Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Identification and Analysis of the FAD Gene Family in Walnuts (Juglans regia L.) Based on Transcriptome Data

Kai Liu
Hebei Agricultural University
Shugang Zhao (zshug@hebau.edu.cn)
Hebei Agricultural University https://orcid.org/0000-0002-8786-0382
Shuang Wang
Hebei Agricultural University
Hongxia Wang
Hebei Agricultural University
Zhihua Zhang
Hebei Agricultural University

Research article

Keywords: Walnut, fatty acid, RNA-Seq, FAD gene family, FAD3

Posted Date: November 21st, 2019

DOI: https://doi.org/10.21203/rs.2.17547/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Version of Record: A version of this preprint was published on April 15th, 2020. See the published version at https://doi.org/10.1186/s12864-020-6692-z.

Abstract

Background Walnut fatty acids, the main component of walnut kernels, contain a large amount of unsaturated fatty acids, such as linoleic acid and linolenic acid, which are essential fatty acids in humans and have important effects on human growth and health. Fatty acid desaturase (FAD) is widely distributed throughout the biological world. Its main function is to remove hydrogen from carbon chains in the biosynthesis of unsaturated fatty acids to synthesize C=C double bonds.

Results In the current research, 25 members of the JrFAD gene family were identified by bioinformatics analysis; the expression of fatty acid synthesis genes in walnut kernels at different developmental stages was analysed by transcriptome sequencing, and the expression of JrFAD3-1 , an enzyme gene for linolenic acid synthesis, was particularly prominent. The results showed that the relative expression level of FAD3-1 changed dramatically with the kernel development stage, and the expression changes showed a "bell shape". There was a significant positive correlation between the expression of JrFAD3-1 from 90-100 days after anthesis and the content of alpha-linolenic acid from 100-130 days after anthesis, with a correlation coefficient of 0.991. JrFAD3-1 can be considered closely related to Betula pendula and Corylus heterophylla .

Conclusion 25 walnut kernels FAD genes were identified and comprehensive analyzed for the first time. The function of a walnut kernels FAD3-1 gene was also characterized from its location in the phylogeny. This work lays a theoretical foundation for the regulation of unsaturated fatty acid synthesis and provide techniques and methods for the creation of new germplasm.

1 Background

Walnut oil not only contains linoleic acid, linolenic acid, palmitic acid, stearic acid, oleic acid and other essential fatty acids [1], but the walnut kernel oil content can be as high as 63%, higher than that of other leading oil crops. In addition, the unsaturated fatty acid is as high as 90% [2] of walnut kernel oil, which can not only inhibit the absorption of cholesterol in the small intestine but also inhibit the re-absorption of bile acid, promote the degradation and elimination of cholesterol in the liver, change the distribution of cholesterol in the body, and accelerate the transfer of plasma cholesterol to vascular tissues [3]. In addition, polyunsaturated fatty acids are also prodrugs of prostaglandins, which have blood pressure lowering, platelet adhesion reducing, and anticoagulant effects [3]. The body cannot synthesize linoleic acid and α-linolenic acid, which can only be obtained by ingesting food [4]. If linoleic acid and α-linolenic acid are lacking, cholesterol is esterified with saturated fatty acids, causing it to accumulate in arterial blood vessels and the cholesterol metabolism is reduced; then, the inner wall of the arteries are hardened, leading to atherosclerosis [5].

Fatty acid desaturase (FAD) is widely present in the biological world [6-12]. Its main function is to remove hydrogen from carbon chains in the biosynthesis of unsaturated fatty acids to synthesize C=C double bonds. In plants, FADs can be classified into two categories according to their solubility, soluble

desaturases and membrane-integrin desaturases [13]. The stearoyl-ACP-desaturase (SAD) unique to higher plants is the only known FAD dissolved in a matrix in the biological world. All other types of desaturases belong to the membrane-integrins and are localized to the endoplasmic reticulum (FAD2 and FAD3) and plastids (FAD6, FAD7 and FAD8, etc.). The first step in the synthesis of unsaturated fatty acids is the introduction of a double bond at the Δ9 position [14], and the enzymes that catalyse the reaction include stearoyl-CoA-desaturase (SCD) [15] and SAD[16]. SAD, catalysed by ferredoxin, catalyses the stearoyl carrier protein (stearoyl-ACP) to remove two hydrogens to form oleoyl ACP [17]. The second unsaturated bond is introduced by the Δ12 desaturases based on the introduction of the first unsaturated bond, including both FAD2 and FAD6 [18]. Δ15 desaturases, including FAD3, FAD7 and FAD8, play a key role in further decreasing the degree of saturation [19-22]. Δ15 saturated enzyme and coenzyme common catalytic and oleoyl glyceride to alpha linolenic acid glyceride transformation. The enzyme encoded by the *FAD3* gene is a member of the ω-3 fatty acid dehydrogenase family, a membrane integrin distributed in the endoplasmic reticulum that catalyses the introduction of a third unsaturated bond. The FAD3 amino acid sequence is highly conserved with three histidine-rich conserved domains, which together with the divalent iron ions constitute the reaction centre of fatty acid dehydrogenase [23]. The Lys-Lys-X-X (KKXX) recovery signal contained at the C-terminus of the amino acid sequence is consistent with the properties of the endoplasmic reticulum integrin [23]. Since the first *FAD3* was isolated from *Arabidopsis* [24], researchers have successively cloned genes from plants such as rapeseed, perilla, peony and *Eucommia ulmoides*. In recent years, the functional regulation of *FAD3* in the synthesis and metabolism of unsaturated fatty acids has become a research hotspot.

Walnut (*Juglans regia* L.) is an important economic tree species throughout the world and is widely cultivated worldwide as a traditional wood and oil crop. Walnut plays an important role in the national economy and people's lives. Fatty acids, which are the main nutritional component accumulated in walnut kernel, contain a large amount of unsaturated fatty acids, such as linoleic acid and linolenic acid, which are essential fatty acids for human beings and have an important role in human growth and health [2]. With the release of walnut genome data, genes involved in the metabolism of walnut nutrients have been gradually discovered. A number of key enzyme genes involved in walnut fatty acid metabolism, such as ACCase, have been gradually cloned and characterized. In recent years, the cultivation area of walnuts has expanded each year, and the production of nuts has risen sharply. In particular, the development and utilization of walnut oil has become a hot topic. In this research, walnut kernels at different stages of maturity were used as test materials and to analyse the key gene family involved in unsaturated fatty acid synthesis based on bioinformatics analysis and transcriptome sequencing. This study explored the formation mechanism of walnut fatty acids, especially the anabolic mechanism of unsaturated fatty acids, and provides a theoretical basis for the regulation of walnut fatty acid synthesis. Finally, this study can also provide a theoretical reference for the rational development of walnut oil and oil-producing walnut varieties.

2 Results

2.1 **Identification and analysis of FAD genes in the walnut genome**

We identified the FAD family using *Arabidopsis* FAD family protein sequence information to construct the hidden Markov model. The walnut protein data were searched, and 33 FAD family genes were screened in walnut. Using Pfam domain analysis, 25 FAD family genes were finally obtained, encoding 30 protein sequences, which were numbered according to their annotation information in walnut (Table 1).

Because of the limited depth of analysis of walnut genome data, we do not know the specific distribution of these genes in the 16 pairs of chromosomes in walnut. The shortest length among these *FAD* genes is 1262 bp, and the longest length is 14280 bp. The CDS lengths of these genes ranges from 453 bp-1368 bp, the lengths of FAD protein translation range from 150 to 455 amino acids, and their MW (molecular weight) range from 32.48 kDa (JrSLD-2) to 52.06 kDa (JrFAD7).

The length of *JrFAD3-1* was 2402 bp, its coding region was 1143 bp, and it encoded 380 amino acids, including 8 exons and 7 introns.

Table 1 Basic information of the FAD family in walnut

Gene	Gene ID	Gene				Amino acid		
		Gene length (bp)	Exon number	Intron number	CDS length (bp)	Protein length (AA)	PI	MW (KDA)
JrADS3	108998330	2497	5	4	1164	387	9.95	44.56
JrALD	109021249	2055	1	0	1344	447	8.56	51.34
JrDAL	109001694	4240	2	1	1149	382	8.78	44.11
JrDES1-1	109011517	2103	2	1	981	326	7.27	37.91
JrDES1-2	108990634	2623	2	1	978	325	7.83	37.74
JrFAD2-1	109021930	3411	3	2	1167	388	8.72	44.38
JrFAD2-2	109021930	3411	3	2	1167	388	8.72	44.38
JrFAD2-3	109011954	6737	3	2	1152	383	8.83	43.90
JrFAD2-4	109011954	6737	3	2	1152	383	8.83	43.90
JrFAD3-1	109002248	2402	8	7	1143	380	8.26	43.88
JrFAD3-2	108989905	2166	8	7	1053	350	7.79	40.51
JrFAD3-3	108989905	2166	8	7	1143	380	7.37	43.61
JrFAD3-4	108989903	2129	8	7	1077	358	7.8	41.43
JrFAD3-5	108989903	2129	8	7	1167	388	7.39	44.5
JrFAD4	108983937	1262	1	0	924	307	9.11	34.68
JrFAD6	108993197	4917	10	9	1329	442	9.18	51.38
JrFAD7	109007160	3038	8	7	1368	455	8.94	52.06
JrFAD8-1	108994930	2963	8	7	1368	455	8.98	51.77
JrFAD8-2	109015144	2965	8	7	1368	455	8.98	51.77

JrSAD-1	109005061	5401	3	2	1191	396	6.44	45.26
JrSAD-2	108984606	5286	4	3	1191	396	6.39	5.16
JrSAD-3	108984606	5286	4	3	1074	357	5.72	40.99
JrSAD-4	108984606	5286	4	3	999	332	5.4	38.05
JrSAD-5	108988959	2413	4	3	1176	391	7.72	44.43
JrSAD6	109012153	1549	2	1	1185	394	5.98	44.63
JrSAD-6	108989429	2075	3	2	1206	401	8.71	45.49
JrSAD-7	108988967	1981	3	2	1170	389	8.26	44.32
JrSAD-8	109018931	14280	4	3	1002	333	6.16	38.08
JrSLD-1	109021451	1905	1	0	1344	447	6.64	32.64
JrSLD-2	108988023	1846	1	0	1344	447	6.42	32.48

2.2 **Analyses of the evolution, exon-intron structure and motif distribution of JrFAD family members**

Using MEGA 5.0 software with the maximum likelihood method, the walnut FAD protein sequence was constructed together with the *Arabidopsis* FAD protein sequence to construct a phylogenetic tree (Fig. 1), indicating that the FAD gene families of *Juglans regia* and *Arabidopsis thaliana* are similar. There are four main subfamilies: the SAD desaturase subfamily, Δ7/Δ9 desaturase subfamily, Δ12/ω3 desaturase subfamily and "front-end" desaturase subfamily.

So far, SAD is the only subfamily of soluble enzymes found in the FAD family, and the remaining types of fatty acid desaturases are membrane integrins [25-26]. There are 7 copies of the *SAD* gene in *Arabidopsis*, and 9 in walnut, and these genes are well clustered in the unified phylogenetic tree. In the Δ12/ω3 desaturase subfamily, six branches of the walnut ω-6 desaturase and two from *Arabidopsis* were grouped together as Δ12 fatty acid desaturases; three *Arabidopsis FAD3* genes and eight annotated ω-3 desaturase genes in walnuts together formed the Δ12/ω3 desaturase subfamily; in the five *FAD3* genes of walnut, *FAD3-1* was far from the other four evolutionary distances. By comparing the gene expression levels during the development of walnut kernels, it was found that the expression of *FAD3-1* only was detected in the developing kernels, and the other four genes were not expressed or were expressed at very low levels, indicating that *FAD3-1* is a key enzyme gene that catalyses the synthesis of linolenic acid in walnut kernels.

To further investigate the structural evolution of walnut FADs, we first analysed the exon intron structure (Fig. 2C). Analysis showed that the FAD genes did not contain introns, except for *JrALD, JrSLD-1, JrSLD-2* and *JrFAD4,* which had from 1-9 introns. Genes in the same subfamily have similar intron and exon structures. *FAD6* differs from other Δ12 desaturase genes in that it has the largest number of introns and exons. Both *DESI* genes had two exons, while the same group of *ALD* and *SLD* genes had only one exon. The number of exons in the SAD subfamily ranged between 2 and 4. *JrSAD-8* contained 4 exons, the highest amount, while *SAD6*, *SAD-3* and *SAD-4* contained two exons, and the remaining genes contained three exons.

Subsequently, we used MEME software to analyse the conservation of these protein sequences and structures (Fig. 2B). Most of the 20 conserved motifs (Fig. 2D) found belonged to the typical transmembrane helix region and unknown functional complex region in fatty acid desaturases. No common conserved motifs were found among the 30 members of the walnut FAD family, but the distribution of conserved motifs was very similar within different subfamilies. In the SAD subfamily, all genes except *JrSAD-8* contain conserved motifs 1, 3, 4, 7, and 10; in the Δ12/ω3 subfamily, except for *JrFAD6*, the genes contain conserved motifs 2, 4, 5, 6, 8, and 9.

2.3 **Expression of genes related to unsaturated fatty acid synthesis**

The expression of genes related to unsaturated fatty acid biosynthesis and arachidonic acid metabolism peaked at 110 d after anthesis based on transcriptome sequencing. The genes related to alpha-linolenic acid metabolism were first decreased and then increased and reached their maximum value at 110 days after anthesis (Supplementary Fig. 1).

Seventeen genes enriched in the metabolic pathway of unsaturated fatty acid biosynthesis were selected (Fig. 3). *KCS2* (encoding 3-ketoacyl-CoA thiolase), *fabG* (encoding 3-oxoacyl-ACP reductase) are involved in carbon chain elongation; *fadE* (encoding acyl-ACP desaturase), *SSI2* (encoding acyl-ACP desaturase), *FAD2* (encoding fatty acid desaturase 2), *FAD3-1*(encoding fatty acid desaturase 3), *Acox1*(encoding peroxisomal acyl-CoA oxidase) and *SAD* (encoding stearoyl-ACP-desaturase)are involved in the desaturation process. The expression of 17 transcripts of these 8 genes can be found to be roughly divided into two categories. The expression level was lower at 70 d after flowering and higher at 110 d after flowering. There were 7 transcripts of *fabG, KCS2, fadE, FAD2* and *FAD3-1*, which were mainly involved in the biosynthesis of linoleic acid, linolenic acid and isounsaturated fatty acids. The other general trend is that the expression level was relatively high at 70 d after flowering and then gradually decreased. 10 transcripts of *fabG*, *KCS2*, *SAD*, *SSI2* and *FAD2* were mainly involved in the biosynthesis of oleic acid and linoleic acid. The expression of *FAD3-1* increased rapidly during the period from 70 d to 110 d after flowering. The dehydrogenation of linoleic acid to α- linolenic acid began 70 d after flowering while the α-linolenic acid content was almost zero, but the genes encodeing enzyme that can catalyze the dehydrogenase of linoleic acid to linolenic acid were highly expressed at this time. With the rapid increase in the expression of *FAD3-1*, the content of α-linolenic acid in the kernel began to increase gradually.

2.4 **Tissue-specific expression of FAD family genes**

By semi-quantitative detection (Fig. 4), *SLD-1* was found to be expressed in all 8 tissues, among which the expression levels were higher in catkins, old branches, mature leaves and kernels; *DAL* was most highly expressed in young and mature leaves. *JrFAD3-1* was expressed in catkins, young leaves and kernels, and the highest expression level was observed in mature embryos; *ADS3*, *SAD-2* and *SAD6* were only expressed in mature embryos. It can be preliminarily concluded that the FAD family is characterized by the desaturase subfamily and that the Δ7/Δ9 desaturase subfamily is specifically expressed in the embryo, while the "front-end" desaturase subfamily is expressed in all tissues. The Δ12/ω3 desaturase subfamily is highly expressed in mature embryos. However, the determination of more specific expression patterns still requires further research and verification.

2.5 **Expression of JrFAD3-1 and accumulation of α-linolenic acid in kernels at different developmental stages**

The relative expression of *JrFAD3-1* increased slowly at 70 days after flowering in the ‘QingXiang’ kernel, and the relative expression increased rapidly after 90 days after flowering, peaking at 100 days after flowering (Fig. 5). Then, the expression quickly decreased, and it gradually stabilized at a lower level after 120 days. We also detected the content of α-linolenic acid in identical samples and found that the content of α-linolenic acid was maintained at a low level from 70 days to 95 days after flowering and gradually accumulated from 95 days to 120 days after flowering. It increased rapidly from 120 days to 130 days after flowering and peaked at 40.92 mg/g (Fig. 5), after which it decreased slightly. Combining the two results showed a 30-day difference between peak *JrFAD3-1* gene expression and peak of α-linolenic acid levels. The expression of *JrFAD3-1* at 70-100 days and 100-130 days was analysed. There was a significant positive correlation between the expression of *JrFAD3-1* and the content of α-linolenic acid at 5 periods (from 90-100 days after anthesis) and the content of α-linolenic acid at 5 periods (from 100-130 days after anthesis), with a correlation coefficient of 0.991. Therefore, it is possible that the decrease in *JrFAD3-1* expression and the increase in water content in the later stage are factors leading to the decline in the α-linolenic acid content in the later stage.

2.6 **Multiple sequence alignment of JrFAD3-1 proteins**

Multiple sequence alignment of the FAD3 proteins in 32 different species was performed using MEGA 5.0. The results show that FAD3 protein sequence was conserved between monocotyledonous and dicotyledonous plants and yeast (Fig. 7). Conservative domain prediction was performed by a database search of the Pfam protein domain family, and all FAD3 proteins contained the fatty acid desaturase motif among 32 different species (Fig. 6). Through comparative analysis, 62 completely conserved regions in the 32 protein sequences were found. It can be further confirmed that the primary structure of the JrFAD3-1 protein is closely related to the content of α-linolenic acid.

2.7 **Phylogenetic evolution of FAD3 proteins**

The retrieved homologous proteins were multiplexed with the JrFAD3-1 protein with MEGA 5.0 (Fig. 8), and the reported *Glycine max* (Genbank: ABP01320.1; ABV00679.1; ACD56667.1), *Phaseolus lunatus*

(Genbank: ADP07952.1), *Vigna unguiculata* (Genbank: ABY60736.1), *Betula pendula* (Genbank: AAN17504.1), *Corylus heterophylla* (Genbank: AEF80000.1), *Crepis alpina* (Genbank: ABA 55806.1), *Solanum lycopersicum* (Genbank: ABX24525. 1), *Eucommia ulmoides* (Genbank: ARQ20744.1), *Linum usitatissimum* (Genbank: AQT18955.1), *Jatropha curcas* (Genbank: ABX82798.1), *Vernicia fordii* (Genbank: AAC98967.2), *Picea abies* (Genbank: CAC18722.1), *Paeonia suffruticosa* (Genbank: AVZ47050.1), *Salvia hispanica* (Genbank: ARE67897.1), *Perilla citriodora* (Genbank: AQZ42316.1), *Perilla frutescens* (Genbank: AAD15744.1), *Physaria fendleri* (Genbank: AWX65625.1; AWX65624.1), *Orychophragmus violaceus* (Genbank: AWY93817.1), *Chorispora bungeana* (Geneb Ank: AKN35208.1), *Descurainia sophia* (Genbank: 104815269), *Brassica juncea* (Genbank: ADJ58020.1), *Brassica oleracea* (Genbank: AGH20189.1), *Brassica napus* (Genbank: AAT09135.1) *Sinapis alba* (Genbank: AGH20189.1), *Oryza sativa* indica group (Genbank: BAA28358.1), *Triticum aestivum* (Genbank: BAA28358.1), *Candida tropicalis* (Genbank: ADN42964.1), and *Arabidopsis thaliana* (Genbank: NP_180559.1) all have a homology FAD proteins. The domains all have conserved amino acid residue sequences. Thirty-one plant-derived FAD3 proteins are grouped into one group, and the outer group of *Candida tropicalis* is a single group; JrFAD3-1 is closely related to *Betula pendula* and *Corylus heterophylla* at 100% confidence. This further indicates that the conserved structure of FADs is consistent with classical plant architecture.

3 Discussion

To date, the FAD gene family has been identified and characterized in many plants, such including 17 FADs in *Arabidopsis* [24], 29 FADs in soybean[27], 31 FADs in peanuts[28], and 41 FADs in Raymond cotton [29]. The *FAD* gene was also isolated from land cotton[30], sesame[31], sunflower [32] and other plants. In this study, 24 FAD family members were identified and analysed based on their protein properties. Then, analysis of the genetic structure showed that each small branch has a similar genetic structure, and the subfamily is highly conservative. For example, the Δ12/ω3 desaturase subfamily contains 7-8 exons, consistent with *Arabidopsis* [24] and *Carya cathayensis* [33]. There are 2-3 exons in the SAD subfamily, consistent with peanut [28]. Studies have shown that the origin of genetic introns is more ancient, and the insertion of introns is the result of exon rearrangement and plays an important role in gene evolution [34, 35]. By analysing the gene sequences, it can be found that the *JrFAD* family is highly conserved, but there are large differences between the subfamilies. The FAD family of walnut and *Arabidopsis thaliana* is composed of four major members of the SAD desaturase subfamily, the Δ7/Δ9 desaturase subfamily, the Δ12/ω3 desaturase subfamily and the “front-end” desaturase subfamily [24]. However, there are large differences in the copy number of genes in each subfamily. The most common gene copy number of 9 was found in the Δ9 desaturase subfamily in *Arabidopsis* [24], but the gene copy number is only one in walnuts. Similar to the results in walnuts, Δ9 desaturase was not identified in the whole genome of soybean; instead, *SAD* [27], which also introduced the first unsaturated bond at Δ9 desaturase, was detected. This suggests that in more evolved higher plants, the Δ9 desaturase, which is widely distributed in the biological world, is replaced by the SAD subfamily with similar functions and more copy numbers, which may be related to the evolution of the *SAD* gene family in plants[24]. The

number of members in the Δ12/ω3 desaturase subfamily is also higher in walnut than in *Arabidopsis*, probably because these genes have not undergone related genome-wide replication during evolution.

Thus far, there is increasing evidence that the linolenic acid content is closely related to the *FAD3* gene [36-39]. In *Arabidopsis* seeds, the specific overexpression of endogenous *FAD3* increased the α-linolenic acid content in the grain from 19% to 40% [40], while overexpression of *FAD3* genes in tomato resulted in α-linolenic acid levels being significantly increased in leaves and fruits of transgenic plants [40]. Ni *et al.* [41] proved that when *BnFAD3* was overexpressed, the content of α-linolenic acid in rapeseed increased. Studies have shown that RNAi technology inhibits the expression of the tobacco *FAD3* gene and can significantly reduce the accumulation of α-linolenic acid [42]. Liao *et al.*[43] inhibited the expression of *FAD2*, *FAD3* and *FATB* in rapeseed by constructing a multi-gene interference vector containing the seed-specific promoter NapinA, which could increase the oleic acid content by more than 16%. The content of linolenic acid in walnut kernels was high, and 5 copies were identified in the genome-wide range. However, only *JrFAD3-1* expression was found in the seed kernel during development with the peak expression occurring after approximately 100 days. Although *JrFAD3-1* is not a seed-specific gene, expression of this gene was significantly upregulated during seed development, showing a trend of first increasing and then decreasing, and tending to decline as the seed matures, consistent with the findings of previous studies [44-46]. This change may be related to the important role of *FAD3* in the synthesis of unsaturated fatty acids in walnut [47]. In this study, after the expression of *JrFAD3* in walnuts, there may have been two reasons for the change in α-linolenic acid content after 30 days. First, a process in the cytoplasm may be involved. Protein must play a role in mRNA production and must be transported from the cytoplasm to the nucleus, and protein expression requires a process such as post-modification, that is, there is a time lag in the process of gene expression [48-49]. Second, α-linolenic acid content is directly related to the amount of jasmonic acid produced, while ω-3 fatty acid dehydrogenase (FAD) can catalyse the synthesis of α-linolenic acid from linoleic acid [50]. There was a *JrFAD3-1* gene encoding Δ15 fatty acid desaturase in the seed, which catalysed the conversion of linoleoyl CoA to α-linoleyl CoA. Then, α-linolenic acid was transformed into jasmonic acid [51]. At this point, the protein content reached its highest level at 110 days after flowering and showed a positive correlation with the oil content [52]. It is also possible that the α-linolenic acid content caused by the lack of substrate in the early stage was not obvious in the early embryo development of walnut. As the embryo matured, the expression of *JrFAD3-1* was downregulated, and the α-linolenic acid content increased. In this process, the desaturation of the gene gradually decreased after the expression peaked, which may have been caused by the decrease in FAD3 activity and the consumption of α-linoleic acid. Through the verification of both positive and negative aspects, the gene function can be explained more comprehensively, which provides a theoretical basis for establishing a complete metabolic network.

4 Materials And Methods

4.1 Plant Materials

The test material was collected from Experimental Field of Hebei Agricultural University in 2016. The ‘Qingxiang’ variety, which has a full fruit bearing period, was selected as the test material. All materials have been permissions.

The samples used for transcriptome sequencing were ‘Qingxiang’ kernels collected 70, 90 and 110 days after anthesis.

The *FAD3-1* gene was cloned in kernels collected 100 days after flowering.

FAD3-1 gene expression analysis was conducted with ‘Qingxiang’ samples from 10 different developmental stages, and 8 different tissue: pistils, catkins, young branches (bark), old branches (bark), young leaves, mature leaves, young embryos and mature embryos. After the fruits were collected, the embryos were quickly removed by peeling and immersed in liquid nitrogen for quick freezing, and stored at -80 ℃ for RNA extraction, expression analysis and determination of fatty acid content.

The kits used for RNA extraction, DNA extraction, gel recovery and plasmid extraction were from Tiangen (product numbers: DP441, DP350, DP209 and DP103, respectively). The drugs used for reverse transcription, PCR and real-time PCR were from Takara (product numbers: RR047A, RR901A and RR820A, respectively).

4.2 Cloning of the JrFAD3-1 gene

The *JrFAD3-1* gene was cloned in kernels collected 100 days after flowering. *JrFAD3-1* partial gene sequence was identified from walnut genome data. So, homology-based cloning method was used to isolate *JrFAD3-1* coding sequence (CDS) from ‘Qingxiang’ walnut. According to the gene sequences of Soybeans and Castor beans. Premier 5.0 software was used to design specific primers. The *JrFAD3-1* gene specific primers as follow:

JrFAD3-1-F: 5'- GAGAAAGGAGGAGAGAATGC-3',

JrFAD3-1-R: 5'- GATAGCCTTGCTCTTCAAA-3'.

4.3 Identification of *FAD* Gene Family and Bioinformatics Analysis

The hidden Markov model was constructed using HMMER 3.1b software from the *Arabidopsis* FAD family protein sequence in the UniProt database and the protein sequence of walnut was downloaded from NCBI (www.ncbi.nlm.nih.gov/genome/?term=txid51240[orgn]). The conserved domains of the proteins were analysed using the PFAM protein family database [53] of the European Institute of Bioinformatics, using the online analysis tool Gene Structure Display Server[54] and MEME pair (http://meme-suite.org/) for the walnut FAD family. The *JrFAD* and *AtFAD* family phylogenetic trees were constructed using the maximum likelihood (ML) method in MEGA 5.0 software with a bootstrap value of 1000. The JrFAD and 31 other branches of phylogenetic trees were constructed using the maximum likelihood method in MEGA 6.0 software with a bootstrap value of 1000.

4.4 **Data analysis of transcriptome sequencing**

An Illumina high-throughput sequencing platform (HiSeq X Ten) was used for sequencing.

The RNA coding potential was predicted using CPAT v.1.2.2 software. The predicted coding transcripts were aligned with the GenBank NR, Swiss-prot, eggNOG, InterPro, GO, and KEGG databases for functional annotation. Differentially expressed genes were identified using R with the edgeR package, and q-values < 0.05 indicated differentially expressed genes.

4.5 **Tissue-specific expression in the FAD family**

For RNA isolation experiments, all samples were immediately frozen in liquid nitrogen and extracted using a Tiangen Plant RNA Extraction Kit according to the protocol. First-strand cDNA was synthesized with a Tiangen Inverse Transcription Kit. The gene-specific primers used for RT-PCR and qRT-PCR were designed using NCBI (Table 2). Five genes belonging to four subfamilies were selected to analyse the expression of the FAD family in different tissues.

The PCR conditions were as follows: 94°C for 5 min; 33 cycles of 94°C for 30 s, 54°C for 30 s, and 72°C for 1 min; and 72°C for 10 min.

Gene name	Sequence （5'–3'）
ACT2	F: TCCACCATGTTCCCTGGTAT; R: ACCTCCCAATCCAGACACTG
SLD-1	F: ACTTGGTTTGGGAGGCTGTC; R: TCGTCCTAACAACCCCAAGC
ADS3	F: TTCCTTGTGTGGGGAATGGG; R: AAATGCAAGCAATGCCACCC
SAD-2	F: TTTTTCCCCTTCTCAACCTCCAA; R: TGGTGTAAAAGGCTTCTTTGGATTG
SAD6	F: AAGCTGGAGGATTTGACGGG; R: CCAGCTAAACTTAACGCCGC
FAD3-1	F: ATCCACCATGATATCGGCACC; R: GAAGGGGAATTGGCCCAGA

Table 2 Primers used in this study

4.6 **Gene expression analysis by quantitative RT-PCR**

The qRT-PCR experiments were performed on an FX96 Touch™ Real-Time PCR Detection System (Bio-Rad) using a SYBR Green I Master Kit (Tiangen). *JrActin* was used as the reference gene to calculate the

relative fold differences based on the comparative cycle threshold ($2^{-\Delta\Delta ct}$). The PCR solution (20 μl) contained 10 μl of SYBR-Green PCR Master Mix (Tiangen), 0.2 μM of each primer, 100 ng of cDNA template, and nuclease-free water. The PCR conditions were as follows: 95°C for 5 min; 45 cycles of 94°C for 10 s, 60°C for 30 s, and 72°C for 30 s; followed by a final extension at 72°C for 3 min.

4.7 Determination of fatty acid content

The determination of unsaturated fatty acids referred to Bai [55]

4.8 Statistical Analysis

The data were analysed using SPSS 17.0 software.

5 Conclusions

25 walnut kernels *FAD* genes were identified and comprehensive analyzed for the first time. The function of a walnut kernels *FAD3-1* gene was also characterized from its location in the phylogeny. This study provides significant information for further in-depth molecular investigations into the biological of these genes.

Abbreviations

ADS:Arabidopsis desaturase; FAD:Fatty acids desaturase; SAD: Stearyl acyl-carrier-protein desaturase; DES: delta(7)-sterol-C5(6)-desaturase; ALD: acyl-lipid (9-3)-desaturase; SCD: stearoyl-CoA-desaturase; ACCase: Acetyl coenzyme A carboxylase; KCS2: 3-ketoacyl-CoA thiolase; fabG: 3-oxoacyl-ACP reductase; fadE: acyl-CoA dehydrogenase; ALD: acyl-lipid (9-3)-desaturase; ACP: acyl-carrier protein; DAL: delta(12)-acyl-lipid-desaturase; DES:sphingolipid delta(4)-desaturase; SLD: delta(7)-sterol-C5(6)-desaturase; ML: maximum likelihood.

Declarations

Ethics approval and consent to participate

Walnut is one of widespread fruit trees in China, and it is not an endangered species. The walnut trees were from the Experimental Field of Hebei Agricultural University. No specific permits are required for sample collection on walnut.

Consent for publication

Not applicable.

Availability of data and materials

All data and materials are presented in the main paper and additional supporting file.

Competing interests

The authors declare that they have no competing interests.

Funding

Supported by the Natural Science Foundation of Hebei Province (C2019204270), the Science and Technology Research Project of the Universities of Hebei Province (QN2019159), and Research Foundations for Returned Scholars from Overseas of the Human Resources Dept. of Hebei Province (C20190345). These funding bodies had no role in the design of the study, sample collection, analysis or interpretation of data and in writing the manuscript.

Authors' contributions

HW and ZZ designed the research and wrote the paper. KL and SZ performed the experiments, analyzed the data and wrote the paper. SW participated in the data analysis.

KL and SZ are equal contributors. All authors read and approved the final the manuscript.

Acknowledgements

Not applicable.

Author details

[1] Mountainous Areas Research Institute, Hebei Agricultural University, Baoding 071001, China.

[2] College of Life Sciences, Hebei Agricultural University, Baoding 071001, China.

[3] Research Center for Agricultural Engineering Technology of Mountain District of Hebei, Baoding 071001, China.

[4] National Engineering Research Center for Agriculture in Northern Mountainous Areas, Baoding 071001, China.

[#] Corresponding author E-mail: zshug@hebau.edu.cn;whx@hebau. edu. cn; zzh@hebau. edu. cn.

[†] Kai Liu and Shugang Zhao contributed equally to the work.

References

1. Hsieh K, Huang AHC. Endoplasmic Reticulum, oleosins, and oils in seeds and tapetum cells. Plant Physiology. 2004; 136(3): 3427-3434.

2. Zhao SG, Zhao YP, Wang HX, Zhang ZH. Physieal/chemieal chaareteirs ties and fatty acid composition of walnut oil. Journal of the Chinese Cereals and Oils Association. 2008; 23(2): 85-88.
3. Anderson KJ, Teuber SS, Gobeille A, Cremin P, Waterhouse AL, Steinberg FM. Walnut polyphenolics inhibit in vitro human plasma and LDL oxidation. Journal of Nutrition. 2001; 131(11): 2837-2842.
4. Vingrys AJ, Armitage JA, Weisinger HS, Bui BV, Sinclair AJ, Weisinger RS. The role of omega-3 polyunsaturated fatty acids in retinal function. Fatty Acids. 2001; 12: 193-217.
5. Boelsma E, Hendriks HF, Roza L. Nutritional skin care: health effects of micronutrients and fatty acids. American Journal of Clinical Nutrition. 2001; 73(5): 853-864.
6. Lu YD, Chi XY, Li ZX, Yang QL, Li FC, Liu SF, Gan QH, Qin S. Isolation and characterization of a stress-dependent plastidial Δ^{12}-fatty acid desaturase from the antarctic microalga chlorella vulgaris NJ-7. Lipids. 2010; 45(2): 179-187.
7. Xue Y, Chen B, Win AN, Fu C, Lian J, Liu X, Wang R, Zhang X, Chai Y. Omega-3 fatty acid desaturase gene family from two ω-3 sources, *Salvia hispanica* and *Perilla frutescens*: cloning, characterization and expression. *Plos One*. 2018; 13(1): e0191432.
8. Majou D. Evolution of the human brain: the key roles of *DHA* (ω-3 fatty acid) and Δ6-desaturase gene. OCL. 2018; 25(4): A401.
9. Miller CN, Polasek LK, Oliveira AC, Frost CJ, Maniscalco JM. Milk fatty acid composition of perinatal and foraging steller sea lions: examination from pup stomachs. Canadian Journal of Zoology. 2018; 96(2). Available from:https://doi.org/10.1016/j.plefa.2015.11.005.
10. Horiguchi S, Nakayama K, Iwamoto S, Ishijima A, Minezaki T, Baba M, Kontai Y, Horikawa C, Kawashima H, Shibata H, Kagawa Y, Kawabata T. Associations between a fatty acid desaturase gene polymorphism and blood arachidonic acid compositions in Japanese elderly. Prostaglandins Leukotrienes and Essential Fatty Acids. 2016; 105: 9-14.
11. Torresfranklin ML, Repellin A, Huynh VB, Agnès d'Arcylameta, Zuilyfodil Y, Phamthi AT. Omega-3 fatty acid desaturase (*FAD3, FAD7, FAD8*) gene expression and linolenic acid content in cowpea leaves submitted to drought and after rehydration. Environmental and Experimental Botany. 2009; 65(2/3): 162-169.
12. Feng JY, Dong YT, Liu W, He QL, Daud MK, Chen JH, Zhua SJ. Genome-wide identification of membrane- bound fatty acid desaturase genes in *Gossypium hirsutum* and their expressions during abiotic stress. Scientific Reports. 2017; 7: 45711.
13. Shanklin J, Cahoon EB. Desaturation and related modifications of fatty acids. Annual Review of Plant Physiology & Plant Molecular. 1998; 49(49): 611-641.
14. Rawlings BJ. Biosynthesis of fatty acids and related metabolites. Natural Product Reports. 1998; 15(4): 275-308.
15. Ntambi JM, Buhrow SA, Kaestner KH, Christy RJ, Sibley E, Kelly TJ, Lane MD. Differentiation-induced gene expression in *3T3-L1* preadipocytes. Characterization of a differentially expressed gene encoding stearoyl-CoA desaturase. Journal of Biological Chemistry. 1988; 263(33): 17291-17300.

16. Shanklin J, Somerville C. Stearoyl-acyl-carrier-protein desaturase from higher plants is structurally unrelated to the animal and fungal homologs. Proceedings of the National Academy of Sciences. 1991; 88(6): 2510-2514.
17. Kachroo A, Shanklin J, Whittle E, Lapchyk L, Hildebrand D, Kachroo P. The *Arabidopsis* stearoyl-acyl carrier protein-desaturase family and the contribution of leaf isoforms to oleic acid synthesis. Plant Molecular Biology. 2007; 63(2): 257-271.
18. Liu XL, Jiang XZ, Zhang ML Huang JZ. Cloning and high expression of Δ12-desaturase genes from *Mucor circinelloides* EIM-10. Science and Technology of Food Industry. 2015; 36(21): 196-199+213.
19. Browse J, Mcconn M, James D, Miquel M. Mutants of *Arabidopsis* deficient in the synthesis of alpha-linolenate. Biochemical and genetic characterization of the endoplasmic reticulum linoleoyl desaturase. Journal of Biological Chemistry. 1993; 268(22): 16345-16351.
20. Gibson S, Arondel V, Iba K, Somerville C. Cloning of a temperature-regulated gene encoding a chloroplast ω-3 desaturase from *Arabidopsis thaliana*. Plant Physiology. 1994; 106(4): 1615-1621.
21. Mcconn M, Hugly S, Browse J, Somerville Chris. A mutation at the *fad8* locus of *Arabidopsis* identifies a second chloroplast ω-3 desaturase. Plant Physiology. 1994; 106(4): 1609-1614.
22. Iba K, Gibson S, Nishiuchi T, Fuse T, Nishimura M, Arondel V, Hugly S, Somerville C. A gene encoding a chloroplast ω-3 fatty acid desaturase complements alterations in fatty acid desaturation and chloroplast copy number of the *fad7* mutant of *Arabidopsis thaliana*. Journal of Biological Chemistry. 1993; 268(32): 24099-24105.
23. Los DA, Murata N. Structure and expression of fatty acid desaturases. Biochimica Et Biophysica Acta. 1998; 1394(1): 3-15.
24. Arondel V, Lemieux B, Hwang I, Gibson S, Goodman H, Somerville C. Map-based cloning of a gene controlling ω-3 desaturation in *Arabidopsis*. Science. 1992; 258(5086): 1353-1355.
25. 25. Xiao LJ, Zhang Y, Wu T, Ning DL, Mao YL. Fatty acid content and nutritional evaluation of Yunnan purple kernel walnut. China Oils and Fats. 2014; 39(9): 94-97.
26. 26. Li WB, Lei F, Song W, Luo J, Bi WS, Liu Y, Xu LX, Zhao X. Accumulation dynamics of fatty acid in soybean seed and screening of candidate gene related to linolenic acid metabolism. Journal of Northeast Agricultural University. 2017; 48(11): 1-8.
27. Chi XY, Yang QL, Lu YD, Wang JY, Zhang QF, Pan LJ, Chen MN, He YN, Yu SL. Genome- wide analysis of fatty acid desaturases in soybean (*Glycine max*). Plant Molecular Biology Reporter. 2011; 29(4): 769-783.
28. Ruan J, Shan L, Li XG, Guo F, Meng JJ, Wan SB, Peng ZY. Genome-wide Identification and expression pattern analysis of peanut *FAD* gene family. Shandong Agricultural Sciences. 2018; 50(6): 1-9.
29. Liu W. Genome-wide identification and gene structure and function analysis of *CDPK* and *FAD* gene families of two diploid cotton species. Zhejiang university. 2015; 4: 60-92.
30. Pirtle IL, Kongcharoensuntorn W, Nampaisansuk M, Knesek JE, Chapman KD, Pirtle RM. Molecular cloning and functional expression of the gene for a cotton Δ12-fatty acid desaturase (*FAD2*). Biochim Biophys Acta. 2001; 1522(2): 122-129.

31. Jin UH, Lee JW, Chung YS, Lee JH, Yi YB, Kim YK. Characterization and temporal expression of a ω-6 fatty acid desaturase cDNA from sesame (*Sesamum indicum* L.) seeds. Plant Science. 2001; 161(5): 935-941.
32. Martínezrivas JM, Sperling P, Lühs W, Heinz E. Spatial and temporal regulation of three different microsomal oleate desaturase genes (*FAD2*) from normal-type and high-oleic varieties of sunflower (*Helianthus annuus* L.). Molecular Breeding. 2001; 8(2): 159-168.
33. 33. Huang CY, Huang YJ, Wu JF, Huang R, Luan YM, Zhang SM, Wang ZJ, Zhang QX, Huang JQ. *SAD* and *FAD* genes regulate the ratio of unsaturated fatty acid components in *Carya cathayensis.* Acta Horticulturae Sinica. 2018; 45(2): 250-260.
34. Cavaliersmith T. Intron phylogeny:a new hypothesis. Trend in Gentics. 1991; 7(5): 145-148.
35. Rzhetsky A, Ayala FJ, Hsu LC, Chang C, Yoshida A. Exon/intronstructure of aldehyde dehydrogznase genes supports the “introns-late”theory. Proceeding of National Academy of science USA. 1997; 94(13): 6820-6825.
36. Hita WD, Carlson TJ, Booth JR, Kinney AJ, Stecca KL, Yada NS. Cloning of a higher-plant plastid ω-6 fatty acid desaturase cDNA and its expression in a Cyanobacterium. Plant Physiol, 1994; 105(2): 634-641.
37. Banilas G, Moressis A, Nikoloudakis N, Hatzopoulos P. Spatial and temporal expressions of two distinct oleate desaturases from olive (*Olea europaea* L.). Plant Science.2005; 168(2): 547-555.
38. Teixeira MC, Coelho N, Olsson ME, Brodelius PE, Carvalho IS, Brodelius M. Molecular cloning and expression analysis of three ω-6 desaturase genes from purslane (*Portulaca oleracea* L.). Biotechnol Letters. 2009; 31(7): 1089-1101.
39. Mikkilineni V, Rocheford TR. Sequence variation and genomic organization of fatty acid desaturase-2 (*fad2*) and fatty acid desaturase-6 (*fad6*) cDNAs in maize. Theoretical and applied genetics. 2003; 106(7): 1326-1332.
40. Puttick D, Dauk M, Lozinsky S, Smith MA. Overexpression of a *FAD3* desaturase increases synthesis of a polymethylene-interrupted dienoic fatty acid in seeds of *Arabidopsis thaliana* L. lipids. 2009; 44(8): 753-757.
41. Ni XF, Zhao CZ, Li H, Zhang M. Construction of *TmFAD2* and *BnFAD3* double-gene expression vector and fatty acid analysis of transgenic plants. Acta Agriculturae Boreali-occidentalis Sinica. 2014; 23(10): 112-119.
42. Murakami Y, Tsuyama M, Kobayashi Y, Kodama H, Iba K. Trienoic fatty acids and plant tolerance of high temperature. Science. 2000; 287(5452): 476-479.
43. Liao XJ, Wang XJ, Wang B, Liu AZ. Cloning of the promoter of the key enzyme gene *PvFAD3* from Sacha Inchi and its transcriptional activation analyses. Plant Physiology Journal. 2017; 53(1): 79-88.
44. Wang Z, Wang SY, Han B, Xu J, Yu DY, Zhang GH. Analysis of the expression of fatty acid desaturase gene (*fads*) in oil Flax. Molecular Plant Breedin*g*. 2019; 1-12. Available from: http://kns.cnki.net/kcms/detail/ 46.1068.S.20190424. 1717. 031. html.

45. Wang SY, Han B, Zhou SM, Xu J. Expression of functional genes related to fatty acid synthesis of oil flax. Chinese Journal of Oil Crop Sciences. 2016; 38(6): 771-777.
46. Banik M, Duguid S, and Cloutier S. Transcript profiling and gene characterization of three fatty acid desaturase genes in high, moderate, and low linolenic acid genotypes of flax (*Linum usitatissimum* L.) and their role in linolenic acid accumulation. Genome. 2011; 54(6): 471-483.
47. Vrinten P, Hu Z, Munchinsky MA, Qiu RX. Two *fad3* desaturase genes control the level of linolenic acid in flax seed. Plant Physiology. 2005; 139(1): 79-87.
48. He Y, Zeng J, Wu M,Zhang CK, Zhang Y. Stability for discrete-time genetic regulatory network with time- verying interval delay. Control Theory & Applications. 2012; 29(11): 1465-1470.
49. Yang RD. Introduction to modern biological science and bioengineering. *Guangzhou: south* China university of technology press. 2006; 6(2): 82-83.
50. Makandar R, Nalam V, Chaturvedi R, Jeannotte R, Sparks AA, Shah J. Involvement of salicylate and jasmonate signaling pathways in *Arabidopsis* interaction with *Fusarium graminearum*. Molecularplant-microbeinteractions. 2010; 23(7): 861-870.
51. Suza WP, Avila CA, Carruthers K, Kulkarni S, Goggin FL, Lorence A. Exploring the impact of wounding and jasmonates on ascorbate metabolism. Plant Physiology and Biochemistry. 2010; 48(5): 337-350.
52. Chen H, Pan CD, Wang B, Hu Y, Xiao ZZ, He M. The relationship among nutrients' accumulation and dynamic changes of fatty acids in seed development of walnut. Journal of Hebei Agricultural University. 2016; 39(1): 57-62 +74.
53. Finn RD, Coggill P, Eberhardt RY, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Research. 2016; 44(Database issue): D279-D285.
54. Guo AY, Zhu QH, Chen X, Luo JC. GSDS: a gene structure display server. Hereditas. 2007; 29(8): 1023-1026.
55. Bai CK. Fatty acids analysis in *Decaisnea insignis* and *Akebia trifoliate* seed oil by GC-MS. Acta Botanica Boreali-Occidentalia Sinica. 2007; 27(5): 0859-0863.

Figures

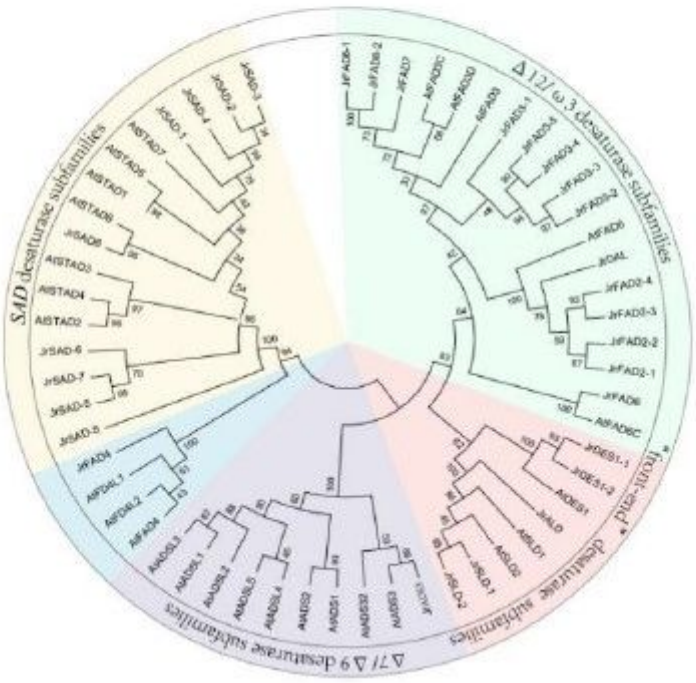


Fig. 1 Phylogenetic analysis of *Arabidopsis thaliana* and *Juglans regia* FADs

Figure 1

Phylogenetic analysis of Arabidopsis thaliana and Juglans regia FADs

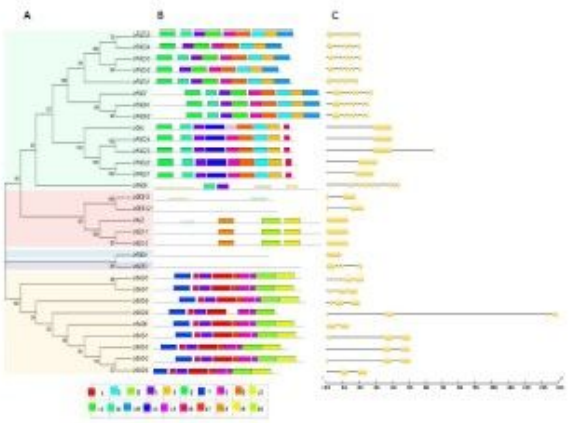


Fig. 2 Gene structure and conserved motif analysis of the FAD family from Juglans regia
Note: A: Phylogenetic analysis of JrFAD family; B: Conserved motif of JrFAD family; C Gene structure of JrFAD family



Fig. 2D Motifs of walnut FAD family proteins

Figure 2

Gene structure and conserved motif analysis of the FAD family from Juglans regia Note: A: Phylogenetic analysis of JrFAD family; B: Conserved motif of JrFAD family; C Gene structure of JrFAD family

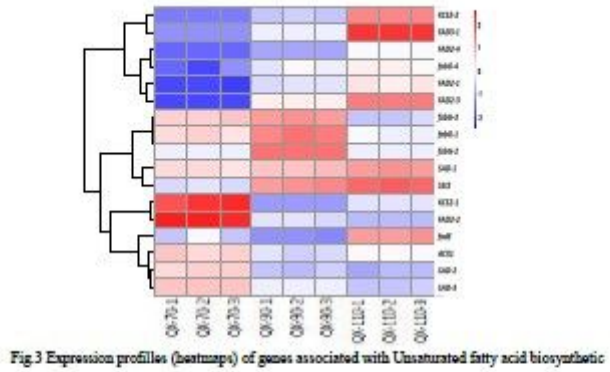


Fig.3 Expression profilles (heatmaps) of genes associated with Unsaturated fatty acid biosynthetic

Figure 3

Expression profilles (heatmaps) of genes associated with Unsaturated fatty acid biosynthetic

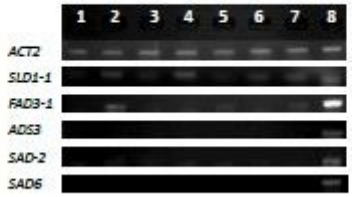


Fig. 4 Tissue-specific expression of walnut *FAD* family

Note：1: pistils; 2: catkin; 3: young stem; 4: older stem; 5: young leaves: 6: mature leaves; 7: young embryo; 8: mature embryo

Figure 4

Tissue-specific expression of walnut FAD family Note：1: pistils; 2: catkin; 3: young stem; 4: older stem; 5: young leaves: 6: mature leaves; 7: young embryo; 8: mature embryo 1 2 3 5 4 6 7 8 SLD1-1 FAD3-1 ADS3 SAD-2 SAD6 ACT2

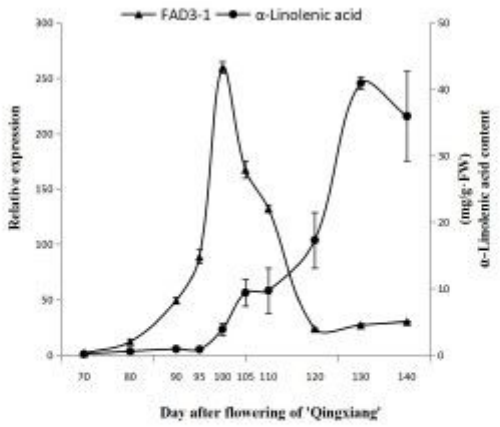


Fig.5 Expression of *FAD3-1* and content of α-Linolenic acid in 'Qingxiang'

Figure 5

Expression of FAD-31



Fig. 6 Conservation of FAD3 proteins in 32 different

Figure 6

Conservation of FAD3 proteins in 32 different



Fig. 7 Multiple sequence alignment of FAD3 proteins in 32 different species

Figure 7

Multiple sequence alignment of FAD3 proteins in 32 different species

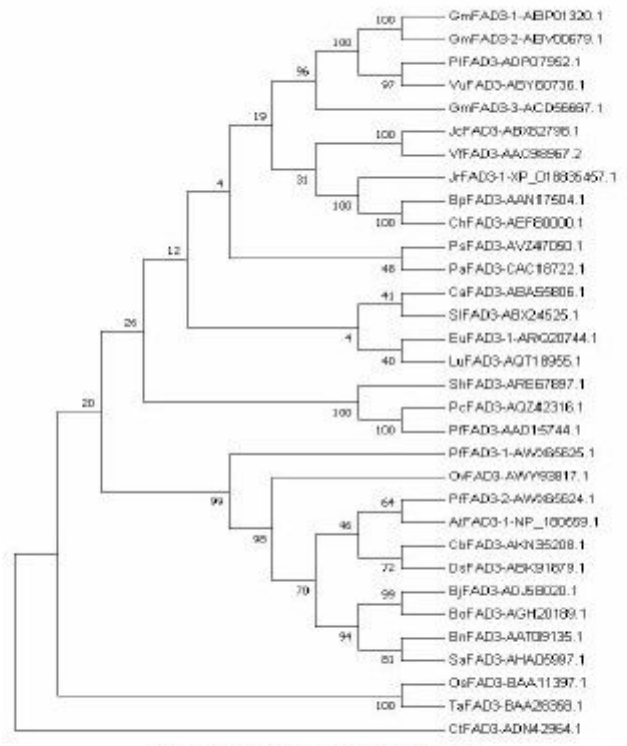


Fig. 8 phylogeny of FAD3 proteins in different species

Figure 8

phylogeny of FAD3 proteins in different species

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFig.1.pdf