

#### **Research Article**

**Open Access** 

# Cloning and Bioinformatics Analysis of Genes Related to Key Enzymes in Oil Metabolism of *Carya illinoensis* (Wangech.) K. Koch

Jia Zhanhui <sup>1,2</sup>, Jia Xiaodong <sup>1,2</sup>, Xu Mengyang <sup>1,2</sup>, Mo Zhenghai <sup>1,2</sup>, Yang Xufeng <sup>1,2</sup>, Zhai Min <sup>1,2</sup>, Xuan Jiping <sup>1,2</sup>, Zhang Jiyu <sup>1,2</sup>, Wang Gang <sup>1,2</sup>, Wang Tao <sup>1,2</sup>, Guo Zhongren <sup>1,2</sup>

1 Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing, 210014, China

2 The Jiangsu Provincial Platform for Conservation and Utilization of Agricultural Germplasm, Nanjing, 210014, China

Co-corresponding authors email: <u>jiaxiaodong@cnbg.net</u>; <u>xuanjiping@cnbg.net</u>

Bioscience Method, 2022, Vol.13, No.1 doi: 10.5376/bm.2022.13.0001

Received: 25 Nov., 2021

Accepted: 07 Jan., 2022

Published: 18 Jan., 2022

Copyright © 2022 Jia et al., This is an open access article published under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Preferred citation for this article:

Jia Z.H., Jia X.D., Xu M.Y., Mo Z.H., Yang X.F., Zhai M., Xuan J.P., Zhang J.Y., Wang G., Wang T., and Guo Z.R., 2022, Cloning and bioinformatics analysis of genes related to key enzymes in oil metabolism of *Carya illinoensis* (Wangech.) K. Koch, Bioscience Method, 13(1): 1-13 (doi: 10.5376/bm.2022.13.0001)

**Abstract** Taking mixed sample of 115 d and 135 d *Carya illinoinensis* 'Pawnee' as experimental materials, after RT-PCR amplification, cloning and sequencing, the gene sequences of *CiSAD* gene, *CiFAD2* gene and *CiGPAT* gene related to the key enzymes of oil metabolism in *Carya illinoinensis* were obtained, and bioinformatics analysis was carried out. The length of *CiSAD* gene was 1 240 bp, containing a complete open reading frame (ORF) of 1 194 bp, encoding 397 amino acids. *CiFAD2* was 1 329 bp, containing an ORF of 1 155 bp, encoding 384 amino acids. *CiGPAT* was 1 671 bp, containing an ORF of 1 626 bp, encoding 541 amino acids. All proteins encoded by the cloned genes have corresponding domains, indicating that the proteins encoded by these genes should have corresponding functions in organisms.

Keywords Carya illinoensis (Wangech.) K. Koch; Oil metabolism; Gene cloning; Bioinformatics analysis

Pecan (*Carya illinoensis* (Wangech.) K. Koch) originated in the United States. At present, the pecan central producing area is the United States, and distributed in the Mexico, Italy, France, Israel, Japan, China and other places (Zhang and Lv, 1998, Guangxi Forestry Science, 27(4): 202 - 206). *Carya illinoensis* can be traced back to the Cretaceous, and it is one of the important dry fruit trees in the world (Wang et al., 2009; 2010). Compared with *Juglans regia* and *Carya cathayensis* in China, *Carya illinoensis* has thin shell with high yield (1 500~2 250 kg/ha) and high oil yield, which is easy to extract kernel. And *Carya illinoensis* contains higher unsaturated fatty acids (UFA) and various amino acids beneficial to human body than olive, but also rich in vitamin B1, B2 and vitamin E (Pan, 2008, Xiandai Horticulture (2): 12-13; Hong, 2007). The highest content of nutrients in *Carya illinoensis* kernels is oil. Mature kernels contain 55%~75% oil, and are mainly triglycerides (TAG) composed of glycerol and fatty acids.

The oil content of *Carya illinoensis* is high. Among them, oleic acid accounts for a high proportion, which is an excellent material for studying the regulation mechanism of lipid metabolism, especially the regulation mechanism of monounsaturated fatty acid (MUFA) metabolism. Stearoyl-ACP desaturase (SAD) catalyzes the dehydrogenation of palmitic acid to produce a double bond to form oleic acid, which affects the proportion of saturated fatty acid (SFA) and MUFA and is the key enzyme in the lipid metabolism pathway of *Carya illinoensis*. At present, *SAD* gene has been isolated from various plants such as *Arabidopsis thaliana* (Ping et al., 2008), *Glycine max* (Byfield et al., 2006), *Arachis hypogaea* L. (Florin et al., 2010; Dong et al., 2012), *Elaeis guineensis* (Shah et al., 2000), *Camellia oleifera* (Zhang et al., 2008; Zhang et al., 2008) *Brassica campestris* (Knutzon et al., 1992), *Solanum tuberosum* (Taylor et al., 1992; Li et al., 2015), *Xanthoceras sorbifolia* (Zhao et al., 2015). Omega-6 fatty acid desaturase (FAD2) catalyzes the dehydrogenation of oleic acid to produce the second double bond to form linoleic acid, which is another key enzyme in the lipid metabolism pathway of *Carya illinoensis*. At present, *FAD2* gene has been isolated from various plants such as *Xanthoceras sorbifolia* (Zhao et al., 2015), *Brassica napus* (Xiong et al., 2002), *Glycine max* (Li et al., 2007), *Olea europaea* var *sylvestris* (Georgios et al., 2005), *Gossypium hirsutum* (Kargiotidou et al., 2008). Glycerol-3-phosphate acyltransferase (GPAT) catalyzes the first-step of acylation reaction of various glyceride such as triacylglycerol and phosphatidylglycerol, and can



acylate the sn-1 position of triacylglycerol (Shockey et al., 2016), which plays a very important role in the accumulation of vegetable oil. In *Carya illinoensis*, GPAT catalyzes the first step of the TAG synthesis reaction to produce 1-acyl glycerol ester. There are three types of GPAT in plant cells, which are located on plastids, cytoplasm and mitochondria, respectively. *GPAT* genes have been cloned from *Arabidopsis thaliana* and other plants (Nishida et al., 1993).

In the transcriptome study of *Carya illinoensis* kernels during fruit quality formation, we found that the unigeness encoding *SAD*, *FAD2* and *GPAT* genes in the lipid metabolism pathway of *Carya illinoensis* were highly expressed, indicating that they may play an important role in the lipid metabolism of *Carya illinoensis* (Jia at al., 2018). In this study, the gene sequences of these three key enzyme genes *SAD*, *FAD2* and *GPAT* were cloned and analyzed by bioinformatics, to lay the foundation for the in-depth study of high oil mechanism and breeding of *Carya illinoensis*.

# **1 Materials and Methods**

# 1.1 Materials

# 1.1.1 Plant materials

The test material was *Carya illinoinensis* 'Pawnee', which was cultivated in the Production Base of Nanjing Luzhou Pegan Co., Ltd., Luhe District, Nanjing. The tree-age was  $8\sim12$  a, with the spacing in the rows and spacing between rows of 5.0 m×7.0 m. 6 plants with good growth and relatively consistent tree potential were selected, and 2 plants were divided into one plot, which were repeated 3 times. In 2014, samples were taken at 115 d and 135 d after anthesis, and 2 healthy and full fruits without pests and diseases were taken from each plant in four directions of east, south, west and north. Samples were put into the ice box and taken back to the lab quickly. The samples were stored overnight in a refrigerator at -20°C, and then cut open to take kernels. The kernels were frozen and mixed in liquid nitrogen and stored in the refrigerator at -80°C.

# 1.1.2 Reagent

RNAsimple Total RNA Kit (TIANGEN BIOTECH), HiScript® II Q RT SuperMix, (Vazyme), 2 × Taq Master Mix (Vazyme). The primers were synthesized by Shanghai Invitrogen Biotechnology Co., Ltd.

# 1.1.3 Instrument

SimpliAmp<sup>™</sup> PCR, Hettich centrifuge (Andreas Hettich GmbH & Co. KG, Tuttlingen, Germany), pipette (Eppendorf, Germany).

# 1.2 Methods

# 1.2.1 Total RNA extraction

Taking mixed sample of 115 d and 135 d *Carya illinoinensis* 'Pawnee' as experimental materials. Total RNA was extracted according to the instructions of RNAsimple Total RNA Kit of Tiangen Biotech (Beijing) Co., Ltd. The ratio of OD260/OD280 of total RNA was determined by ultraviolet spectrophotometer to determine the purity and concentration of RNA. Detected the integrity of RNA with 1% agarose gel electrophoresis.

# 1.2.2 Reverse transcription

The first strand of cDNA was synthesized according to the HiScript® II Q RT SuperMix for qPCR (+gDNA wiper) instructions as follows: The template RNA, RNase free ddH<sub>2</sub>O and  $4 \times$  gDNA wiper Mix were added to the RNase-free centrifuge tube, and gently mixed with a pipette. The genomic DNA was removed at 42°C for 2 min. Then added 5 × HiScript II qRT SuperMix II to prepare a reverse transcription reaction system, gently mixed with a pipette. Reverse transcription was performed at 50°C for 15 min and 85°C for 5 sec in turn.

# 1.2.3 Primer design

Primers (Table 1) were designed with the help of the unigene sequence obtained from the previous transcriptome (Jia et al., 2018).

Gene name	Primer name	Nucleotide sequence (5'-3')	
CiSAD	SAD-F	TCGACAAACAGAAGACAAC	
	SAD-R	GCATAAGGACCTCCACTC	
CiFAD2	FAD2-F	ACCCAACAACACCGAAAC	
	FAD2-R	ATCTCCTGGATTGCACATAG	
CiGPAT	GPAT-F	CACAGCCAGCATCCGTA	
	GPAT-R	CACCCAAAACGTGTTCG	

Table 1 Sequence of primers for pecan

1.2.4 PCR amplification of target gene fragment

Refer to the instructions of  $2 \times \text{Taq}$  Master Mix (Vazyme), and the reaction system was as follows (Table 2).

Table 2 Reaction of PCR amplification

Reaction	Volume (µL)	
ddH <sub>2</sub> O	To 50	
$2 \times Taq$ Master Mix	25	
Primer 1 (10 μM)	2	
Primer 2 (10 μM)	2	
Template DNA	5	

1.2.5 Recovery, ligation, transformation, identification and sequencing of the target fragment

The target fragment was recovered and purified by Biospin agarose gel DNA recovery kit, and the steps were completely in accordance with the instructions. After the recovery, 3  $\mu$ L of recovered and purified DNA products were taken for 1% agarose gel electrophoresis detection, and the quality of recovered DNA products was detected by ultraviolet spectrophotometer. The recovered DNA solution can be immediately used for ligation reaction or stored at -20°C.

The successfully recovered target fragment was ligated to pMD18-T vector (TaKaRa, Japan), and the reaction system was shown in Table 3. After mixing, reacted at 16°C for more than 4 h or overnight. Then transformed into *Escherichia coli* and performed colony PCR. The PCR product length of the appropriate monoclonal sent to Genepioneer Biotechnologies for sequencing.

8 8		
Reaction	Volume (µL)	
Insert DNA	5.0	
Solution I	4.5	
pMD18-T vector	0.5	
Total	5	

Table 3 Reaction of target fragment

1.2.6 Sequence analysis

The nucleotide sequences obtained by sequencing were compared by BLASTn in NCBI, and then the deduced amino acid sequences were compared by BLASTp. The plant amino acid sequences with high homology were selected for homology sequence alignment and phylogenetic tree construction. Physicochemical property of the protein was analyzed by Swiss-Prot online tool. The isoelectric point, GRAVY value prediction website is https://web.expasy.org/compute\_pi/. Protein hydrophobicity/hydrophilicity prediction website is https://web.expasy.org/protscale/. DNAman 7.0 software was used for homologous sequence alignment. CDD (Conserved Domain Database) in CNBI was used for prediction of conservative domains. MEGA 5.1 software was used to construct phylogenetic tree.



# 2 Results and Analysis

#### 2.1 Cloning and sequence analysis of CiSAD gene from Carya illinoinensis

A target gene fragment with a length of about 1 300 bp (Figure 1) was obtained by PCR amplification. After recovery and sequencing, it was found that the sequence length was 1 240 bp. After NCBI alignment, it contained a complete open reading frame (ORF) of 1 194 bp, encoding 397 amino acids, which was named *CiSAD*. The accession number of this sequence in the GenBank database was BankIt2129876 Seq1 MH588443.



Figure 1 Pecan CiSAD and CiFAD2 gene fragment amplified by PCR

The physicochemical property of the protein was predicted by Swiss-Prot online tool. Results showed that the molecular weight of the protein was 44.99 kDa and the isoelectric point (pI) was 6.77, the grand average of hydropathicity (GRAVY) value was -0.351, indicating that the protein is a hydrophilic protein.

The alignment analysis results of amino acid sequence in *CiSAD* gene from *Carya illinoinensis* and other species showed that the *CiSAD* gene from *Carya illinoinensis* had high homology with oil plants such as *Juglans regia*, *Jatropha curca*, *Ricinus communis* and so on. Especially *Juglans regia* with the highest homology of 93%, and the homology with other plants was also greater than 80%. The CiSAD amino acid sequence deduced from *Carya illinoinensis* was subjected to multiple alignment with SAD amino acid sequences with high homology such as *Juglans regia*, *Olea europaea* var *sylvestris*, *Manihot esculenta*, *Jatropha curcas*, *Ricinus communis*, *Hevea brasiliensis*, *Citrus sinensis*, *Populus trichocarpa* and *Arabidopsis thaliana*. The results are shown in Figure 2. By comparing the SAD amino acid sequences of different species, combined with the CDD database of NCBI, it was found that the SAD protein contains two conserved domains: an acyl-ACP desaturase conserved domain with 256 amino acids from 65 to 320 and a ferritin-like family conserved domain with 129 amino acids from 144 to 272 (Figure 3). These two conserved regions are shared by acetyl-ACP desaturase and are highly conserved in plants.

The phylogenetic tree of *Carya illinoensis*, high-homologous plants and common oil crops was constructed with MEGA 5, and it was mainly divided into three branches. The first branch consisting of woody oil crops such as *Hevea brasiliensis* and *Manihot esculenta* and containing *Carya illinoensis*, the second branch consisting of *Arabidopsis thaliana* and the third branch consisting of herbaceous oil crops such as *Brassica napus* and *Helianthus annuus* (Figure 3).



#### Bioscience Method 2022, Vol.13, No.1, 1-13 http://bioscipublisher.com/index.php/bm

CSSAD PtSAD AtSAD	MQMCALHLYSS.FCTGNCALRPRGCNPTLHRLRSP.ATVRFAMTAAPAVPFTALK MINCALHLNSS.FYTENCALRRSGGCSPTLHRLQSP.TTVRFAMAAAAFTALK MQTLAFYFS.HVTTQLACTPRGLRG.ATTLRRPPP.ISSVAF MQVSLHLNN.HQLPWTVQS.RHRLRPPSPSAVRIRPLPS.ITAVAAFPTP MLAAHYLN.TWIPHT.HHHR.LTVKSSRLPS.ITAVAAFPTP MLAAHYLN.TWIPHT.HHSLQPP.QLRRLKP.ITAVAAFPTP MSAHCLSLSLSSHHNFWTQH.HSLQPP.JCLRLKP.IAVAAFPT MCASQFLVTS.TKRFAACSPLVPNFAHHRLTAAS.STVRFRQISAVAAFPTP MQVSHFLSG.HQLSCTTHQQKMILHRFSPPS.STVRFRPSDRSITVAVAAFPQ. MQVSHFLSG.FTTQWATLMP.SPSTFIASRPRGFAK.ISAVAAFVRP	54 51 44 47 38 46 65 50 52 44
CiSAD JrSAD OeSAD MeSAD JcSAD RcSAD HbSAD CsSAD PtSAD AtSAD	HQKITHSMEFEKMEVEKSIENWASQSVLFILKEVEQCWQEQNMVPDFTLFEDESSQVRELRKRTAELPIEYFVL HQKITHSMEFEKIEVEKSIENWASQSVLFILKEVEQCWQEQNIVPDFTLFEEFCEQVRELRKRTAELPIEYFVL QQVTHSMEFEKVEVEKSIESWASQCTLFILKEVEQCWQEQNFIPDFAMFEDETTQVRELRKRTAELPDEYFVL QQVTHSMEFEKVETEKSIENWAQQVLFFLKEVDICWQEQNFIPDFAMFEDETTQVRELRKRTAELPDEYFVL LKHQKTHSMEFEKVETEKSIENWAQQVLFFLKEVDICWQEQNFIPDFALFEDETTQVRELRRTAELPDEYFVL LKHQKTHSMEFEKVETEKSIENWATQVVLFILKEVDCWQEQNFIPDFALFEDETTQVRALRRTAELPDEYFVL LKHQKTHSMEFEKVETEKSIENWATQVVLFFLKEVDICWQEQNFIPDFALFEDETTQVRALRRTAELPDEYFVL LKHQKTHSMEFEKVETESSIENWATQVVLFFLKEVDLCWQEQNFIPDFALFEDETTQVRALRRTAELPDEYFVL LKHQKTHSMEFEKVETESSIENWATQVVLFFLKEVDLCWQEQNFIPDFALFEDETTQVRALRRTAELPDEYFVL LKHQKTHSMEFEKVETESSIENWATQVVLFFLKEVDLCWQEQNFIPDFALFEDETTQVRALRRTAELPDEYFVL LKHQKTHSMEFEKVETESSIENWATQVVLFFLKEVDLCWQEQNFIPDFALFEDETTQVRALRRTAELPDEYFVL LKHQKTHSMEFEKVETESSIENWATQVVLFFLKEVDCWQEQNFIPDFALFEDETTQVRALRRTAGIPDIYFVL LKHQKTHSMEFEKVETESSIENWATQVVLFILKEVECWQEQNFIPDFLFSSETTQVRALRRTAGIPDIYFVL LKHQKTHSMESEKIETKSIDNWATQVVLFILKEVDCWQEQAFFIPDFTLFSSETTQVRALRRTAGIPDIYFVL LKHQKTHSMESEKIETKSIDNWATQVVLFILKEVDCWQEQAFFIPDFTLFSSETTQVRALRRTAGIPDIYFVL	130 127 119 127 116 124 143 123 130 123
CiSAD JrSAD OeSAD MeSAD JcSAD RcSAD HbSAD CsSAD PtSAD AtSAD	VGDMITELAIPTYQTMINTLDGVRDETGASLSEWAVWTRAWTAEENRHGDLIRTYLYLSGRVDMIMTERTVQYLLGAGM VGDMITELAIPTYQTMINTLDGVRDETGASLSEWAVWTRAWTAEENRHGDLIRTYLYLSGRVDMIMTERTVQYLLGAGM VGDMITELAIPTYQTMINTLDGVRDETGASRNEWAVWTRAWTAEENRHGDLIRTYLYLSGRVDMAMTERTVQYLLGAGM VGDMITELAIPTYQTMINTLDGVRDETGASRNEWAVWTRSWTAEENRHGDLIRTYLYLSGRVDMAMTERTVQYLLGAGM VGDMITELAIPTYQTMINTLDGVRDETGASRNEWAVWTRSWTAEENRHGDLIRTYLYLSGRVDMIMTERTVQYLLGAGM VGDMITELAIPTYQTMINTLDGVRDETGASRNEWAVWTRSWTAEENRHGDLIRTYLYLSGRVDMIMTERTVQYLLGAGM VGDMITELAIPTYQTMINTLDGVRDETGASRNEWAVWTRSWTAEENRHGDLIRTYLYLSGRVDMIMTERTVQYLLGAGM VGDMITELAIPTYQTMINTLDGVRDETGASRNEWAIWTRAWTAEENRHGDLIRSFLYLSGRVDMIMTERTVQYLLGAGM VGDMITELAIPTYQTMINTLDGVRDETGASSNEWAIWTRSWTAEENRHGDLIRSFLYLSGRVDMIMTERTVQYLLGAGM VGDMITELAIPTYQTMINTLDGVRDETGASSNEWAIWTRAWTAEENRHGDLIRSFLYLSGRVDMIMTEKTVQYLLGAGM VGDMITELAIPTYQTMINTLDGVRDETGASSEWAIWNRSWTAEENRHGDLIRSFLYLSGRVDMIMTEKTVQYLLGAGM VGDMITELAIPTYQTMINTLDGVRDETGASSEWAIWNRSWTAEENRHGDLIRSFLYLSGRVDMIMTEKTVQYLLGAGM VGDMITELAIPTYQTMINTLDGVRDETGASSEWAIWNRSWTAEENRHGDLIRSFLYLSGRVDMIMTEKTVQYLLGAGM VGDMITELAIPTYQTMINTLDGVRDETGASSEWAIWNRSWTAEENRHGDLIRSFLYLSGRVDMIMTEKTVQYLLGAGMD	210 207 199 207 196 204 223 203 210 203
CiSAD JrSAD OeSAD MeSAD JcSAD RcSAD HbSAD CsSAD PtSAD AtSAD	PGTENNPYLGEVYTSFQERATEVSHGNTARIAKDGGDELLARICGTIAADEKRHENAYAKIVEKILEVDETGAMIAVADM PGTENNPYLGEVYTSFQERATEVSHGNTARIAKGDELLARICGTIAADEKRHENAYAKIVEKILEVDETGAVIAVADM PGTENNPYLGEVYTSFQERATEVSHGNTARIAKEGGDEVLARICGTIAADEKRHENAYAKIVEKILEVDTNETMIAIG PGTENNPYLGEVYTSFQERATEVSHGNTARIAKEGGDEVLARICGTIAADEKRHENAYAGIVEKIVEVDENGAMIAIAD PGTENNPYLGEVYTSFQERATEVSHGNTARIAKEGGDEVLARICGTIAADEKRHENAYAGIVEKIVEVDETGAVVAIG PGTENNPYLGEVYTSFQERATEVSHGNTARIAKEGGDEVLARICGTIAADEKRHENAYAGIVEKIVEVDETGAVVAIG PGTENNPYLGEVYTSFQERATEVSHGNTARIAKEGGDEVLARICGTIAADEKRHENAYARIVEKILEVDETGAVVAIG PGTENNPYLGEVYTSFQERATEVSHGNTARIAKEGGDEVLARICGTIAADEKRHENAYARIVEKILEVDETGAVANIADM PGTENNPYLGEVYTSFQERATEVSHGNTARIAKEGGDEVLARICGTIAADEKRHENAYARIVEKILEVDETGAVANIADM PGTENNPYLGEVYTSFQERATEVSHGNTARIAKEGGDEVLARICGTIAADEKRHENAYARIVEKILEVDETGAVANIATADM PGTENNPYLGEVYTSFQERATEVSHGNTARIAKEGGDEVLARICGTIASDEKRHENAYARIVEKILEVDETGAVANIADL PGTENNPYLGEVYTSFQERATEVSHGNTARIAKEGGDEVLARICGTIASDEKRHENAYARIVEKILEVDETGAMATADL PGTENNPYLGEVYTSFQERATEVSHGNTARIAKEGGDEVLARICGTIAZDEKRHENAYARIVEKILEVDETGAMATADL PGTENNPYLGEVYTSFQERATEVSHGNTARIAKEGGDEVLARICGTIAZDEKRHENAYARIVEKILEVDETGAMIATADM	290 285 279 287 276 284 303 283 290 283
CiSAD JrSAD OeSAD MeSAD JcSAD RcSAD HbSAD CSSAD PtSAD AtSAD	MRKKITMFAHLMEDGEDFELFEHESAVAQRLGVYTANDYADINEFLYGRWGLEKLEGLTGEGRFAQDFVCGIAFRIRKLQ MRKKITMFAHLMEDGEDFRLFEHFSAVAQRLGVYTANDYADILEFLYGRWGLEKLEGLTGEGRFAQDFVCGIAFRIRKLQ MRKKITMFAHLMYDGEDFNLFEHFSAVAQRMGVYTADDYADILEFLIGRWRLEKLEGLTGEGRFAQDFVCGIAFRIRKLQ MRKKITMFAHLMYDGEDFNLFEHFSAVAQRMGVYTADDYADILEFLIGRWRLEKLEGLTGEGRFAQDFVCGIAFRIRKLQ MRKKITMFAHLMYDGEDFNLFEHFSAVAQRMGVYTADDYADILEFLIGRWRLEKLEGLTGEGRFAQDFVCGIAFRIRKLQ MRKKITMFAHLMYDGEDFNLFEHFSAVAQRMGVYTADDYADILEFLIGRWRLEKLEGLTGEGRFAQDFVCGIAFRIRKLQ MRKKITMFAHLMYDGEDFNLFEHFSAVAQRLGVYTADDYADILEFLIGRWRLEKLEGTGEGRFAQDFVCGIAFRIRKLQ MRKKITMFAHLMYDGEDFNLFEHFSAVAQRLGVYTADDYADILEFLIGRWRLEKLEGTGEGRFAQDFVCGIAFRIRKLQ MRKKITMFAHLMYDGEDFNLFEHFSAVAQRMGVYTADDYADILEFLIGRWRLEKLEGTGEGRFAQDFVCGIAFRIRKLQ MRKKITMFAHLMYDGEDFNLFEHFSAVAQRMGVYTADDYADILEFLIGRWRLEKLEGTGEGRFAQDFVCGIAFRIRKLQ MRKKITMFAHLMYDGEDFNLFEHFSAVAQRLGVYTADDYADILEFLYGRWGLEKLEGLTGEGRFAQDFVCGIAFRIRRLQ MRKKITMFAHLMYDGEDFNLFEHFSAVAQRLGVYTADDYADILEFLYGRWRLEKLEGLTGEGRFAQDFVCGIAFRIRRLQ MRKKITMFAHLMTDGEDFNLFEHFSAVAQRLGVYTADDYADILEFLYGRWRLEKLEGLTGEGRFAQEFVCGIAFRIRRLQ MRKKITMFAHLMHDGEDFNLFEHFSAVAQRLGVYTADDYADILEFLYGRWRLEKLEGLTGEGRFAQEFVCGIAFRIRRLQ MRKKITMFAHLMHDGEDFNLFEHFSAVAQRLGVYTADDYADILEFLYGRWRLEKLEGLTGEGRFAQEFVCGIAFRIRRLQ	370 365 359 367 356 364 383 363 370 363
CiSAD JrSAD OeSAD JcSAD RCSAD HbSAD CSSAD PtSAD AtSAD	ERADEIAKKME. PRGVKFSWIFNREVK BRADEIAKKME. PRGVKFSWIFNREVK ERADDRAKKMK.AQSVKFSWIFNREVT BRADEKAKKLG.FKGVKFSWIFNREIP BRADEKAKKME.PQGVKFSWIFNREVP BRADERAKKMK.PQSAKFSWIFNREVP BIADQRAKKLK.PLJVKFSWIFNREVP BIADQRAKKLK.PLJVKFSWIFNRES BIADQRAKKKK.PLJVKFSWIFNKLS BRADEKAKKMK.FLSAFSWIFNKEVA BRADEKAKKMK.FLSAFSWIFNKEVA	396 391 385 393 382 390 409 389 396 390

Figure 2 Multiple sequence alignment of amino acid homology of CiSAD

Note: CiSAD, *Carya illinoensis*; JrSAD, *Juglans regia* (xp\_018849207.1); OeSAD, *Olea europaea* var sylvestris (XP\_022879085.1); MeSAD, *Manihot esculenta* (XP\_021603146.1); JcSAD, *Jatropha curcas* (XP\_012066083.1); RcSAD, *Ricinus communis* (NP\_001310674.1); HbSAD, *Hevea brasiliensis* (XP\_010999822.1); CsSAD, *Citrus sinensis* (XP\_006472924.1); PtSAD, *Populus trichocarpa* (XP\_002307478.2); AtSAD, *Arabidopsis thaliana* (NP\_175048.1). Red and green underlines represent the conserved domain of acyl-ACP desaturase family and ferritin-like family

#### 2.2 Cloning and sequence analysis of CiFAD2 gene from Carya illinoinensis

The amplification of the CiFAD2 sequence of *Carya illinoinensis* is shown in Figure 1. The length of the target gene is 1 329 bp, containing an ORF of 1 155 bp and encoding 384 amino acids. The molecular weight of the deduced protein is 44.30 kDa, and the isoelectric point (pI) is 8.79. The grand average of hydropathicity (GRAVY) value is -0.063, indicating that the protein is a hydrophilic protein. The accession number of this sequence in GenBank database is BankIt2133004 Seq1MH613768.





0.05

Figure 3 Phylogenetic tree constructed using CiSAD of pecan

Note: HbSAD, Hevea brasiliensis (XP\_010999822.1); MeSAD, Manihot esculenta (XP\_021603146.1); RcSAD, Ricinus communis (NP\_001310674.1); JcSAD, Jatropha curcas (XP\_012066083.1); PtSAD, Populus trichocarpa (XP\_002307478.2); CsSAD, Citrus sinensis (XP\_006472924.1); CiSAD, Carya illinoensis; JrSAD, Juglans regia (xp\_018849207.1); OeSAD, Olea europaea var sylvestris (XP\_022879085.1); AtSAD, Arabidopsis thaliana (NP\_175048.1); BnSAD, Brassica napus (CAA65990.1); HaSAD, Helianthus annuus (CAC80359.1); AhSAD, Arachis hypogaea (AAD48495.1); GmSAD, Glycine max (AAA92462.1)

The amino acid sequence of CiFAD2 from *Carya illinoensis* was aligned with BLASTp in NCBI, and it was found that CiFAD2 had the highest homology with *Juglans regia* (97%). And the homology with *Xanthoceras sorbifolium, Corylus avellana, Quercus suber, Rhus chinensis, Citrus sinensis, Vitis vinifera, Sesamum indicum, Coffea canephora* are also more than 80%. The CiFAD2 amino acid sequence of *Carya illinoensis* was compared with the FAD2 amino acid sequence of plants with high homology. The results are shown in Figure 4. By searching CDD database in NCBI combined with literature reports, it was found that the protein contains 3 histidine conserved domains, namely sequence HLLPH, HECGH and HVAHH (HXXXH or HXXHH). The conserved domain is a putative di-iron ligand shared by the membrane-FADS-like superfamily, which is the binding site of iron ions, and the active center of FAD enzyme catalysis (Shanklin and Cahoon, 1998). The phylogenetic tree was constructed using MEGA 5 (Figure 5). It was found that *Carya illinoensis* was first combined with *Juglans regia*, and then combined with *Corylus avellana*. While herbal oil crops such as *Glycine max* and *Sesamum indicum* had a far evolutionary relationship with *Carya illinoensis*.

# 2.3 Cloning and sequence analysis of CiGPAT gene from Carya illinoinensis

The target sequence was obtained by PCR amplification (Figure 6). The length of *CiGPAT* sequence from *Carya illinoinensis* is 1 671 bp. After NCBI alignment, it contained an ORF of 1 626 bp, encoding 541 amino acids. The molecular weight of the derived protein is 61.25 kDa, isoelectric point (pI) is 8.97, and the GRAVY value is 0.159, indicating that the protein is a hydrophobic protein. The accession number of this sequence in GenBank database is BankIt2133004 Seq2 MH613769.



# Bioscience Method 2022, Vol.13, No.1, 1-13

http://bioscipublisher.com/index.php/bm

CiFAD2 JrFAD2 XsFAD2 QsFAD2 RcFAD2 CcFAD2 CcFAD2 VvFAD2 CaFAD2 SiFAD2 CsFAD2	MGAGGQMSVDVVNKSEEQKATLQRVFHTQPFFTLGQLKKAIPFHCFKRSLFRSFSVAYDIFFAFTFYTLATSYFHLLFH MGAGAQMTVVNKSEEQKATLQRVFHTKPFFTLSQLKKAIPFHCFQRSLFRSFSVAYDIFFAFTFYTATSYFHLLFH MGAGGRMS.ASSSENGEKGSSIQRVFHTKPFFTLSQLKKAIPFHCFQRSLFRSFSVAYDIFFAFTFYTATSYFHLLFH MGAGGRMS.SSSENGEVGKNTLQRVFHTKPFFTLSQLKKAIPFHCFQRSLTRSFSVAYDISLAFTFYTATSYFHLLFH MGAGGRMS.SPKKNGEQKNTLQRVFHTKPFFTLSQLKKAIPFHCFQRSLTRSFSVAYDISLAFTFYTATSYFFLFY MGAGGRMS.SATKTEDKKNPLKRVFYSKPFTVGDIKKAIPFHCFGRSLTRSFSVAYDITLVSTFYTASSYFPLFY MGAGGRMS.SATTKTEDKKNPLKRVFYSKPFTVGDIKKAIPFHCFGRSTFRSFSVAYDISLAFTFYTATSYFSLLFS MGAGGRLSVFNSPVQRVFHTKPFFTLGQIKKAIPFHCFGRSTFRSFSVAYDISLAFTFYTATSYFSLLFS MGAGGRLSVFNSPVQRVFHTKPFFTLGQIKKAIPFHCFGRSTFRSFSVAYDISLAFTFYTATSYFSLLFS MGAGGPSLQREKKESLIQFAFHTKPFFTLGDIKKAIPFHCFRRSLIRSFSVVYDISLAFTFYTATSYFHLLF MGAGGPSLQREKKESLIQFAFHTKPFFTLSQIKKAIPFHCFRRSLIRSFSVVYDIVTFTIYTATSYFHLLFS MGAGGPSLQREKKESLIQFAFHTKPFFTLSQIKKAIPFHCFRRSLIRSFSVVYDIVTFTIYTATSYFHLLFS	80 78 79 78 79 79 72 78 79 75
CiFAD2 JrFAD2 XsFAD2 QsFAD2 RcFAD2 CcFAD2 VvFAD2 CaFAD2 SiFAD2 CsFAD2	PLSYFAW PIYW TIQGC II TGVWVIAHECGHHAFSDYQWVDDTVGIVLHSALIVPYFSWKISHRRHHSNTGSIERDEVFVP PLSYFAW PIYW TIQGC II TGVWVIAHECGHHAFSDYQWVDDTVGIVLHSALIVPYFSWKISHRRHHSNTGSIERDEVFVP PYAYIAW PIYW II QGC II TGVWVIAHECGHHAFSDYQWVDDTVGIVLHSALIVPYFSWKYSHRRHSNTGSIERDEVFVP PLSYFAW PIYWAFQGC II TGVWVIAHECGHHAFSDYQWVDDTVGI LHSALIVPYFSWKYSHRRHSNTG SIERDEVFVP PLYIAW PYYWFQGC VI TGVWVIAHECGHHAFSDYQWVDDTVGI LHSALIVPYFSWKYSHRRHSNTG SIERDEVFVP PIAYIAW PYYWFQGC TGVWVIAHECGHHAFSDYQWVDDTVGI LHSALIVPYFSWKYSHRRHSNTG SIERDEVFVP PYNYIAW PYYWIFQGC TGVWVIAHECGHHAFSDYQWVDDTVGI LHSALIVPYFSWKYSHRRHSNTG SIERDEVFVP PYNYIAW PYYWIFQGC TGVWVIAHECGHHAFSDYQWVDDTVGI LHSALIVPYFSWKYSHRRHSNTG SIERDEVFVP PYNYIAW PYYWIFQGC TGVWVIAHECGHHAFSDYQWDDTVGI LHSALIVPYFSWKISHRRHSNTG SIERDEVFVP PLSYIAWSIYYALQGC II TGVWVIAHECGHHAFSDYQWDDTVGI LHSALIVPYFSWKISHRRHSNTG SIERDEVFVP PLSYIAWSIYYALQGC II TGVWVIAHECGHHAFSDYQWDDTVGI LHSALIVPYFSWKISHRRHSNTG SIERDEVFVP PYCYIAW PIYWAFQGC YCTGIWVIAHECGHHAFSDYQWDDTVGI LHSALIVPYFSWKYSHRRHSNTG SIERDEVFVP PYCYIAW PIYWAFQGC YCTGIWVIAHECGHAFSDYQWDDTVGI LHSALIVPYFSWKYSHRRHSNTG SIERDEVFVP PYCYIAW PIYWAFQGC YCTGIWVIAHECGHHAFSDYQWDDTVGI LHSALIVPYFSWKYSHRRHSNTG SIERDEVFVP PYCYIAW PIYWAFQGC YCTGIWVIAHECGHHAFSDYQWDDTVGI LHSALIVPYFSWKYSHRRHSNTG SIERDEVFVP PYCYIAW PIYW IF QCC YCTGIWVIAHECGHHAFSDYQWDDTVGI LHSALIVPYFSWKYSHRRHSNTG SIERDEVFVP PYCYIAW PIYW FF YF	160 158 159 158 159 159 152 158 159 155
CiFAD2 JrFAD2 XsFAD2 QsFAD2 RcFAD2 CcFAD2 VvFAD2 CaFAD2 SiFAD2 CsFAD2	KFKSKMPWFSKYFNNEFGRVITILVTIILGWPLYIAFNVSGRPYDREACHYDFYGPIYSGRERIQIYISGAGIFAATYVL KPKTKMPWYSKYFNNEFGRVITILVTIILGWPLYIAFNVSGRPYDREACHYDFYGPIYSVRERIQIYISGAGIFAATYVL KTKAEIPWFSKYINNEFGRAUTIVITITIGWPLYIAFNVSGRPYDREACHFDPYGPIYSDRERIQIFISDTGIFATTYVL KTKSKMQWYSKYINNEFGRAITILVTITIGWPLYIAFNVSGRPYDREACHFDPHGPIYSDRERIQIYISGAGIFTATYVL KTKSKTPWFAKYFNNFFGRAITILVTITIGWPLYIAFNVSGRPYDREACHFDPHGPIYSDRERIQIYISGAGIFTATYVL KTKSKTPWFAKYFNNFFGRAITILVTITIGWPLYIAFNVSGRPYDREACHFDPHGPIYSDRERIQIYISGAGIFTATYVL KTKSKTPWFAKYFNNFFGRAITILVTITIGWPLYIAFNASGRPYDREACHFDPHGPIYSDRERIQIYISDAGIFTATYVL KFKSKMAWYSKYINNFFGRVITITITITIGWPLYIAFNASGRPYDREACHYDPYGPIYNDRERIQIYISDAGIFTAYVL KFKSKMWYSKYINNEFGRVITITITITIGWPLYIAFNASGRPYDREACHYDPYGPIYNDRERIQIYISDAGIFTASYVL KFKSKMPWFSKYFNNEFGRVITILITITIGWPLYIAFNVSGRPYDREACHYDPYGPIYSDRERIQIYISDAGIFTASYVL KFKSKMPWFSKYFNNEFGRVITILITITIGWPLYIAFNVSGRPYDREACHYDPYGPIYSNRERCQIFVSDAGVFATTYVL KFKSKMPWFSKYFNNEFGRVITILITITIGWPLYIAFNVSGRPYDREACHYDPYGPIYNDRERIQIFTSJAGIFAAVVL KFKSKPWFSKYFNNEFGRVITILITITIGWPLYIAFNVSGRPYDREACHYDPYGPIYSNRERCQIFVSDAGVFATTYVL KFKSRVFYSKYFNNEFGRVITILITITIGWPLYIAFNVSGRPYDREACHFDFYGPIYNDRERIQIFTSJAGIIAAVCVL KFKSEVFFYSKYFNNEFGRVITILVTITIGWPLYIAFNVSGRPYDREACHFDFYGPIYNDRERIQIFTSJAGIIAAVCVL KFKSEVFFYSKYFNNEFGRVITILVTITIGWPLYIAFNVSGRPYDREACHFDFYGPIYNDRERIQIFTSJAGIIAAVCVL	240 238 239 238 239 239 232 238 239 235
CiFAD2 JrFAD2 XsFAD2 QsFAD2 RcFAD2 CcFAD2 VvFAD2 CaFAD2 SiFAD2 CsFAD2	YSAAMAKGIAN LVCVYGVPLIIVN SFLVTITYLQHTHPSLPHYLSSEWDWIKGAIATADRDYGVLNKVFHNITDTHVAHH YSAAMAKGIAN LVCVYGVPLIIVNGFLVTITYLQHTHPSLPHYLSSEWDWIKGAIATADRDYGVLNKVFHNITDTHVAHH YNIAIAKGIAN LVCVYGVPLIIVNGFLVIITFLQHTHPSLPHYLSSEWDWIKGAIATVDRDYGVLNKVFHNITDTHVAHH YQVAIAKGIAN LVCYYGVPLIIVNGFLVIITYLQHTHFALPHYLSSEWDWIRGAIATADRDYGVLNKVFHNITDTHVAHH YRVAMAKGIAN LVCTYGVPLIIVNGFLVIITYLQHTHFALPHYLSSEWDWIRGAIATADRDYGVLNKVFHNITDTHVAHH YRVAMAKGIAN LVCTYGVPLIIVNGFLVIITYLQHTHFALPHYLSSEWDWIRGAIATADRDYGVLNKVFHNITDTHVAHH YRVAMAKGIAN LVCTYGVPLIIVNGFLVIITYLQHTHFALPHYLSSEWDWIRGAIATADRDYGVLNKVFHNITDTHVAHH YRVAMAKGIAN LVCTYGVPLIIVNGFLVIITYLQHTHFALPHYLSSEWDWIRGAIATADRDYGVLNKVFHNITDTHVAHH YRVAMAKGIAN LVCTYGVPLIIVNGFLVIITYLQHTHFALPHYLSSEWDWIRGAIATVDRDYGVLNKVFHNITDTHVAHH YRVALVKGIAN LVFTYGPLIIVNGFLVIITYLQHTHFALPHYLSSEWDWIRGAIATVDRDYGVLNKVFHNITDTHVAHH YRVALVKGIAN LVGTYGVPLIIVNGFLVIITYLQHTHFALPHYLSSEWDWIRGAIATVDRDYGVLNKVFHNITDTHVAHH YRVALVKGIAN LVGTYGVPLIIVNGFLVIITYLQHTHFALPHYLSSEWDWIRGAIATVDRDYGVLNKVFHNITDTHVAHH	320 318 319 318 319 319 312 318 319 315
CiFAD2 JrFAD2 XsFAD2 QsFAD2 CcFAD2 CcFAD2 VvFAD2 CaFAD2 SiFAD2 CsFAD2	LFSTMPHYHAMEATKAIKPILGEYYREDGTSFYKALWREAKECIYVEFDERAPSKGVFWYQKK LFSTMPHYHAMEATKAIKPILGEYYREDGTSFYKALWREAKECIYVEFDERSPKGVFWYQKK LFSTMPHYHAMEATKAIKPILGEYYQFDGTPFYKAIWREAKECIYVEFDERSPTKGVFWYRNK LFSTMPHYHAMEATKAIKPILGEYYQFDGTPFYKAIWREAKECIYVEFDEGTSNKGVFWYRNK LFSTMPHYHAMEATKAIKPILGEYYQFDGTPFYKAMWREAKECIYVEFDEGSPTKGVFWYRNK LFSTMPHYHAMEATKAIKPILGEYYQIDGTPFYKAMWREAKECIYVEFDEGGKGVFWYRNK LFSTMPHYHAMEATKAIKPILGEYYQIDGTPFYKAMWREAKECIYVEFDEGGKGVFWYRNK LFSTMPHYHAMEATKAIKPILGEYYQIDGTPFYKAMWREAKECIYVEFDEGG.KGVFWYRNK LFSTMPHYHAMEATKAIKPILGEYYQIDGTPFYKAWWREAKECIYVEFDEGGSDKGVFWYRNK LFSTMPHYHAMEATKAIKSILGKYYQFDGTPYYKAWWREAKECIYVEFDEGSPKGVFWYRNK LFSTMPHYHAMEATKAIKSILGKYYQFDGTPYYKAWWREAKECIYVEFDESTPLKGVFWYRNK LFSTMPHYHAMEATKAIKSILGKYYQFDGTPYYKAWWREAKECIYVEFDESTPLKGVFWYRNK LFSTMPHYHAMEATKAIKSILGKYYQFDGTPYYKAWWREAKECIYVEFDESTPLKGVFWYRNK	383 381 382 381 382 380 375 381 382 378

Figure 4 Multiple sequence alignment of amino acid homology of CiFAD2

Note: CiFAD2, Carya illinoensis; JrFAD2, Juglans regia (XP\_018834629.1); XsFAD2, Xanthoceras sorbifolium (AGO32050.1); QsFAD2, Quercus suber (XP\_023886251.1); RcFAD2, Rhus chinensis (AIC34705.1); CcFAD2, Coffea canephora (CDP17521.1); VvFAD2, Vitis vinifera (XP\_002285640.1); CaFAD2, Corylus avellana (AIT96965.1); SiFAD2, Sesamum indicum (XP\_011080227.1); CsFAD2, Citrus sinensis (XP\_006492661.1). \*, represent the conserved domain of histidines





Figure 5 Phylogenetic tree constructed using CiFAD2 gene of pecan

Note: CiFAD2, *Carya illinoensis*; JrFAD2, *Juglans regia* (XP\_018834629.1); CaFAD2, *Corylus avellana* (AIT96965.1); QsFAD2, *Quercus suber* (XP\_023886251.1); PgFAD2, *Punica granatum* (AAO37754.1); XsFAD2, *Xanthoceras sorbifolium* (AGO32050.1); RcFAD2, *Rhus chinensis* (AIC34705.1); CsFAD2, *Citrus sinensis* (XP\_006492661.1); VvFAD2, *Vitis vinifera* (XP\_002285640.1); HbFAD2, *Hevea brasiliensis* (AAY87459.1); GmFAD2, *Glycine max* (NP\_001347010.1); SiFAD2, *Sesamum indicum* (XP\_011080227.1); CcFAD2, *Coffea canephora* (CDP17521.1); OeFAD2, *Olea europaea* var *sylvestris* (XP\_022875273.1)



Figure 6 Pecan *CiGPAT* gene fragment amplified by PCR

The amino acid sequence of CiGPAT from *Carya illinoensis* was aligned with BLASTp, and it was found that CiGPAT had the highest homology with *Juglans regia* (94%). And the homology with *Quercus suber, Vitis vinifera* and other plants are between 62%~71%. The CiGPAT amino acid sequence of *Carya illinoensis* was compared with the GPAT amino acid sequence of plants with high homology. The results are shown in Figure 7. By searching CDD database in NCBI combined with literature reports, it was found that the C-terminal of the protein contained Lysophospholipid acyltransferase (LPLAT), PlsC, Acyltransferase, AGP acyltrn, PLN02177



and other domains. Among them, LPLAT is the functional domain of acyltransferase involved in TAG biosynthesis, Acyltransferase is the functional domain of acyltransferase involved in phospholipid synthesis, and Plsc is the functional domain of glycerol triphosphate acyltransferase involved in lipid synthesis (Heath and Rock, 1998; Nagiec et al., 1993). Using MEGA 5 to construct a phylogenetic tree, it was found that *Carya illinoensis* and *Juglans regia* first clustered together, then clustered with *Quercus suber*, *Populus trichocarpa* and *Populus euphratica* clustered together, and the farthest relationship was *Ziziphus jujube* and *Glycine max* (Figure 8).

CIGPAT JrGPAT QSGPAT DZGPAT RCGPAT HbGPAT PtGPAT JCGPAT PeGPAT MeGPAT	. MMSKRY FKTLELFEYC ILFRKPKYFPGGLHSNDHLYQFQCQKYSTASIALRSTLSSKTIMESVEG MSKPY FKTLELFEYC VLFRKPKYFFAGLLSNNHLNQFQCQKYRTASIALRSTLSSKTIMESVEG MSKAFE FKTLELFEYC ILFGKVRGFHTNLS. NNHATTQLLKYQKYFALALRSTLSSKTIMESVEG MSKAFE FKTLELFENR ILFGKVRGFHTNLS. NNHATTQLLKYQKYFALALRSTLSSKTL VENVEE MAKFFMKYFS FKTLEFFEFR	66 65 61 66 81 64 74 66
CIGPAT JrGPAT QSGPAT RCGPAT HbGPAT JCGPAT JCGPAT MeGPAT	ALLKSSSLFPYFMLVAFEAGSFLRAFVILLLYFFICLVSEETGLRIMVMVGFFGIKKESFRVGSSVLPKFFLEDVGLEIFNQVLR ALLKSSSLFPYFMLVAFEAGSFLRAFVILLYFFICLVSEETGLRIMVMVGFFGIKKESFRVGSSVLPKFFLEDVGLEIFNQVLR VLKSFSLFPYFMLVAFEAGSLRAFVIFTLYFLICLVSEETGLKIMVMVGFFGIKKESFRVGSSVLPKFFLEDVGLEIF ALLKSSSLFPYFMLVAFEAGSILRAFVIFTLYFFICLVGEDGIKIMVMVGFFGIKKESFRVGSSVLPKFFLEDVGLEAF LLKSSSLFPYFMLVAFEAGSILRAFVIFTLYFFVCLADEETGIKIMVMVGFFGIKKESFRVGRAVLPKFFLEDVGLEAF LLKSSSLFPYFMLVAFEAGSILRAFVIFTLYFFVCLADEETGIKIMVMVGFFGIKKESFRVGRAVLPKFFLEDVGLEAF LLKSSSLFPYFMLVAFEAGSILRAFVIFTLYFFVCLADEETGIKIMVMVGFFGIKKESFRVGRAVLPKFFLEDVGLEAF LLKSSSLFPYFMLVAFEAGSILRAFVIFTLYFFICLVGGDMGIKIMVMVGFFGIKKESFRVGRAVLPKFFLEDVGLEAF LLKSSSLFPYFMLVAFEAGSILRAFVIFTLYFFICLVGGDMGIKIMVMVGFFGIKKESFRVGRAVLPKFFLEDVGLEAF LLKSSSLFPYFMLVAFEAGSILRALVIFTLYFFICLVGGDMGIKIMVMVGFFGIKKESFRVGRAVLPKFFLEDVGLEAF LLKSSSLFPYFMLVAFEAGSILRALVIFTLYFFICLVGGDMGIKIMVMVGFFGIKKESFRVGRAVLPKFFLEDVGLEAF LLKSSSLFPYFMLVAFEAGSILRALVIFTLYFFICLVGQDMGIKIMVMVGFFGIKKESFRVGRAVLPKFFLEDVGLEAF LLKSSSLFPYFMLVAFEAGSILRALVIFTLYFFICLVGQDMGIKIMVMVGFFGIKKESFRVGRAVLPKFFLEDVGLEAF LLKSSSLFPYFMLVAFEAGSILRALVIFTLYFFICLVGQDMGIKIMVMVGFFGIKKESFRVGRAVLPKFFLEDVGLEAF LLKSSLFPYFMLVAFEAGSILRALVIFTLYFFICLVGQDMGIKIMVMVGFFGIKKESFRVGRAVLPKFFLEDVGLEAF LEDVGLEAF	151 150 148 155 150 165 148 158 150
CIGPAT JrGPAT QSGPAT DZGPAT RCGPAT HbGPAT PtGPAT JCGPAT PeGPAT MeGPAT	RGVRKVSVSE. LFIVMIES IKCYLEIDVVVGEPIKVFHGYFVGLMEEEEKNIIGLEEILRENKAVGTNIIGINGLHKVHG.H RGVRKVSVSE.LFIVMIES BLRDYLEIDVVVGSELKVFRGYFVGLMEEKKKNIIGLEEILKENKAVGTNIIGINGLYKVHD.H RGGKVAVSK.LFOVMVEEBLREYLEIDVVVGEIKVFHGYFVGLIEEKKKDMIWLEEILVKN.RVDSNMIGISGFNVSID.D KGGKVAVSD.FPQVMIES VRAYLEIDFVVGEIKVFGYFUGVMEEKKRIKAALEEIGRENLGCYIGISGFKKSLD.Q RGGRKVAVSD.FPQVMIES BLRDYLEIDCVVGRELKSVCGYFVGVMEEKNRKALEEIGRENLGCYIGISGFKSSID.D RGGRKVAVSD.FPQVMIES BLRDYLEIDCVVGRELKSVCGYFVGVMEEKNRKALEEIGRENLGCYIGISGFKSSID.D RGGRKVAVSD.FPQVMIES BLRDYLEIDCVVGRELKSVCGYFVGVMEEKNRKALSUSSID.LEKIISTQGNVVGISRFNRSID.D RGGRKVAVSD.FPQVMIES BLRDYLEIDCVVGRELKSVCGYFVGLMEEKKRKALSUSSID.LEELIG.KDSVISHDDIIGISCFNRSID RGRRVAVSD.FPQVMIES BLRDYLEIDCVVGRELKSVCGYFVGLMEEKKRKALSUSSID.LEELIG.KDSVISHDDIIGISCNSSID RGRRVAVSDHFPQVMIES BLRDYLGDCVVGRELKSCGYFUGLEEKKKMILVEEILG.KDSVISHDDIIGISQLNSSID RGRRVAVSDHFPQVMIES BLRDYLDVCVVGRELKTGGYFLGLEEKKKMILVEEILG.KDSVISHDJ.IIGISQLNSSID RGRRVAVSDHFPQVMIES BLRDYLDVCVVGRELKTGGYFLGLEEKKKMILVEEILG.KDSVISHDJ.IIGISQLNSSID RGRRVAVSDHFPQVMIES BLRDYLDVCVVGRELKTGGYFLGLEEKKKMILVLEEILG.KDSVISHDJ.IIGISQLNSSID	232 231 228 235 228 232 247 226 239 232
CiGPAT JrGPAT QsGPAT RcGPAT HbGPAT JcGPAT JcGPAT MeGPAT	HÇLFSQHCKEIYIVDKALKESWQNLPREKYPRPLIFHDGE TALIPTPLNALALFVWVPFGEVUSISELLISLSLFTDISTELLAF HÇLFSQHCKEIYIVDKALKESWQNLPREKYPRPLIFHDGE TALIPTFLNALALFVWVPFGEVUSISELFISLSLFTDVSNELNAF HÇLFSQ.CKEIYIVGKALKESWQRLPREKYPRPLIFHDGE ALPTFLLAVALFWAPFGEVUSISELFISLSLFTDISTELAF YHLFSH.CKEIYLVRKAEKSWQRLPROKYPNPLIFHDGE TAFQPTLLATIMEMWVPFGTISIETVISIETUSSISELFIS HÇLFSY.CKEIYLVRKAEKSWQRLPROKYPNPLIFHDGE TAFQPTLLATIMEMWVPFGTISIETVISIETUSSISELFIS HLFSH.CKEIYLVRKAEKSWQRLPROKYPNPLIFHDGE TAFQPTLLATIMEMWVPFGTISIETVIS HLFSH.CKEIYLVRKAEKSWQRLPROKYPNPLIFHDGE TAFQPTLATLALFWVPFGTIAVIRVVSITLFYSISTETLFY HLFSH.CKEIYLVRKAEKSWQRLPROKYPNPLIFHDGE TAFPTHATLALFWVPFGTIALIFAAVALSLFYSMSTETLTF HHLFSH.CKEIYLVRKOKSWQRLPROKYPNPLIFHDGE TAFPTFLATLALFWVPFGTIALIFAAVALSLFYSMSTETTTF HHLFSH.CKEIYLVRKOKSWQRLPROKYPNPLIFHDGE TALRPTFLATLALFWVPFGFUTALIFAAVALSLFYSMSTETTTF HHLFSH.CKEIYLVRKOKSWQRLPROKYPNPLIFHDGE TALRPTFLATLALFWVPFGFUTALIFAAVALSLFYSMSTETTTF HHLFSH.CKEIYLVRKOKSWQRLPROKYPNPLIFHDGE TALRPTFLATLALFWVPFGFUTALIFAAVALSLFYSMSTETTTF HFLFSH.CKEIYLVRKOKSWQRLPROKYPNPLIFHDGE TAFPTFLATLALFWVPFGFUTALIFAAVALSLFYSMSTETTTF	317 316 312 319 312 316 331 310 323 316
CIGPAT JrGPAT QsGPAT DzGPAT RcGPAT HbGPAT JcGPAT PeGPAT MeGPAT	LGMHLTAP.AMPKTPHSLYPNTSKEFIHKSKGVLYACNHKTILDEIFLSVILKSSSLRGVSVSISRVSDVLSPIKTVRITRRDO IGMHLTAPTSMPKTPHSLYPNTSKE.LHKSKGVLYACNHKTILDEIFLSVILKNSLRGVSVSISRVSDVLSPIKTVRITRRDO SGMYLTTMEKNSYSLS.NTEKG.KKSKGLIYVCNHKTILDEVYLVILN.TSVTGVVTSVSRISTISTLSTKTVRISRRDO SGLHLELSIPQTTQHPHSLQSSN.NRKTKGRLYVCNHRTILDELYLSFILG.KDLTAVTYSISRVSDHAPIKTVRIARRDO SGIRCTISKSGTS.SSALKSEN.KFR.GGLLYVCNHRTILDELYLSFIK.KNFTAVTYSISRVSDHAPIKTVRIARRDO SGIRCTIS.KSGTS.NTEKTGRLYVCNHRTILDELYLSFIK.KNFTAVTYSISRVSDHAPIKTVRIARRDO SGIRCTIS.KSGTS.NTEKTGRLYVCNHRTILDELYLSFIK.KNFTAVTYSISRVSDHAPIKTVRIARRDO SGIRCTIS.KSGTS.NTEKTGRLYVCNHRTILDELYLSFIK.KNFTAVTYSISRVSDHAPIKTVRIARRDO SGIRCTIS.KFTSGYSPTFYEN.KFR.KGLLYVCNHRTILDELYLSFIK.KNFTAVTYSISRVSDHAPIKTVRITRRDE TGVKVAIS.KFKLSSSVLFSSKEN.ESK.KGLLYVCNHRTILDELYLSFAK.KNFTAVTYSISRVSDHSPIRTVRITRRDE GERCTVS.EFNN.SKEN.KILIYVCNHRTILDELYLSFAK.KNFTAVTYSISRUSBILSPIRTVRITRRDE GERCTVS.KFKLSPSVLRSSKEN.ESK.KGLLYVCNHRTILDELYLSFAK.KNFTAVTYSISRUSBILSPIRTVRITRRDE GERCTVS.KFRNSRYSPTTFNDD.KFK.KGLLYVCNHRTILDELYLSFAK.KNFTAVTYSISRUSBILSPIRTVRITRR	401 400 391 400 390 396 411 387 403 396
CiGPAT JrGPAT QsGPAT RcGPAT HbGPAT JcGPAT JcGPAT MeGPAT	LARTIESLINCGEMFICPEGTTCRESYLLRESPLEADIAGE TVPVAIDSHTMLYGTTACGLKCLDEFFELMNEADSYTVCFDEK LARTIESLINKGEMVICPEGTTCREAYLLRESPLEADVAERIVVPVAIDSHTMLYGTTACGLKCLDEFFELMNEADSYTVCFDEK LGRMVESLINCGELVVCPEGTTCREPYLLRESPLEADWAERIVVPVAIDSHTMLYGTTACGLKCLDEFFELMNEADSYTUCFDEK LRKKMENLINCGELVVCPEGTTCREPYLLRESPLEADWSEITVPVANDSHVSMEYGTTACGLKCLDEFFELMNEADSYNVCMTEG LAMMMEKLISCGELVVCPEGTTCREPYLLRESPLEADWSEITVPVALNSHVSMEYGTTACGLKCLDEFFELMNEHFYTVCTIG LGKMMEKLISCGELVVCPEGTTCREPYLLRESPLEADWSEITVPVALNSHVSMEYGTTACGLKCLDEFFELMNEHFYTVCTIG LGKMMEKLISCGELVVCPEGTTCREPYLLRESPLEADWSEITVPVALNSHVSMEYGTTACGLKCLDEFFELMNEHFYTTC LGKMMEKLISCGELVVCPEGTTCREPYLLRESPLEADWSEITVPVALLAFVSMEYGTTACGLKCLDEFFELMNEFFSYTICTUEG LGKMMEKLISCGELVVCPEGTTCREPYLLRESPLEADWSEITVPVALLAFVSMEYGTTACGLKCLDEFFELMNEFFSYTICTUEG LGKMMEKLISCGELVVCPEGTTCREPYLLRESPLEADWSEITVPVALLAFVSMEYGTTACGLKCLDEFFELMNEFFSYTICTUEG LGKMMEKLISCGELVVCPEGTTCREPYLLRESPLEADWSEITVPVALLATVSMEYGTTACGLKCLDEFFELMNEFFSYTICTUEG LGKMMEKLISCGELVVCPEGTTCREPYLLRESPLEADWSEITVPVALLATVSMEYGTTACGLKCLDEFFELMNEFFSYTICTUEG LGKMMEKLISCGELVVCPEGTTCREPYLLRESPLEADWSEITVPVALLATVSMEYGTTACGLKCLDEFFELMNEFFSYTICTUEG LGKMMEKLISCGELVVCPEGTTCREPYLLRESPLEADWSEITVPVALLATVSMEYGTTACGLKCLDEFFENNEFFSYTICTUEG LGKMMEKLISCGELVVCPEGTTCREPYLLRESPLEADWSETVPVALLATVSMEYGTTACGLKCLDEFFENNEFSYTICTUEG LGKMMEKLISCGELVVCPEGTTCREPFLLRESPLEADWSENFTVPVALLATVSMEYGTTACGLKCLDEFFENNEFSYTICTUEG LGKMMEKLISCGELVVCPEGTTCREPFLLRESPLEADWSETTPVALNTVSMEYGTACGLKCLDEFFENNEFSYTICTUEG LGKMMEKLISCGELVVCPEGTTCREPFLLRESPLEADWSETVPVALLATVSMEYGTACGLKCLDEFFENNEFSYTICTUEG	486 485 476 485 475 481 496 472 488 481
CIGPAT JrGPAT QsGPAT DzGPAT RcGPAT HbGPAT PtGPAT JcGPAT MeGPAT	VSGSCTY.LNGEKSKTFVANHVQCEIGKALGFECTKLTRKDKYVILAGNEGIVSRK VSGSCTY.LNGEKSKTFVANHVQCEIGKALGFECTKLTRKDKYVILAGNEGIVCHK VSGLSKSCSTSSNGGVSKTFVANHVQSEIGKALGFECTKLTRKDKYLIVAGNEGKV VSGLSACHDHSERSRTFVANHVQSEIGKALGFECTKLTRRDKYLIVAGNEGKV VSGLSKSNTFVANHVQSEIGKIGFECTKLTRRDKYLIVAGNEGVSCFTSKF VGGLSKSNTFVANHVQSEIGKIGFECTKLTRRDKYLIVAGNEGTSSSSSFAKH VSGLSTCQDSD.KSRTFVANYVQSEIGKALSFECTKLTRRDKYLIVAGNEGTCNQR VSGLSTCQDSD.KSRTFVANHVQSEIGKALSFECTKLTRRDKYLIVAGNEGTSSSSSSSSAVK	541 540 532 541 532 544 552 525 544 545

Figure 7 Phylogenetic tree constructed using CiGPAT of pecan

Note: CiGPAT, *Carya illinoensis*; JrGPAT, *Juglans regia* (xp\_018849207.1); QsGPAT, *Quercus suber* (XP\_023927330.1); DzGPAT, *Durio zibethinus* (XP\_022769787.1); RcGPAT, *Ricinus communis* (XP\_015572336.1); HbGPAT, *Hevea brasiliensis* (XP\_021644822.1); PtGPAT, *Populus trichocarpa* (PNT50557.1); JcGPAT, *Jatropha curcas* (XP\_012082935.1); PeGPAT, *Populus euphratica* (XP\_011002733.1); MeGPAT, *Manihot esculenta* (XP\_021614154.1)





Figure 8 Phylogenetic tree constructed using CiGPAT of pecan

Note: HbGPAT, *Hevea brasiliensis* (XP\_021644822.1); MeGPAT, *Manihot esculenta* (XP\_021614154.1); JcGPAT, *Jatropha curcas* (XP\_012082935.1); RcGPAT, *Ricinus communis* (XP\_015572336.1); PtGPAT, *Populus trichocarpa* (PNT50557.1); PeGPAT, *Populus euphratica* (XP\_011002733.1); CcGPAT, *Corchorus capsularis* (OMP11935.1); DzGPAT, *Durio zibethinus* (XP\_022769787.1); TcGPAT, *Theobroma cacao* (EOX95331.1); VvGPAT, *Vitis vinifera* (XP\_002271508.1); QsGPAT, *Quercus suber* (XP\_023927330.1); CiGPAT, *Carya illinoensis*; JrGPAT, *Juglans regia* (xp\_018849207.1); ZjGPAT, *Ziziphus jujube* (XP\_015891692.1); GmGPAT, *Glycine max* (AAA92462.1)

# **3** Discussion

SAD gene plays an important role in the process of plant oil accumulation, so there are relatively many studies on SAD gene. SAD gene has been cloned from a variety of plants. The coding region of most SAD genes is  $1.170 \sim 1$ 190 bp, encoding precursor proteins containing about 390 amino acid residues (Knutzon et al., 1992; Taylor et al., 1992; Shah et al., 2000; Byfield et al., 2006; Ping et al., 2008; Zhang et al., 2008; Florin et al., 2010; Dong et al., 2012; Li et al., 2015; Zhao et al., 2015). The CiSAD gene cloned in this study was 1 194 bp in length, encoding 397 amino acids, including initiation codon ATG and termination codon TAA. The CiSAD amino acid sequence cloned in this study has high homology with other plants, especially with the same family plant Juglans regia, indicating that the SAD gene in plants is relatively conservative. Studies have shown that SAD gene mainly exists in plant chloroplast matrix, with tissue expression specificity and the highest expression level in seeds (Shanklin and Somerville, 1991). Inhibiting the expression of SAD gene in Brassica napu could increase the content of stearic acid in Brassica seeds (Knutzon et al., 1992). Similarly, silencing of the SAD gene in Gossypium could increase stearic acid content by 20% (Liu et al., 2002), while overexpression of SAD gene in Lycopersicon esculentum could increase oleic acid content by 60% (Zaborowska et al., 2002). The study of 508 maize inbred lines showed that the SAD gene had the greatest impact on the proportion of C18:0/C18:1 in seeds. One nonsynonymous single-nucleotide polymorphism in exon 3 and one 5-bp insertion/deletion in the 3' untranslated region were further shown to contribute to the natural variation in C18:0/C18:1 in maize (Han et al., 2017). In summary, SAD gene plays an important role in the accumulation of UFA in plants and can regulate the ratio of SFA to UFA. In this study, CiSAD gene from Carya illinoinensis was cloned, which laid a foundation for the regulation of oil metabolism in Carva illinoinensis in the future.



FAD2 belongs to desaturase gene family, its content and activity determine the composition and proportion of UFA in oil. At present, high oleic acid lines of almost all common oil crops have been obtained by transgenic or gene mutation techniques (Auld et al., 1992; Jung et al., 2000; Bruner et al., 2001; Schwartzbeck et al., 2001; Buhr et al., 2002; Liu et al., 2002), the acquisition of these lines was achieved by inhibiting the expression of FAD2 gene, which was the most outstanding achievement in the metabolic engineering of oil crops. Desaturase contains three highly conserved histidine-rich regions: His Box I, His Box II and His Box III. Among them, the 3 histidine conserved domains of the omega-6 fatty acid desaturase are HXXH, HXXHH and HXXHH. In this study, the analysis of the functional domain of FAD2 protein showed that the CiFAD2 amino acid sequence contained 3 histidine conserved clusters, namely HLLPH, HECGH and HVAHH, which combined with di-iron (Moche et al., 2003) and constituted the active center catalyzed by desaturase. Studies have shown that the deletion or substitution of His Box I and His Box II may lead to the decrease of enzyme activity, and the deletion or substitution of His Box III may lead to the loss of enzyme activity (Zhang, 2011). Libisch et al. (2000) recombined the amino acids in the HIS I and HIS II regions of the Borago officinalis \Delta6-fatty acid desaturase and the HIS III region of the Borago officinalis  $\Delta 8$ -fatty acid desaturase. The results showed that the recombinant enzyme lost the function of catalyzing the dehydrogenation of C18 fatty acids but could catalyze the dehydrogenation of palmitoleic acid (C16:1) and tetradecanoic acid (C I4:1).

Arabidopsis *GPAT* gene family contains 10 members, namely *AtATS* and *AtGPAT1*~9. Among them *AtGPAT9* may play a catalytic role in glycerophosphate (Gidda et al., 2009). At the same time, GPAT in other plants also plays an important role in improving seed oil content and oil quality (Chi et al., 2015; Paya-Milans et al., 2016). In this study, the results of homology analysis showed that the *GPAT* gene has species distinctive. The *GPAT* gene has a high degree of variation and low homology among different plants, while it is more conservative among plants of the same family and genus. Heath and Rock (1998) found that H and D in Motif HXXXXD structure are important functional sites for acyltransferase activity. Functional domain prediction also showed that CiGPAT not only has LPCAT1-like domain of lysophospholipid acyltransferase (LPLAT), but also has PlsC, Acyltransferase, AGP acyltrn, PLN02177 domain, indicating that CiGPAT has the functional domain of acyltransferase.

#### Authors' contributions

JXD and XJP are the experimental designers of this research. JZH is the executor of this research, completed the data analysis and manuscript writing. XMY, YXF, ZM participated in the design of the study and performed the statistical analysis. MZH, XJP, ZJY, WG, and WT guided data analysis and paper writing. JXD, XJP, and GZR conceived of the study, guided the experimental design, data analysis, writing and revision. All authors read and approved the final manuscript.

#### Acknowledgments

This study was supported by the National Natural Science Foundation of China (31901347), Jiangsu Provincial Department of Science and Technology-Special Project of Science and Technology in Northern Jiangsu (SZ-YC2019038), and Independent Scientific Research Project of Institute of Botany, Jiangsu Province and Chinese Academy of Sciences (JSPKLB202055).

#### Reference

Auld D.L., Heikkinen M.K., Erickson D.A., Sernyk J. L., and Romero J. E., 1992, Rapeseed mutants with reduced levels of polyunsaturated fatty acids and increased levels of oleic acid, Crop Science, 32: 657-662

https://doi.org/10.2135/cropsci1992.0011183X003200030016x

Bruner A.C., Jung S., Abbott A.G., and Powell G.L., 2001, The naturally occurring high oleate oil character in some peanut varieties results from reduced oleoyl-PC desaturase activity from mutation of asparate 150 to asparagine, Crop Science, 41: 522-526 <a href="https://doi.org/10.2135/cropsci2001.412522x">https://doi.org/10.2135/cropsci2001.412522x</a>

Buhr T., Sato S., Ebrahim F., Xing A., Zhou Y., Mathiesen M., Schweiger B., Kinney A., Staswick P., and Clemente T., 2002, Ribozyme termination of RNA transcripts down-regulate seed fatty acid genes in transgenic soybean, Plant Journal, 30: 155-163
<u>https://doi.org/10.1046/j.1365-313X.2002.01283.x</u>
PMid:12000452

Byfield G.E., Xue H., Upchurch R.G., 2006, Two genes from soybean encoding soluble Δ<sup>9-</sup>stearoyl-ACP desaturase, Crop Science, 46: 840-846 https://doi.org/10.2135/cropsci2005.06-0172

Chi X., Yang Q., Pan L., et al., 2015, Isolation and expres-sion analysis of glycerol-3-phosphate acyltransferasegenes from peanuts (*Arachis hypogaea L.*), GrasasYAceites, 66(3): e09

https://doi.org/10.3989/gya.1190142



- Dong J., Wan Y., and Liu F., 2012, Sequence analysis of ∆9-stearoyl-ACP desaturase gene (SAD) in peanut, Acta Agronomica Sinica, 38(7): 1167-1177 https://doi.org/10.3724/SP.J.1006.2012.01167
- Florin S., Yael B., Arnon B., Ilan H., and Ran H., 2010, Identification and molecular characterization of homeologous Δ<sup>9</sup>-stearoyl-acyl carrier protein desaturase3 genes from the allotetraploid peanut (*Arachis hypogaea*), Plant Molecular Biology Reporter, 29: 232-241 <u>https://doi.org/10.1007/s11105-010-0226-9</u>
- Georgios B., Anastassios M., Nikos N., and Polydefkis H., 2005, Spatial and temporal expressions of two distinct oleate desaturase from olive (*Olea europaea* L.), Plant Science, 168: 457-555

https://doi.org/10.1016/j.plantsci.2004.09.026

Gidda S.K., Shockey J.M., Rothstein S.J., Dyer J.M., and Mullen R.T., 2009, Arabidopsis thaliana GPAT8 and GPAT9 are localized to the ER and possess distinct ER retrieval signals: Functional divergence of the dilysine ER retrieval motif in plant cells, Plant Physiology Biochemistry, 47: 867-879 <u>https://doi.org/10.1016/j.plaphy.2009.05.008</u>

PMid:19539490

Han Y., Xu G., Du H., Hu J., Liu Z., Li H., Li J., and Yang X., 2017, Natural variations in stearoyl-acp desaturase genes affect the conversion of stearic to oleic acid in maize kernel, Theoretical Applied Genetics, 130(1): 151-161 <u>https://doi.org/10.1007/s00122-016-2800-5</u>

PMid:27717956

Heath R.J., and Rock C.O., 1998, A conserved histidine is essential for glycerolipid acyltransferase catalysis, Bacteriol, 180(6): 1425-1430 https://doi.org/10.1128/JB.180.6.1425-1430.1998

PMid:9515909 PMCid:PMC107040

- Hong D.D., 2007, AFLP and SSR analysis of pollinated generation between *Carya cathayensis* and C. *illinoinensis*, Zhejiang Forestry University, pp.36-39
- Jia X., Li M., Luo H., et al., 2018, Transcriptome survey reveals candidate genes involved in lipid metabolism of *Carya illinoinensis*, Int J Agric Biol, 20(5): 991-1004
- Jung S., Swift D., Sengoku E., Patel M., Teulé F., Powell G., Moore K., and Abbott A., 2000, The high oleate trait in the cultivated peanut (*Arachis hypogaea* L.). I. Isolation and characterization of two genes encoding microsomal oleoyl-PC desaturases, Molecular General Genetics, 263(5): 796-805 <u>https://doi.org/10.1007/s004380000244</u>

PMid:10905347

Kargiotidou A., Deli D., Galanopoulou D., Tsaftaris A., and Farmaki T., 2008, Low temperature and light regulate delta 12 fatty acid desaturases (FAD2) at a transcriptional level in cotton (*Gossypium hirsutum*), Journal of Experiment Botany, 2008, 59: 2043 – 2056 <u>https://doi.org/10.1093/jxb/ern065</u>

PMid:18453533 PMCid:PMC2413273

Knutzon D.S., Thompson G.A., Radke S.E., Johnson W.B., Knauf V.C., and Kridl J.C., 1992, Modification of a Brassica seed oil by antisense expression of a stearoyl-acyl carrier protein desaturase gene, Proceeding of the National Academy of Sciences of the United States of America, 89(7): 2624-2628 <u>https://doi.org/10.1073/pnas.89.7.2624</u>

PMid:1557366 PMCid:PMC48714

Li F., Bian C., Xu J., Pang W., Liu J., Duan S., Lei Z., Jiwan P., Jin L., 2015, Cloning and functional characterization of sad genes in potato, PloS One, 10(3): e0122036

https://doi.org/10.1371/journal.pone.0122036

PMid:25825911 PMCid:PMC4380360

Li H.H., Flachowsky H., Fischer T.C., Hanke M., Forkmann G., Treutter D., Schwab W., HoVmann T., and Szankowski I., 2007, Maize *Lc* transcription factor enhances biosynthesis of anthocyanins, distinct proanthocyanidins and phenylpropanoids in apple (*Malus domestica* Borkh), Planta, 226: 1243-1254 <u>https://doi.org/10.1007/s00425-007-0573-4</u>

PMid:17618453

Libisch B., Michaelson L.V., Lewis M.J., et al., 2000, Chimeras of delta 6- fatty acid and delta8- sphingolipid desaturases, Biochem Biophys Res Commun https://doi.org/10.1006/bbrc.2000.4023

PMid:11162428

Liu Q., Singh S.P., and Green A.G., 2002, High-stearic and high-oleic cottonseed oils produced by hairpin RNA-mediated post-transcriptional gene silencing, Plant Physiology, 129(4): 1732-1743

https://doi.org/10.1104/pp.001933

PMid:12177486 PMCid:PMC166761

Moche M., Shanklin J., Ghoshal A., and Lindqvist Y., 2003, Azide and acetate complexes plus two iron-depleted crystal structures of the di-iron enzyme deha9 stearoyl-acyl carrier protein desaturase: Im- plications for oxygen activation and catalytic intermediates, Journal of Biological Chemistry, 278(27): 25072-25080

https://doi.org/10.1074/jbc.M301662200

PMid:12704186

Nagiec M.M., Wells G.B., and Lester R.L., 1993, A suppressor gene that enables Saccharomyces cerevisiae to grow without making sphingolipids encodes a protein that resembles an *Escherichia coli* fatty acyltransferase, The Journal of Biological Chemistry, 268(29): 22156 – 22163 <u>https://doi.org/10.1016/S0021-9258(20)80661-9</u>



Nishida I., Tasaka Y., Shiraishi H., and Murata N., 1993, The gene and the RNA for the precursor to the plastid-located glycerol-3-phosphate acyltransferase of Arabidopsis thaliana, Plant Molecular Biology, 21: 267-277 https://doi.org/10.1007/BF00019943 PMid:7678766 Paya-Milans M., Aznar-Moreno J.A., Balbuena T.S., et al., 2016, Sunflower HaGPAT 9-1 is the predominant GPAT during seed development, Plant Science, 252: 42-52 https://doi.org/10.1016/j.plantsci.2016.07.002 PMid:27717477 Ping Z., Joseph W.B., Robert G.U., Edward W., John S., and Ralph E.D., 2008, Mutations in a  $\Delta^9$ -stearoyl-ACP- desaturase gene are associated with enhanced stearic acid levels in soybean seeds. Crop Science, 48: 2305-2313 https://doi.org/10.2135/cropsci2008.02.0084 Schwartzbeck J.L., Jung S., Abbott A.G., Mosley E., Lewis S., Pries G.L., and Powellg L., 2001, Endoplasmic oleoyl-PC desaturase references the second double bond, Phytochemistry, 57(5): 643-652 https://doi.org/10.1016/S0031-9422(01)00081-4 Shah F.H., Rashid O., and San C.T., 2000, Temporal regulation of two isoforms of cDNA clones encoding delta 9-stearoyl-ACP desaturase from oil palm (Elaies guineensis), Plant Science, 152: 27-33 https://doi.org/10.1016/S0168-9452(99)00209-5 Shanklin J., and Cahoon E.B., 1998, Desaturation and related modifications of fatty acids, Annual Review of Plant Physiology and Plant Molecular Biology, 49: 611-641 https://doi.org/10.1146/annurev.arplant.49.1.611 PMid:15012248 Shanklin J., and Somerville C., 1991, Stearoyl-acyl-acyl-acrier-protein desaturase from higher plants is structurally unrelated to the animal and fungal homologs, Proceeding of the National Academy of Sciences of the United States of America, 88: 2510-2514 https://doi.org/10.1073/pnas.88.6.2510 PMid:2006187 PMCid:PMC51262 Shockey J., Regmi A., Cotton K., Adhikari N., and Browse J., 2016, Bates P D. Identification of Arabidopsis GPAT9 (At5g60620) as an essential gene involved in triacylglycerol biosynthesis, Plant Physiology, 170(1): 163-179 https://doi.org/10.1104/pp.15.01563 PMid:26586834 PMCid:PMC4704598 Taylor M.A., Smith S.B., Davies H.V., Buth L.R., 1992, The primary structures of the precursor of a cDNA clone of the stearoyl-acyl carrier protein desaturase gene from potato (Solanum tuberosum L.), Plant Physiology, 100: 533-534 https://doi.org/10.1104/pp.100.1.533 PMid:16652995 PMCid:PMC1075583 Wang M., Li X.Z., Ning D.L., Zhang Y., and Li Y.J., 2009, The research summary and developmental trend of Carya illinoensis, Forest Inventory and Planning, 34(6): 93-95 Wang M., Ning D.L., Li X.Z., Zhang Y., and Li Y.J., 2010, The survey research and development trends of Carya illinoensis, Forest By-Product and Speciality in China, 105(2): 84-86 Xiong X.H., Guan C.Y., Li X., Wang X.J., Zhou X.Y., and Li J.Y., 2002, Cloning of a sequence encoding fad2 from Bassica napus and constructing of antisence fad2 expression vector, Chinese Journal of Oil Crop Science, 24(2): 1-4 Zaborowska Z., Starzycki M., Femiak I., Swiderski M., and Legocki A.B., 2002, Yellow lupine gene encoding stearoyl-ACP desaturase-organization, expression and potential application, Acta Biochimica Polonica, 49: 29-42 https://doi.org/10.18388/abp.2002 3818 PMid:12136953 Zhang D., Tan X., and Hu F., 2008, The cDNA cloning and characteristic of stearoyl-ACP desaturase gene of Camellia Oleifera, Acta Horticulturae, 769(769): 55-61 https://doi.org/10.17660/ActaHortic.2008.769.5 Zhang D.Q., Tan X.F., Cheng H.P., Zeng Y.L., Jiang Y., Li W., and Hu F.M., 2008, Full-length cDNA cloning and bioinformatic analysis of Camellia oleifera SAD, Scientia Silvae Science, 44(2): 155-159 Zhang Y., 2011, Clone and functional verification of Δ9 - and Δ15 -fatty acid desaturase gene from Lipomyces kononenkoae, Wuhan: Huazhong Agricultural University Zhao N., Zhang Y., Li Q., Li R., Xia X., Qin X., and Guo H., 2015, Identification and expression of a stearoyl-ACP desaturase gene responsible for oleic acid accumulation in Xanthoceras sorbifolia seeds, Plant Physiology Biochemistry, 87(5): 9-16 https://doi.org/10.1016/j.plaphy.2014.12.009

PMid:25528221

Zhao N., Zhang Y., Li Q.Q., Li R.F., and Guo H.H., 2015, Sequence and functional analysis of *FAD2* gene from *Xanthoceras sorbifolia* seeds, Journal of Beijing Forestry University, 37(2): 87-93

https://doi.org/10.1016/j.plaphy.2014.12.009