy vector and sequenced. The length of 5'-UTR was found to be 80 bp (Fig. 4). The 3'-untranslated region on amplification generated a fragment of ~200 bp, which was cloned and sequenced. The length of 3'-UTR was found to be 176 bp, which had a poly A tail at its

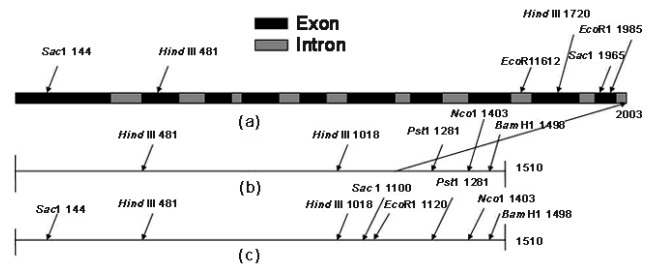


Fig. 5—Restriction map of *Brassica juncea DGAT* (a) Partial genomic sequence, (b) *BJ DGAT 1* cDNA, and (c) *BJDGAT2* cDNA. [The numbers indicated are the position of the restriction sites in the gene]

3'end (Fig. 4). The total length of the *DGAT 1* and *DGAT 2* cDNAs was 1768 bp with identical 5' and 3'-UTR (Fig. 4).

Sequence analysis of cDNA clones

Pair-wise alignment of these two cDNAs showed 85.1% sequence identity across an ORF of 1512 bp coding region beginning with start codon ATG and ending with stop codon TGA. It encoded for a polypeptide of 503 amino acids. *BjDGAT1* was 85.1% and 96% identical to *BjDGAT2* at nucleotide and amino acid level, respectively. These cDNAs showed polymorphism for *EcoRI*. Homology search analysis (Blast 2 server developed by NCBI, USA)²⁰ showed that at nucleotide level *BjDGAT1* was 98%, 95%, and 90% identical to *BnDGAT1*, *BnDGAT2*, and *A. thaliana DGAT* respectively, whereas *BjDGAT2* was 94%, 91%, and 82% identical to *BnDGAT1*, *A. thaliana DGAT*, and *G. max* respectively. Overall sequence identity varied from 98-79% across several plant species in case of *BjDGAT1* and 94-80% in case of *BjDGAT2* respectively. Restriction site analysis of

both cDNAs showed sites for important enzymes like *EcoR* I, *Hind* III, *Pst* I, *Nco*I, *Bam*H I and *Sac*I (Fig. 5). The unique feature was the presence of one *EcoR* I and two *Sac*I sites in *BjDGAT2*. *BjDGAT2* was transcribed from the partial genomic sequence of *DGAT* reported in this study. Only one *EcoR*I site was present in the cDNA as compared to corresponding genomic sequence, which had two *EcoR*I sites, as one *EcoR*I site was present in the 8th intronic region.

The primary structure of encoded protein was analyzed using ProtParam tool (<http://au.expasy.org>) and showed a protein of 503 amino acids with a molecular mass of 57.3 kD and a theoretical pI of 8.53, suggesting slightly basic nature at physiological pH. FASTA search at deduced amino acid level showed homology varying between 95% (*B. napus*) to 65% with *R. communis*. *BjDGAT2* cDNA shared maximum homology with *B. napus* *DGAT1*. At nucleotide level, it was 94%, 91%, and 82% identical to *BnDGAT1* and 2, *A. thaliana* *DGAT* and *G. max* *DGAT* respectively. *BjDGAT2* coded for a polypeptide of 503 amino acids, having molecular mass of 57.22 kD and a pI 8.65. At amino acid level homology was found to vary from 88% (*B. napus*) to 65% (*P. frutescens*). Analysis using pfam programme showed *DGAT* protein to belong to MBOAT (membrane bound O-transferase) protein family. A putative diacylglycerol/phorbol ester binding motif **HKWXXRHXYXP** which is unique to *DGAT*, while absent in ACATs^{21,22} was observed after 396 amino acids in *BjDGAT2* also. A **FYXDWWN** motif required for acyl CoA, also reported in *BnDGAT1* and ACAT families was found conserved after 374 amino acids in *BjDGAT1* and 2 proteins. The hydropathy plot of the protein encoded by transcripts analyzed by TMHM package (available at <http://cbs.dtu.dk/cgi>) revealed that both had multiple hydrophobic domains. It was predicted that both *BjDGAT1* and 2 had 9 possible transmembrane segments at the same position (between amino acids numbers, 116-135, 159-178, 190-212, 217-239, 298-320, 346-368, 408-430, 440-462, 474-496), as shown in *A. thaliana* *DGAT*²³. N-terminal hydrophilic segment of 115 amino acids was predicted to lie on the cytoplasmic side of the membrane²⁴. A number of putative serine (18), tyrosine (4), and threonine (3) phosphorylation sites are also present in *BjDGAT* protein.

Two *DGAT* transcripts were also reported earlier from a microspore-derived cell suspension culture of

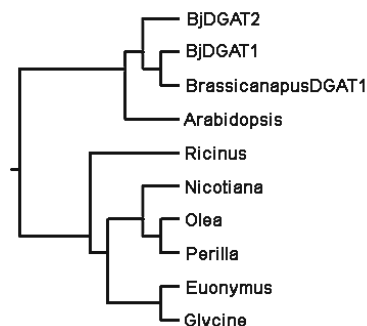


Fig. 6—Dendrogram of *DGAT* constructed from different plant species using the deduced protein sequence of *BjDGAT1* (DQ016106) *BjDGAT2* (DQ016107), *B. napus* (AF164434), *A. thaliana* (AF051849), *R. communis* (AY366496), *N. tabacum* (AF129003), *O. europaea* (AY445635), *P. frutescens* (AF298815), *E. alatus* (AY751297) and *G. max* (AY652765)

*B. napus*¹¹ and in olive drupe tissues¹⁶. The presence of two different *DGAT* transcripts might reflect the existence of at least two copies of the genes consistent with the tetraploid nature of *B. juncea* (AABB). Further, role for these two different *DGAT* enzymes with localization in mustard tissues remains to be determined. Also, it is not known whether these two proteins (*DGAT1* and 2) have the same catalytic activity or their activity depends on other factors. Studies in castor showed 18-fold higher expression of *DGAT2* in contrast to *DGAT1* in seeds than in leaves and expression was temporal-specific during seed development¹⁴.

Phylogenetic relationship of various plant species for *DGAT* sequence was analyzed using constructed dendrogram (Fig. 6) to indicate the *DGAT* amino acids variation among different plant species. In the dendrogram, *DGAT* from *B. juncea*, *B. napus* and *A. thaliana* were grouped together in the same cluster, and those from *N. tabacum*, *R. comunis*, *Olea europaea*, *P. frutescens*, *Euonymus alatus* and *G. max* grouped separately, showing that these were distantly related to the previous one evolutionarily. Clustal W alignment for these proteins revealed that divergence between *DGATs* was concentrated in the N-terminal while many conserved regions were present in the middle and towards the C-terminal region.

Although the regulatory role of *DGAT* in oil synthesis has been suggested⁸, complete understanding of the regulation of Kennedy pathway, which is vital for the rational engineering of transgenic oil crops is still lacking. As *DGAT* is less specific to acyl-CoA, studies on remodeling the active site by site-directed mutagenesis may provide further

insights into the specificities towards unique substrates.

In conclusion, the present study showed *BjDGAT1* cDNA was 85.1% identical to *BjDGAT2* across an ORF of 1512 bp and at deduced amino acid level, 96% identical across 503 overlapping amino acids. A large variation at the N-terminal region of *BjDGAT1* and 2 might reflect their unique property. Both were alkaline in nature (pI, 8.5-8.6), had similar molecular size (56-57 kD), N-terminal hydrophilic segment and nine transmembrane segments. These transcripts encoding *DGAT* could be used in developing transgenic plants, which may prove useful in studying the regulatory mechanism of TAG synthesis, altering seed oil profile and in enhancing the oil levels.

Acknowledgement

Financial assistance by CSIR as Senior Research Fellowship to N. Ilaiyaraja is gratefully acknowledged. This work was supported by ICAR National Fellow Project awarded to Dr. I M Santha.

References

- 1 Drexler H, Spiekermann P, Meyer A, Domergue F, Zank T, Sperling P, Abbadi A & Heinz E (2003) *J Plant Physiol* 160, 779-802
- 2 Cao Y & Huang A H C (1987) *Plant Physiol* 84, 762-765
- 3 Kaup M T, Froese C D & Thompson J E (2002) *Plant Physiol* 129, 1616-1626
- 4 Lardizabal K D, Mai J T, Wagner N W, Wyrick A, Voelker T & Hawkins D J (2001) *J Biol Chem* 276, 38862-38869
- 5 Kalscheuer, R & Steinbuchel A (2003) *J Biol Chem* 278, 8075-8082
- 6 Dahlqvist A, Stahl U, Lenman M, Banas A, Lee M, Sandager L, Ronne H & Stymne S (2000) *PNAS* 97, 6487-6492
- 7 Stahl U, Carlsson AS, Lenman M, Dahlqvist A, Huang B, Banas W, Banas A & Stymne S (2004) *Plant Physiol* 135, 1324-1335
- 8 Hobbs D H, Lu C & Hills M J (1999) *FEBS Lett* 452, 145-149
- 9 Jako C, Kumar A, Wei Y, Zou J, Barton D L, Giblin E M, Covello P S & Taylor D C (2001) *Plant Physiol* 126, 861-874
- 10 Fa-Yun Zhang, Ming-Feng Yang & Yi-Nong Xu (2005) *Plant Sci* 169, 689-694
- 11 Nykiforuk C L, Laroche A & Weselake R J (1999) *Plant Physiol* 121, 1053
- 12 Nykiforuk C L, Laroche A & Weselake R J (1999) *Plant Physiol* 120, 1207
- 13 He X, Chen G Q, Lin J T & McKeon T A (2004) *Lipids* 39, 865-871
- 14 Kroon J T M, Wei W, Simon W J & Slabas A R (2006) *Phytochemistry* 67, 2541-2549
- 15 Milcamps A, Tumaney A W, Paddock T, Pan D A, Ohlrogge J & Pollard M (2005) *J Biol Chem* 280, 5370-5377
- 16 Giannoulia K, Haralampidis K, Poghosyan Z, Murphy D J & Hatzopoulos P (2000) *Biochem Soc Trans* 28, 695-697
- 17 Bouvier-Nave P, Benveniste P, Noirielle A & Schaller H (2000) *Biochem Soc Trans* 28, 692-695
- 18 Murray M G & Thompson W F (1980) *Nucleic Acids Res* 8, 4321-4325
- 19 Sambrook J, Fritsch E F & Maniatis T (1989) *Molecular Cloning: A laboratory manual*, 2nd edn, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York
- 20 Altschul S F, Madden T L, Miler W, Myers E W & Lipman DJ (1990) *J Mol Biol* 215, 403-410
- 21 Zou J, Wei Y, Jako C, Kumar A, Selvaraj G & Taylor D C (1999) *Plant J* 19, 645-653
- 22 Oelkers P, Behari A, Cromley D, Billheimer J T & Sturley S T (1998) *J Biol Chem* 273, 26765-26771
- 23 Bouvier-Nave P, Oelkers P, Benveniste P, Sturley S L & Schaller H (2000) *Eur J Biochem* 267, 85-96
- 24 Weselake R J, Pomeroy M K, Furukawa T L, Golden J L, Little D B & Laroche A (1993) *Plant Physiol* 102, 565-571