

Characterization of Thioesterase Gene in *Sesamum indicum* using Bioinformatics Tools

Spandana B¹ and Prakash DSRS^{2*}

¹Agricultural College, Acharya NG Ranga Agricultural University (ANGRAU), Andhra Pradesh, India

²Department of Human Genetics, Andhra University, Visakhapatnam, India

Abstract

Thioesterases play a central role in determining chain lengths of fatty acids in oil storage tissues and have been isolated from a number of plant sources. TEs can be classified into two evolutionarily related families, namely FatA and FatB, with distinct substrate selectivity. With the help of bioinformatics tools alignment was done among the selected thioesterase genes from related oilseed crops. Degenerate primers were designed by manual, codehop and webprimer methods for the conserved regions. Gradient PCR with designed primers showed amplification with manual and codehop primers in *Sesamum* DNA. Amplified DNA is transformed and plasmids are sequenced. Sequences are searched for homology with BLAST. The homology at the amino acid level for thioesterase sequences was observed to be with *Helianthus annuus* than with the other oil seed crops compared.

Keywords: Degenerate primers; *Sesamum indicum*; Gradient PCR; Thioesterase

Abbreviations: TEs: Thioesterase; BLAST: Basic Local Alignment Search Tool; PCR: Polymerase Chain Reaction.

Introduction

Sesame (*Sesamum indicum*) is a highly-valued oilseed crop of Asia that produces seeds with 50% of dry weight oil, and which is commonly used in oriental cooking. The seed oil of sesame (*Sesamum indicum*) contains mainly four fatty acids viz., palmitic, stearic, oleic and linoleic acids. Sesame oil is unique in containing several lignan antioxidants as well as some tocopherols. Oleic and linoleic acids constitute around 85% of oil composition and their inheritance is controlled by single gene [1].

Major aims for molecular genetic improvement of plant oils have been to improve the total oil content of seeds, and to increase the content of specific target fatty acids in the oil to improve nutritional or functional properties and expand uses. In several instances, a functional genomics approach has been successfully used to identify the genes responsible for the production of specific fatty acids [2]. More recently, the development of genomics technology has opened doors for large-scale approaches aiming to uncover and understand the regulatory factors that govern synthesis and accumulation of fatty acids in tissues. Plant acyl-acyl carrier protein (ACP) thioesterases are of biochemical interest because of their roles in fatty acid synthesis and their utilities in bioengineering of plant oil seeds [3]. Keeping in view the above facts the present study was undertaken to isolate and characterize the candidate thioesterase gene from the sesame. The thioesterase gene plays vital role in fatty acid synthesis for the production of specific fatty acids from the *Sesamum indicum*. Thioesterases play an essential role in chain termination during de novo fatty acid synthesis and in the channeling carbon flux between the two lipid biosynthesis pathways in plants.

Material and Methods

The plant species harboring the thioesterase which have already been reported were collected from the previous literature and from NCBI and EBI data banks. Thioesterase nucleotide sequences of *B. juncea*, *G. hirsutum*, *H. anus*, *S. indicum*, *A. hypogaeae*, *R. communis*, *G. max* were taken from NCBI database and they are converted into

protein sequences by <http://www.expasy.ch/tools/dna.html>. The consensus amino acid and nucleotide sequences of thioesterases (Fat A1) were selected by using the Protein Domain database, Pro Dom (<http://prodes.toulouse.inra.fr/prodom/doc/prodom.html>), NCBI (<http://ncbi.nlm.nih.gov>) and EBI (www.ebi.ac.uk) tools and thioesterase as the query. The most conserved region is identified by multiple sequence alignments by using Clustal W software. The similarity matrix the phylogenetic trees were produced using DNA Star (version 2004). Based on these conserved regions, gene-specific degenerate primers for thioesterases were designed manually and by using few software programmes like Web Primer and Codehop [4,5]. The consensus of these protein and DNA sequence alignments gave rise to 3 sets of primer pairs which are listed in Table 1. Degeneracy in primer sequence was also taken into consideration. Apart from the above mentioned, the universal site for most of the bioinformatics tools were also used in the present study (www.expasy.org). All the sequences were given for multiple sequence analysis by Clustal W. The aligned sequences were given to BLOCKMAKER (http://bioinformatics.weizmann.ac.il/blocks/blockmkr/www/make_blocks.html) with each block at a time to BLOCKMAKER. The blocks obtained with the maximum similarity and minimum degeneracy was selected. These blocks were given to Codehop (<http://bioinformatics.weizmann.ac.il/blocks/codehop.html>) for designing of primers. The primers thus designed were selected based on minimum degeneracy and length and percentage GC.

Three sets of gene specific degenerate primers (manual, Web primer, Codehop) synthesized by Biotech Desk, India Pvt. Ltd. were tested for PCR amplification with one sesame sample PCR reactions

***Corresponding author:** Prakash DSRS, Department of Human Genetics, Andhra University, Visakhapatnam, India, Tel: +9133102078; E-mail: drdevaprakash@gmail.com

Received January 16, 2017; **Accepted** February 02, 2017; **Published** February 14, 2017

Citation: Spandana B, Prakash DSRS (2017) Characterization of Thioesterase Gene in *Sesamum indicum* using Bioinformatics Tools. Adv Crop Sci Tech 5: 256. doi: 10.4172/2329-8863.1000256

Copyright: © 2017 Spandana B, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

were carried out in a DNA thermocycler (Eppendorf, USA) with a heated lid. Each 20 µL reaction volume contained about 100 ng of template DNA, 1X PCR buffer (10 mM Tris-HCl pH 5.3, 50 mM KCl), 2 mM MgCl₂ (Invitrogen, USA), 2 mM dNTP, (Sigma -Aldrich., USA), 5 pmol each of Forward and Reverse primer and 1U of *Taq* DNA polymerase (Stratagene, Germany) and programmed for an initial denaturation step of 5 min at 94°C, followed by 35 cycles of 1 min at 94°C, 1 min at 50°C, extension was carried out at 72°C for 1 min and a final extension at 72°C for 10 min and a hold temperature of 14°C at the end. Negative controls were also run without template DNA to avoid non-specific amplification.

PCR products were electrophoresed on 1.4% agarose gels, in 1X TAE Buffer at 50 V for 3 h and then stained with ethidium bromide (0.5 µg/µL). PCR reactions were repeated three times to ascertain the reproducibility of the reaction.

Gel elution

Fragments ranging from 400-700 bp were eluted by using Column method (Qiagen Kit).

Ligation

Ligation reaction was set up for 10 µL reaction volume containing 1µL of pGEM easy vector (Promega, USA), 5 µL of PCR enriched (gel eluted) product, 1ml of 10 X T₄ DNA ligation buffer (New England Biolabs, USA), 2.4 U/µL of T₄ DNA ligase (New England Biolabs, USA)

Primer pair	Primer pair (5'-3')	Base pairs	Product size(bp)
Thio M	F- TAYGARGTBGGDAKTAAYAARAC R- NGAYTCNACNACYTCYTCDTGDTGGC	23 26	400
Thio W	F- ATGTTGAAGCTTTCGTGT R- TGAAGGCTTCTTCTCCACA	18 20	400
Thio C	F- CAGGAGGTGGGCTGCAAYCAYGCNC R- TTRRTNCARRTGCTGCGGCTGTCCG	25 25	400

Degenerate bases information

N=A/T/G/C
R=A/G
M=A/C
W=A/T
S=G/C
Y=C/T
K=G/T
V=A/G/C
H=A/C/T
D=A/G/T
B=C/G/T

Table 1: Characteristics of the degenerate primers designed and predicted sizes of PCR amplification fragments of known Thioesterase gene.

and incubated at 14°C overnight. The samples were stored at - 20°C until used.

Transformation

Transformation of ligated product was done in the same manner as described elsewhere. The plasmid isolation was done by Alkaline lysis method. The selected plasmids of amplification range (400 bp for thioesterase) after PCR was given for sequencing.

The vector sequences were removed form the sequenced clones using Vecscreen (NCBI) and homology search was carried.

Results

The primers are designed by manually, (by ta webprimer and codehop methods which are given in Table 1.

Thioesterase characterization

The sequences for characterization of Thioesterase gene are selected from the following edible oil yielding species *Brassica rapa*, *Helianthus annuus*, *Ricinus communis* and converted to amino acid sequences to identify the conserved regions which are given in Supplementary file1.

The primers are designed (including C terminal region as N terminal region has more gaps it is excluded) by manual, webprimer and codehop methods which are given in Table 2.

Three sets of gene specific degenerate primers (manual, Web primer, Codehop) synthesized by Biotech Desk, India Pvt. Ltd.) were tested for PCR amplification with one sesame sample. The Figure 1 showing gradient PCR amplification is given below.

Lanes 1-11: *S. indicum* DNA amplified with Codehop designed primer using a gradient from 45-55°C with a 0.9°C raise at each lane.

Lanes 12-21 : *S. indicum* DNA amplified with Web designed primer using a gradient from 45 -55°C with a 0.9°C raise at each lane.

Lanes 22-31: *S. indicum* DNA amplified with manually designed primer using a gradient from 45 -55°C with a 0.9°C raise at each lane.

Accession no	Organism	Max score
NM100724.2	Arabidopsis thaliana FATB (fatty acyl-ACP thioesterases B);	1023
NM870270.1	Brassica napus lipase 2 mRNA, complete cds	25.7
NM022816.2	<i>Mus musculus</i> acyl-CoA thioesterase 10 (Acot10), mRNA	24.7
NM147161.2	<i>Homo sapiens</i> acyl-CoA thioesterase 11 (ACOT11), transcript variant 2, mRNA	22.9

Table 2: Homology results of Thioesterase.

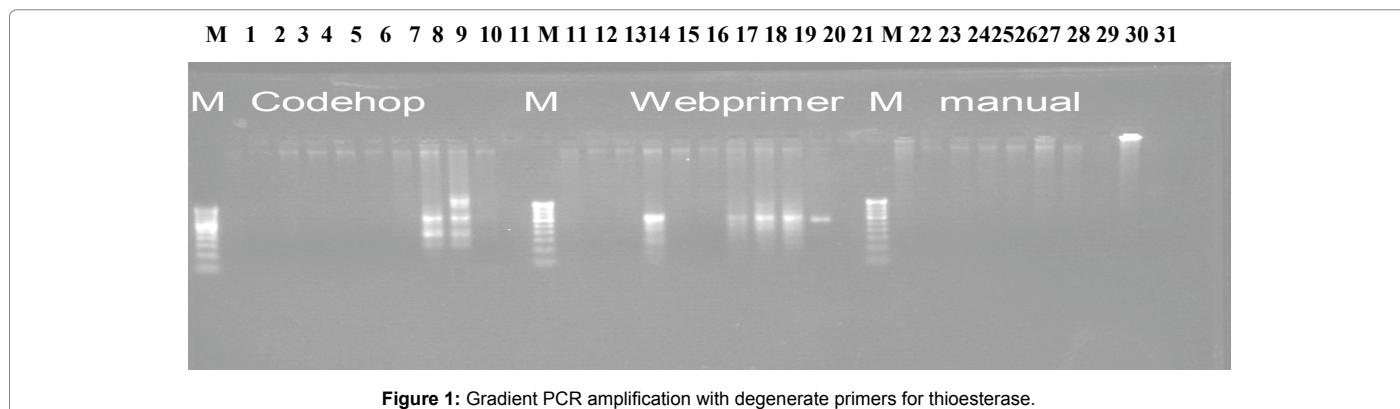


Figure 1: Gradient PCR amplification with degenerate primers for thioesterase.

Lane M-100 bp marker with bands 100 to 1000 bp bands DNA ladder.

Since the designed gene primers of thioesterase gene will amplify the DNA at 400 bp, that band is eluted from the gel, transformed and isolated plasmids from selected colonies were sequenced for further conformation.

Of the 10 PCR positive clones obtained 5 were sequenced. The sequences obtained were subjected to BLAST search.

Search results of Thioesterase in the NCBI for homologs

Four sequenced thioesterase sequences were submitted to the NCBI (National Centre for Biotechnology Information) website for BLAST-N (<http://www.ncbi.nlm.nih.gov/BLAST/>)[6]. The sequence results indicated that the thioesterase gene cloned from *S. indicum* had high homology at the nucleotide and amino acid level with those of *A. thaliana*, *Mus musculus* and *B. napus* (Figures: 2 and 3).

Most of the sequences showed high homology with *Arabidopsis thaliana* FATB (fatty acyl-ACP thioesterases B); acyl carrier/acyl-[acyl-carrier-protein] hydrolase (FATB) mRNA, complete cds.

Arabidopsis lyrata sub sp. *lyrata* hypothetical protein, mRNA *Mus musculus* acyl-CoA thioesterase 10 (Acot10), mRNA.

Mus musculus acyl-CoA thioesterase 9 (Acot9), nuclear gene encoding mitochondrial protein, mRNA.

Rattus norvegicus acyl-CoA thioesterase 9 (Acot9), nuclear gene encoding mitochondrial protein, mRNA.

Homo sapiens acyl-CoA thioesterase 11 (ACOT11), transcript variant 2, mRNA.

Homo sapiens acyl-CoA thioesterase 11 (ACOT11), transcript variant 1, mRNA.

Homo sapiens acyl-CoA thioesterase 9 (ACOT9), nuclear gene encoding mitochondrial protein, transcript variant 1, mRNA.

Brassica napus lipase 2 mRNA, complete cds.

Arabidopsis thaliana FATB (fatty acyl-ACP thioesterases B); acyl carrier/ acyl-[acyl-carrier-protein] hydrolase (FATB) mRNA, complete cds *Arabidopsis lyrata* sub sp. *lyrata* hypothetical protein, mRNA.

Arabidopsis thaliana clone pENTR221-At4g27190 disease resistance protein (At4g27190) mRNA, complete cds.

The obtained sequences are again aligned with the originally

selected sequences of oil seed crops by multiple sequence alignment and results are given in Supplementary File 2.

The homology was based on small portion of entire coding sequence, as the sequence length of the obtained sequence was 126 to 429 bp, where as the length of the selected sequences was 1089 to 1696 bp. The expected DNA fragment range was 400 bp.

The maximum homology of the obtained sequences was observed to be with *Helianthus annuus* indicating that the thioesterase isolated from *S. indicum* was closer to *H. annuus* than the other oil crops compared.

Discussion

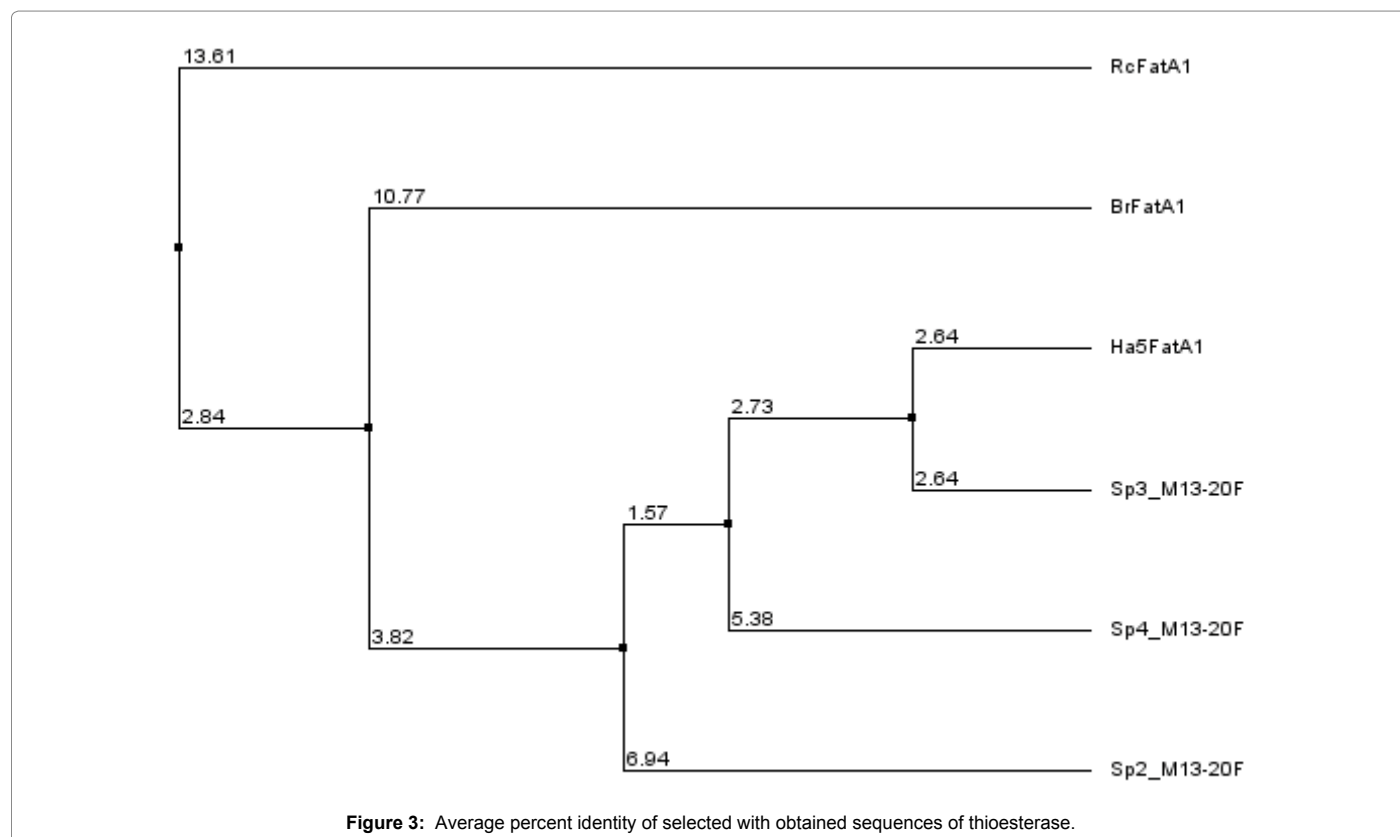
Bioinformatics approaches to lipid research have recently begun using large amounts of mass spectrometry and microarray data. However, to our knowledge, there is no report that describes the investigation of fatty acid structures based on the comprehensive analysis of the gene contents in the genomes. In fact, the prediction of fatty acid structures from genomic information is difficult because of the functional diversity of the key enzymes. Various genes encoding enzymes involved in fatty acid synthesis have been isolated from the species and characterized. Thioesterases are the key enzymes in fatty acid biosynthesis in sesame appear to be showing homology to different oilseed *helianthus*. The reasons are unknown but a more detailed study of these enzymes across the different oilcrops will through some light on the fatty acid biosynthesis pathway. Variation in oil content and composition could also be contributed by variations in this enzyme. A combination of bioinformatics tools were used to predict a 3D model for *Arabidopsis* FatB (AtFatB) comprising a fold similar to that of *E. coli*.

Thioesterase 11 an enzyme similar to that of plant Thioesterases functionally, lacks significant sequence similarity and displays different inhibitor sensitivity [7].

In the present study after analysis of available sequences of Thioesterase related proteins, retrieved thioesterase subgroup clusters and selected six representative sequences of most members of thioesterase which specifically belonged to dicots as *Sesamum* is a eudicot species. In the present study codehop and web primer showed amplification for thioesterase. Thioesterase derived sequences also showed high homology with *Arabidopsis*, *Brassica sp*, *Mus musculus* and *Homosapiens* thioesterase genes. Basic bioinformatics techniques are powerful tools in terms of leading to the discoveries and analysis of novel genes [8]. Even though the results from bioinformatics studies are very helpful in directing and designing the experiments, they need to be supported and confirmed by further experimentation.

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
XM_001111108.2	PREDICTED: Macaca mulatta acyl-CoA thioesterase 12 (ACOT12), m	35.6	35.6	6%	0.91	86%	GM
XM_002744825.1	PREDICTED: Callithrix jacchus acyl-CoA thioesterase 12 (ACOT12), n	35.6	35.6	6%	0.91	86%	GM
NM_130767.2	Homo sapiens acyl-CoA thioesterase 12 (ACOT12), mRNA	35.6	35.6	6%	0.91	86%	UEGM
XM_517670.2	PREDICTED: Pan troglodytes cytosolic acetyl-CoA hydrolase (ACOT1	35.6	35.6	6%	0.91	86%	GM
BC075011.2	Homo sapiens acyl-CoA thioesterase 12, mRNA (cDNA clone MGC:10	35.6	35.6	6%	0.91	86%	UEGM
BC075010.2	Homo sapiens acyl-CoA thioesterase 12, mRNA (cDNA clone MGC:10	35.6	35.6	6%	0.91	86%	UEG
BC089437.1	Homo sapiens acyl-CoA thioesterase 12, mRNA (cDNA clone IMAGE:3	35.6	35.6	6%	0.91	86%	UEG
AK122960.1	Homo sapiens cDNA FI 1166R0 fis clone T11VF200R271 hinhlv similar	35.6	35.6	6%	0.91	86%	UEG

Figure 2: Blast results with Sp1_M13-20F.



References

1. Brar G, Ahuja KL (1979) Sesame: Its culture, genetics, breeding and biochemistry. *Annu Rev Plant Sci* 285-313.
2. Ohlgrogge J, Mhaske VB, Beisson F, Ruuska S (2004) Genomics approaches to lipid biosynthesis. *Proceedings of the 4th International Crop Science Congress, Brisbane, Australia*, 26.
3. Yuan L, Voelker TA, Hawkins DJ (1995) Modification of the substrate specificity of an acyl-acyl carrier protein thioesterase by protein engineering. *Proc Natl Acad Sci* 92: 10639-10643.
4. Wilks AF, Kurban RR, Hovens CM, Ralph SJ (1989) The application of the polymerase chain reaction to cloning members of the protein tyrosine kinase family. *Gene* 85: 67-74.
5. Chen H, Zhu G (1997) Computer program for calculating the melting temperature of degenerate oligonucleotides used in PCR or hybridization. *Biotechniques* 22: 1158-1160.
6. Altschul SF, Madden TL, Shaffer AA, Zhang J, Zhang G, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389-3402.
7. Mayer KM, Shanklin JA (2004) Structural Model of the plant acyl-acyl carrier protein Thioesterase Fat B comprised of two Helix/4 stranded sheet domains, The N terminal Domain containing Residues that affect specificity, the C-terminal Domain containing catalytic residues. *J Biol Chem* 380: 3621-3627.
8. Hu Z, Chen K, Wang L, Yao Q (2005) Identification and characterization of *Bombyx mori* eIF5A gene through bioinformatics approaches. *In Silico Biol* 5: 573-580.

OMICS International: Open Access Publication Benefits & Features

Unique features:

- Increased global visibility of articles through worldwide distribution and indexing
- Showcasing recent research output in a timely and updated manner
- Special issues on the current trends of scientific research

Special features:

- 700+ Open Access Journals
- 50,000+ editorial team
- Rapid review process
- Quality and quick editorial, review and publication processing
- Indexing at major indexing services
- Sharing Option: Social Networking Enabled
- Authors, Reviewers and Editors rewarded with online Scientific Credits
- Better discount for your subsequent articles

Submit your manuscript at: <http://www.omicsonline.org/submit>

Citation: Spandana B, Prakash DSRS (2017) Characterization of Thioesterase Gene in *Sesamum indicum* using Bioinformatics Tools. Adv Crop Sci Tech 5: 256. doi: 10.4172/2329-8863.1000256