

Bioinformatics and Expression Pattern Analysis of Soybean Fatty Acid Desaturase Family Gene

Nan Wu

Jilin Agricultural Science and Technology College

Long Jiang

Jilin Agricultural Science and Technology College

Xiaoming Yu

Jilin Agricultural Science and Technology College

Xiangbo Yang

Jilin Agricultural Science and Technology College

Dianyuan Chen

Jilin Agricultural Science and Technology College

Minghao Yuan (✉ 494352272@QQ.com)

Jilin Agricultural Science and Technology College

Research article

Keywords: FAD2, Bioinformatics, Analysis of Expression Patterns, Oleic acid content

Posted Date: January 6th, 2020

DOI: <https://doi.org/10.21203/rs.2.19966/v2>

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

[Read Full License](#)

Abstract

Background

To clarify the homology and structural differences of seven genes in the soybean fatty acid desaturase 2 (FAD2) family and the relationship between the expression level of each gene and oleic acid content at different stages of grain development, the seven genes in this family were studied through informatics analysis and evaluation of expression patterns.

Results

Database analysis, phylogenetic tree analysis, transmembrane structure prediction, amino acid sequence alignment, intron exon structure analysis, and protein motif analysis were performed for seven genes in the FAD2 family. The results showed that GmFAD2-1 and GmFAD2-2b and Other genes are structurally different, and it is speculated that their functions also have weak changes. By analyzing gene expression patterns at different stages of soybean grain development, the expression levels of the family genes at different stages of grain development first increase, then decrease, and finally decrease. For stability, the correlation analysis showed that the gene expression was negatively correlated with oleic acid content, and the correlation coefficients of GmFAD2-1 and GmFAD2-2b were the largest, showing a significant negative correlation.

Conclusions

Bioinformatics and expression analysis of genes in the FAD2 family. Identify the structural differences of each gene, and the differences in expression levels at different stages of soybean variety and grain development, and initially identify the differences in function size and catalytic efficiency between genes, in order to create new high-oleic soybean varieties using genetic engineering Quality, providing a theoretical basis.

Background

Research interest in the study of fatty acid desaturase (FAD2) genes is high throughout the world. The main breeding target for soybean fatty acids is improvement of the stability of fatty acid storage, and thus, studies have primarily focused on germplasm innovation of soybean to produce high oleic acid content and reduce the linoleic acid and linolenic acid content, changing their proportion^[1]. At present, many FAD genes have been identified in crops such as peanut^[2], cotton^[3], and soybean^[4]. In 2014, William Haun^[5] et al. studied the soybean FAD2 family and increased the oleic acid content from 20–80%. At the same time, the content of linoleic acid, which is harmful to human health was reduced (from 50–4%), thus improving the quality of soybeans. In 2017, Wen^[6] et al. raised the oleic acid content of flax seeds from 16% to more than 50% and decreased the linolenic acid and linoleic acid content from 35–9%. In 2017, Yan^[7] et al. studied the transcription factor GmWRI1a, which is involved in fatty acid

synthesis and glycolysis pathways. Compared with wild-type materials, the oleic acid content was decreased by 11.44%.

In 2012, Schlueter^[8] et al. provided a detailed description of the soybean FAD2 gene, indicating that $\Delta 12$ -fat dehydrogenase (FAD2) is a plastid ω -6 dehydrogenase in soybean. The protein, which is predicted to be primarily distributed on the plastid and endoplasmic reticulum membrane, is responsible for the formation of a hydrogen bond at a specific position in the fatty acid chain of the oleic acid, such as carbon no. 12 and no. 13. Linoleic acid is a key enzyme that determines the amount of linoleic acid synthesis. $\Delta 12$ FAD is a membrane-bound protein primarily present in the plastid membrane and converts oleic acid C181 to linoleic acid C182. The genes in this family are mainly present on the endoplasmic reticulum of plants and on the cell membrane of plastids^[9–11] and participate in the metabolic synthesis of most unsaturated fatty acids. However, there is no synthetic process to produce linoleic acid in animals, and thus, it must be ingested in the diet^[12–14]. Unsaturated fatty acids have two particularly important functions: maintaining the flow of membrane lipids and serving as precursors for synthesis of arachidonic acid^[15–18]. The first report of FAD2 described the enzyme in Arabidopsis. Later, scientists cloned other genes in the family. These genes have high homology with the Arabidopsis gene and high amino acid sequence identity^[19, 20].

The methods for determining soybean oleic acid content include grain analysis using near-infrared spectroscopy^[21], gas chromatography^[22], and nuclear magnetic resonance spectroscopy^[23]. For grain analysis using near-infrared spectroscopy, soy seeds that are smooth, full, and insect-free are selected as seeds to be tested and placed in the infrared sensing area of the instrument, and the results of at least three replicate seeds are recorded in the computer. In 2008, Zhang^[24] et al. established a method based on neural networks for near-infrared spectroscopy of soybean oil content and demonstrated the feasibility of this method. In 2007, Chai^[25] et al. performed similar research in Heilongjiang. With soybean as the basic material, a near-infrared grain analyser was used to measure the sample to obtain a more complete spectrum database, and the experimental data were analysed using the linear regression method and processed to determine the soybean oleic acid content. In 2017, Wang^[26] et al. used a near-infrared grain analyser to establish a rapid soybean oleic acid content determination system. They selected 289 soybean seeds at Jilin Agricultural University, established a calibration model for soybeans in Jilin Province, and constructed a spectrum of five fatty acid components. The model analysed provides theoretical and data support for the determination of soybean oleic acid.

Results

Bioinformatics Analysis

Gene domain analysis of the soybean FAD2 family. According to the amino acid sequence of the seven genes in the FAD2 family in soybean, the accession numbers are shown in Table1–1, and comparison statistics were performed in the database Ensembl.plants. Sequences without complete open reading

frames were excluded, and the domains of each gene were further analysed by CDD using an online website. ProtParam in the Expasy website was used to analyse the physicochemical properties of the protein structural components and the isoelectric points of various genes in the family.

Table 1-1
1 Function and identification number of soybean fatty acid
desaturase genes

Name	GenBank ID	Function
FAD2-2	L43921	Catalysing oleic acid to form linoleic acid
FAD2-1b	AB188251	Catalysing oleic acid to form linoleic acid
FAD2-1	AY954300	Catalysing oleic acid to form linoleic acid
FAD2-2a	AB188252	Catalysing oleic acid to form linoleic acid
FAD2-2b	AB188253	Catalysing oleic acid to form linoleic acid
FAD2-3	DQ53237	Catalysing oleic acid to form linoleic acid
FAD2-1B	DQ532370	Catalysing oleic acid to form linoleic acid

As shown in Table 1–2, the study found that the gene family is distributed on 6 chromosomes, the number of amino acids encoded by the genes ranges between 216 and 387, and the isoelectric point is in the range of 9.18–10.13, which is alkaline. The FAD2 family of proteins has highly conserved amino acid regions, primarily histidine-rich segments, indicating that the genes in this family have the same ancestor.

Table 1-2
2 Soybean FAD2 family gene database information

Gene cluster	Name	Locus tag	Chromosomal location	CDS (bp)	Protein (aa)	PI
FAD2	FAD2-2	Glyma.03G144500.1	Chr03:36014626..36020913	1152	383	9.86
		Glyma.03G144500.2	forward	1152	383	9.97
	FAD2-1b	Glyma.09G111900	Chr09:22174485..22179274 reverse	1161	386	9.92
	FAD2-1	Glyma.10G278000.1	Chr10:50013484..50015460	1164	387	10.08
		Glyma.10G278000.2	forward	1164	387	9.45
	FAD2-2a	Glyma.15G195200	Chr15:22308021..22308902 reverse	882	293	9.18
	FAD2-2b	Glyma.19G147300	Chr19:40814864..40815855	651	216	9.86
			forward			
	FAD2-3	Glyma.19G147400.1	Chr19:40821993..40824926	1152	383	10.13
		Glyma.19G147400.2	forward	1152	383	9.90
	FAD2-1B	Glyma.20G111000.1	Chr20:35315630..35319063	1164	387	9.42
		Glyma.20G111000.2	forward	1164	387	9.72
		Glyma.20G111000.3		1140	379	9.84
		Glyma.20G111000.4		1164	387	9.77
		Glyma.20G111000.5		1164	387	9.51
		Glyma.20G111000.6		1164	387	9.69
		Glyma.20G111000.7		1164	387	9.73

Transmembrane structure analysis of soybean FAD2 family genes

The transmembrane structure of proteins in the soybean FAD2 gene family was analysed using the software TMHMM (<http://www.cbs.dtu.dk/services/TMHMM/>), as shown in Fig. 1.1, and the results showed that the gene family had 2–6 different transmembrane structures, and the distance between the regions was very small. The main body is primarily located on the endoplasmic reticulum, and both ends have 100 amino acids located outside the membrane, which was predicted by subcellular localization

(Cell-PLoc <http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc/>). As a result, all the genes were found to be localized at the endoplasmic reticulum. This conclusion is consistent.

Phylogenetic tree analysis of FAD2 genes among different species

By using the software MEGA5, the amino acid sequence of each member of the FAD2 family in each species was searched, and the NJ method was used for comparison to study the evolutionary position of each gene in the family and the kinship structure between the genes. As shown in Fig. 1.2, we found that the homology of soybean FAD2 family genes is similar to the Arabidopsis FAD2 family gene^[27]. GmFAD2-2 and GmFAD2-2a have the highest homology; GmFAD2-1 and GmFAD2-1b have the highest homology and are less homologous to other soybean FAD2 genes, and their functions are also weakly differentiated. From the phylogenetic tree, we did not see the emergence of specific populations, which indicates that the FAD2 genes have been relatively conserved throughout evolution.

Analysis of the response element in each gene promoter region

The online site Plant CARE was used for promoter region analysis of each gene in the gene family. As shown in Table 1–3, hormone response elements and illumination responses are present in the soybean FAD gene family. Components, such as meristematic response elements and cis-acting elements, involved in low temperature reactions, were similar to the FAD2 genes in other species when subjected to stress^[28–30].

Table 1-3
3 Analysis of cis-elements in promoters of different genes

Gene	Cis-acting element	Sequence
GmFAD 2–2	Gibberellin response element P-box	CCTTTTG
	Drought-related inducing element MBS	TAAC TG
	Mesogenic tissue expression related components CAT-box	GCCACT
	Auxin response element TGA-element	AACGAC
GmFAD2-1b	Photoresponsive response element I-box	GTATAAGGCC
	Drought-related inducing element MBS	TAAC TG
GmFAD2-1	Anaerobic inducing element ARE	AAACCA
	Partial photo response element Box 4	ATTAAT
	Cis-acting element involved in low temperature reaction LTR	CCGAAA
GmFAD2-2a	Partial photo response element Box 4	ATTAAT
	Mesogenic tissue expression related components CAT-box	GCCACT
	Zein metabolism regulating element O2	GATGATGTGG
	Photoresponsive response element SP1	GGGCGG
GmFAD2-2b	Partial light response module AE-box	AGAACTT
	Anaerobic inducing element ARE	AAACCA
	Participation in endosperm expression of cis-elements GCN4-M	TGAGTCA
GmFAD2-3	Partial photo response element Box 4	ATTAAT
	Gibberellin response element P-box	CCTTTTG
	Drought-related inducing element MBS	TAAC TG
	Zein metabolism regulating element O2	GATGATGTGG
GmFAD2-1B	Salicylic acid cis component TCA-element	CCATCTTTT
	Mesogenic tissue expression related components CAT-box	GCCACT

Analysis of Structural Patterns in the Soybean FAD2 Gene Family

In the Fig. 1.3, the exon structure analysis: yellow represents the exon region; blue represents the untranslated region; and black lines represent introns, since some genes have introns in the structure that are too long. The inconvenience is shown in the figure, and thus, in this experiment, the length of the

default intron is the same. The results showed that the genes in the FAD2 family known in soybean, except for GmFAD2-2a and GmFAD2-2b, were structurally similar, indicating that the genes in the family have high consistency to some extent.

To study the evolutionary relationship between genes in the family, the protein structure was constructed using the online software MEME. From Fig. 1.4, we can see that the FAD2 genes clustered together in the soybean FAD2 gene family, with the exception of the GmFAD2-2a and GmFAD2-2b genes, and have almost the same amino acid type and the same protein motif type, which is consistent with the above results.

Amino acid sequence alignment between species and comparison with the Arabidopsis gene amino acid sequence

In the Fig. 1.5 showed that the protein domains are similar to those of the Arabidopsis gene and contain conserved histidine clusters (red box in the figure): HECGHH, HRRHH, and HVAHH. In GmFAD2-2b, the first histidine cluster and the third histidine cluster become HDCGHH and HVVHH, respectively, but the position of the histidine cluster is unchanged. These three histidine clusters are responsible for the primary FAD2 activity. The main binding sites of Fe^{2+} are necessary for completion of the catalytic reaction. The distance between the first cluster and the second cluster is relatively small, and the distance between the second cluster and the third cluster is large. This result is consistent with the oil of other species, such as sunflower species^[31].

Gene expression pattern analysis

Analysis of expression patterns in different organs under low temperature stress

Through low temperature stress treatment, quantitative analysis of different genes in different organs of soybean was performed using a quantitative fluorescence technique. The results showed that the family genes were mainly divided into two types: the first type comprised GmFAD2-2 and GmFAD2-1b, which showed seed-specific expression, and the remaining genes, which are constitutively expressed, were categorized as the other type. Under low temperature (4 °C) stress treatment, the expression levels of genes in the family were expressed to varying degrees, as shown in Fig. 1.6.

The above results indicate that the expression patterns of the seven genes in the family were different in different soybean tissues and during low temperature treatment. Among them, only GmFAD2-2 showed significant expression in the grain, and the expression level in other tissues was extremely high. The expression level of the gene was up-regulated with an increase in the low temperature treatment time, and the expression level was the highest at 48 hours of low temperature treatment. GmFAD2-1b also showed significant expression in the grain, but with a change in the low temperature treatment time, the expression level was down-regulated, and the expression level was the highest when the plant was treated at low temperature for 6 hours.

Analysis of expression patterns of various genes at different stages of grain development

In the process of soybean oleic acid synthesis, the expression level of FAD2 genes directly affects the level of oleic acid. As shown in Fig. 1.7, the expression levels in soybean seeds changed to different degrees at different stages of soybean development, but the overall trend was consistent. In the early stage of soybean grain development, the gene expression was higher. As the grain matured, the expression level gradually decreased and finally stabilized.

Determination of oleic acid content at different stages of grain development

It can be seen from Fig. 1.8 that in the soybean genotypes "Jike Soybean 20", "Ji Midou 3", "Jike Midou No. 1" and "Jike Fresh Bean 1" evaluated in this experiment, the trend in oleic acid content changes in soybean seeds was basically the same. The oleic acid content in "Ji Midou 3" was highest on the 35th day after flowering, and the period of the highest oleic acid content in "Jike Soybean 20", "Jike Midou No. 1", and "Jike Fresh Bean 1" was found on the 40th day after flowering, but there were differences in oleic acid content between the varieties because the synthesis of oleic acid is affected by differences between genotypes.

Analysis of the differences in FAD2 gene expression levels between different grain development stages (Table 1–4) showed that in the genotype "Jike Fresh Bean 1", the total expression of the genes was the largest, and the difference was significant, resulting in the low oleic acid content of the variety at the early stage. The experimental data were consistent. Among the other varieties, there were no significant differences in the content.

Table 1-4

Significance analysis of FAD2 gene expression at different grain development stages

Genotype	25th day after flowering	30th day after flowering	35th day after flowering	40th day after flowering	45th day after flowering	Total expression
Jike Soybean 20	6.39c	4.42c	4.21b	2.702c	5.091b	22.802
Ji Midou 3	6.15c	5.81b	3.296d	3.599b	5.198b	24.091
Jike Midou No. 1	10.98b	4.42c	3.672c	3.587b	4.432c	27.218
Jike Fresh Bean 1	13.76a	10.34a	5.192a	4.199a	7.308a	40.592

Correlation analysis between gene expression and oleic acid content

By analysing the correlation between the expression level of FAD2 gene expression in grain development and oleic acid (Table 1–5), we found that the gene expression level was negatively correlated with the oleic acid content in "Jike Soybean 20". Among the genes in "Jike Soybean 20", the negative correlation coefficient for GmFAD2-2a was the largest, followed by GmFAD2-2b. The negative correlation coefficient for GmFAD2-1 was the largest in "Ji Midou 3". The negative correlation coefficient for GmFAD2-1 was the largest in "Jike Midou No. 1", and that for GmFAD2-2b was the largest in "Jike Fresh Bean 1", indicating a significant negative correlation.

Table 1-5

Analysis of correlation between FAD2 gene expression and oleic acid during grain development

Gene	Jike Soybean 20	Ji Midou 3	Jike Midou No. 1	Jike Fresh Bean 1
GmFAD2-2	-0.639	-0.821	-0.099	-0.257
GmFAD2-1b	-0.172	-0.012	-0.025	-0.631
GmFAD2-1	-0.068	-0.892*	-0.921*	-0.528
GmFAD2-2a	-0.898*	-0.447	-0.151	-0.604
GmFAD2-2b	-0.875*	-0.689	-0.372	-0.919*
GmFAD2-3	-0.725	-0.757	-0.561	-0.718
GmFAD2-1B	-0.713	-0.896*	-0.780	-0.729

Analysis of the different expression levels of different genes of the same genotype on the same day (Table 1–6) showed that the levels of the GmFAD2-2b gene were significantly different from those of other genes in the varieties "Jike Soybean 20" and "Jike Fresh Bean 1". On the 25th, 30th, and 45th day of flowering, there were significant differences, and the expression level was the highest. The expression level of GmFAD2-1 in the varieties "Ji Midou 3" and "Jike Midou No. 1" was significantly different from those of other genes.

Table 1-6

Significance analysis of the difference in the expression amount of different genes in the same variety on the same day

Genotype	Gene	25th day after flowering	30th day after flowering	35th day after flowering	40th day after flowering	45th day after flowering
Jike Soybean 20	GmFAD2-2	0.94b	0.45d	0.44d	0.58b	0.53f
	GmFAD2-1b	0.68d	0.43c	1.17a	0.67a	0.62e
	GmFAD2-1	0.31e	0.52bc	0.36f	0.39d	0.71d
	GmFAD2-2a	0.97b	0.97a	0.74b	0.26f	0.68ef
	GmFAD2-2b	1.63a	0.84a	0.47de	0.34d	0.95a
	GmFAD2-3	0.85c	0.57b	0.58c	0.27e	0.77c
	GmFAD2-1b	1.01b	0.55b	0.35ef	0.46c	0.84b
Ji Midou 3	GmFAD2-2	0.83c	0.37 g	0.44c	0.51b	0.65 cd
	GmFAD2-1b	0.81c	0.55f	1.25a	0.55b	0.87b
	GmFAD2-1	0.97ab	1.54a	0.57b	0.72a	1.05a
	GmFAD2-2a	0.31d	0.68e	0.34d	0.59b	0.74b
	GmFAD2-2b	0.92b	1.16b	0.31 cd	0.33c	0.65c
	GmFAD2-3	1.01a	0.72d	0.14f	0.48c	0.58d
	GmFAD2-1B	1.02ab	0.95c	0.21e	0.33c	0.77b
Ji Midou No. 1	GmFAD2-2	0.91c	0.45d	0.25f	0.49d	0.65b
	GmFAD2-1b	0.98c	0.45c	0.42d	0.36e	0.77b
	GmFAD2-1	1.83a	1.27a	0.75c	0.59b	0.94a
	GmFAD2-2a	0.75d	0.12e	0.97a	0.69a	0.68c
	GmFAD2-2b	1.43b	0.38d	0.88b	0.11f	0.57d
	GmFAD2-3	0.97c	0.84b	0.24f	0.37e	0.59de
	GmFAD2-1B	0.67d	0.92b	0.33e	0.55c	0.45e
Jike Fresh Bean 1	GmFAD2-2	0.74f	1.05de	0.47f	0.73b	0.70e
	GmFAD2-1b	1.86de	1.61c	0.77c	0.25f	0.78e
	GmFAD2-1	2.04c	1.87b	0.66d	0.45d	1.44b
	GmFAD2-2a	1.93 cd	1.74c	0.56e	0.32e	0.86 cd

Genotype	Gene	25th day after flowering	30th day after flowering	35th day after flowering	40th day after flowering	45th day after flowering
	GmFAD2-2b	2.76a	2.06a	0.98b	0.77b	1.64a
	GmFAD2-3	2.42b	1.15d	1.25a	0.91a	0.87de
	GmFAD2-1B	1.76e	1.08e	0.39f	0.60c	0.971c
	Total average	33.54	25.14	16.03	13.72	22.03

Discussion

Soybeans are an important oil crop, and the oil content is valued throughout the world. Although research on the catalytic function of key enzymes in the soybean fatty acid biosynthesis pathway is relatively clear, research on the nature of expressed genes, especially FAD2, has been inconsistent; the accumulated information is relatively fragmented and lacks an understanding of the genetic basis and genetic behaviour of key enzyme genes that affect specific differences in the catalytic efficiency of different genes at the same locus, which limits the oil composition in soybean grains through joint regulation of each key enzyme gene. The content of oleic acid in soybean seeds and the ratio of fatty acid components are important reference factors for determining the quality of soybeans. However, pod development varies in different soybean receptor varieties due to differences in the seven known FAD2 genes in soybeans. A systematic analysis of the expression patterns at different stages has not been reported. Since the genes in this family have certain structural differences, there are also differences in function and expression regulation patterns. Oleic acid synthesis and that of other unsaturated fatty acids is a multigene regulatory chain reaction process. A single gene can not only affect the synthesis of oleic acid but also affect the synthesis of other unsaturated fatty acids.

This study found that FAD2 is affected by low temperature and directly affects the fluidity of plant cell plasma membranes. Unsaturated fatty acids are important components of plant cell plasma membranes and play an important role in the normal growth of plants. High plant cell plasma membrane fluidity can help plants resist low temperature and other adverse conditions^[32–34]. In 2012, Byfield^[35] found that in the process of soybean germination, the expression level of FAD2 was significantly higher at low temperature than at normal temperature, while at high temperature, the expression level was significantly decreased, and at the same time, the oleic acid content was increased.

Conclusions

In this study, the correlation between FAD2 expression and oleic acid content during the pod development showed that FAD2 plays an important role in the process of oleic acid synthesis, and each gene in the family functions differently. In 2015, Li^[36] found that oleic acid was undetectable in the early stage of soybean pod development but could be detected 16 days after flowering, and in the later stage, the content increased greatly. Oleic acid synthesis is a process involving multiple genes and co-regulators, and thus, inhibiting the expression of multiple genes simultaneously to change the oleic acid content,

expand the scope of its promotion, and select for higher oleic acid content can lead to novel and beneficial soybean germplasm.

Methods

FAD2 family gene bioinformatics analysis

Annotation of FAD2 family genes

A conserved structure was identified based on the National Center for Biotechnology Information (NCBI) database, the soybean database online website soybase (<https://soybase.org/>) and the Pfam (<http://pfam.xfam.org/>) database. Gene location on chromosomes and the conserved domains in seven genes in the soybean FAD family were identified using Hmmer 3.0 software.

FAD2 gene evolutionary tree analysis and sequence alignment

The online program Clustal X (version 3.0) was used to align the sequences of the soybean FAD2 family, and a phylogenetic tree was constructed using MEGA 5.0 software for all functionally defined FAD2 genes in soybean species and between species to classify genes. The neighbour-joining (NJ) method was used to analyse the results.

Analysis of FAD2 gene protein sequences

The online website MEME (<http://meme.nbcr.net/meme>) was used to predict the structure and classification of other protein sequences in addition to the conserved sequences in the gene family. The protein sequences of all the genes in the family were entered in the MEME website, and the parameters were all set to the default parameters.

Analysis of gene expression patterns

Experimental materials

Plant Materials

The soybean genotypes "Jike Soybean 20", "Ji Midou 3", "Jike Fresh Bean 1", and "Jike Midou No. 1", among which the "Jike Fresh Bean 1" oleic acid content is relatively low, while the other genotypes have relatively high oleic acid content, were selected and provided by the Plant Biotechnology Center of Jilin Agricultural University.

Required reagents

RNA extraction kit, real-time PCR kit, DNA standard molecular weight marker, and DL2000 kits were from TaKaRa. Fluorescent quantitative primers were synthesized by Changchun Kumei Biotechnology Co.(Table S1)., Ltd. The plant RNA extraction kits were purchased from Omega. Sodium chloride was

purchased from Thermo Fisher Company, and the other major reagents were obtained domestically and were of pure analytical grade.

Instruments and equipment

A fluorescence quantitative polymerase chain reaction kit (purchased from the United States), cryogenic freezing centrifuge at 4 °C, cryogenic – 80 °C refrigerator, water bath heating pot, and indoor plant culture room were utilized.

Low Temperature Stress Treatment

Clean and full "Jike Soybean 20" soybean seeds were chosen and planted in small plastic buckets with a puncture in the bottom to allow for easy draining. The planted seeds were placed in an artificial climate chamber that simulated the growth environment under natural conditions and were allowed to grow for approximately 2 weeks after germination. When the first three leaves were present, a series of abiotic stress treatments were carried out on the seedlings. Abiotic stress treatment of plants were conducted using the following methods.

Low temperature stress: 20 days after germination, the seedlings were placed in a refrigerator at 4 °C for cold treatment. After treatment for 0 h, 1 h, 3 h, 6 h, 12 h, 24 h, and 48 h, root, stem, leaf and germinated seed parts were taken. The control plants treated as described above were watered every other day. In the treatment process, replicates were prepared for each treatment group to reduce the error generated during the experiment. After the samples were collected, they were placed in a -80 °C refrigerator for storage.

Determination of oleic acid content and relative expression of FAD2 genes at different stages of grain development

The soybean test strains "Jike Soybean 20", "Ji Midou 3", "Jike Fresh Bean 1" and "Jike Midou No. 1" were sown in the experimental field of Jilin Agricultural University, planted at the end of April, harvested at the end of September, and observed for the flowering period; the first sampling was performed on the 25th day after flowering. Every 5 days, the young pods in development were taken, placed in liquid nitrogen immediately, and then stored in a refrigerator at -80 °C. A total of 5 young pods were taken, and 3 replicate samples were used. A DS2500 grain analyser was used to measure the change in oleic acid content at different periods of grain development. Primer5.0 software was used to design fluorescent quantitative primers (Table S1) to quantify the genes in the soybean FAD2 family. The expression levels in pods at different stages of development were determined.

Real-time PCR

In this experiment, cDNA of the internal reference gene β -actin and the target gene was amplified by a two-step method using a fluorescence quantitative analyser, and three replicate experiments were performed for each set of samples. Following the instructions in the Fluorescence Quantitation Kit, the qRT-PCR reaction mixture was prepared as follows (operating on an ice box). To reduce error caused by the pipetting gun during the loading and ensure accuracy, the reaction mixture was uniformly mixed with a pipette and then dispensed into a sterile PCR tube, and finally, reverse transcription was carried out. The cDNA sample was added, centrifuged and placed in the qRT-PCR thermocycler (Table 1–7).

Table 1-7
PCR System

Name	System
2 × qPCR Mix	10 μL
Primer1	2 μL
Primer2	2 μL
RNase-free H ₂ O	4 μL
cDNA	2 μL

The amplification conditions were predenaturation at 95 °C for 10 min, denaturation at 95 °C for 10 s, annealing at 55 °C for 20 s, and extension at 72 °C for 15 s; set to 55 cycles, and the relative expression of the target gene was calculated using the $2^{-\Delta\Delta C_t}$ method.

Determination of oleic acid content in soybean seeds with the near-infrared grain analysis

In this experiment, an NIRS DS2500 grain analyser was used to select the clean grain to be tested, which was put into a measuring cup. The seed volume needed to reach the bottom of the cup and cover the infrared scanning area (If the seed quantity is small, the seed can be ground into powder before being placed in the measuring cup). The sample was placed into the sample tank of the near-infrared grain analyser, according to the previously established laboratory procedure for acquisition of spectra to determine soybean oleic acid content. Analysis of each sample was repeated 3 times, and the measurement results were determined by the software Operator and automatically saved in the computer. The error between the results measured by the near-infrared grain analyser and the gas chromatography method was previously verified to be between 1% and 1.5%, which is within the acceptable range, indicating that the results of the near-infrared detection are authentic and can be widely applied to fatty acids.

Data analysis

All data in this experiment were analyzed using SPSS2.0 software. Each group of experiments was set up to three replicates to reduce experimental errors. ANOVA was used to analyze the correlation between the difference in FAD2 genes expression and changes in oleic acid content among different soybean varieties. Correlation coefficient.

Abbreviations

qRT-PCR:Quantitative reverse transcription-polymerase chain reaction; RCR:Relative consumption rate.

Additional Files

Additional file 1: Table S1. Primers used for the experiments in this study.(pdf 24kb)

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

All the data or materials used during this study are available from the corresponding author upon reasonable request.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

YMH and WN designed the experiments. WN, JL and YXM planned and performed the experiments. YXB and CDY edited the manuscript. All authors discussed the results and commented on the manuscript. All authors have read and approved the manuscript

Funding

Expression pattern of soybean fatty acid desaturase gene *GmFAD2* On the Mechanism of Biosynthesis with Soy Oleic Acid—Funded by the Education Department of Jilin Province.

Acknowledge

We are grateful for the generous grant from the Jilin Agricultural Science and Technology University that made this work possible.

Not applicable.

Author details

1 Jilin Agricultural Science and Technology University, Jilin, 132101, China

References

1. Fu-liang Cao, Huan-li Wang, Wan-wen Yu, Hua Cheng. Research progress on fatty acid desaturases and coding genes in higher plants[J].Journal of Nanjing Forestry University(Natural Science).2012;36:125-132.
2. Jung S,Powell G, Moore K, Abbott A.The high oleate trait in the cultivated peanut(*Arachis hypogaea* L.).II. Molecular basis and genetics of the trait[J].Mol Gen Genet. 2000;263:806-811.
3. Pirle IL,Kongcharoensuntorn W, Nampaisansuk M, Knesek J E, Chapman K D, Pirtile R M, Molecular cloning functional expression of the gene for a cotton Δ -12 fatty acid desaturase (FAD2)[J].Biochim

Biophys Acta.2011;1522:122-129.

4. Heppard E, Kinney A, Stecca K, Miao G. Development and growth temperature regulation of two different microsomal w-6 desaturase genes in soybean[J]. Plant Physiol.1996; 110:311-319.
5. William Haun. Improved soybean oil quality by targeted mutagenesis of the fatty acid desaturase 2 gene family. Plant Biotechnology Journal.
6. Wenzhi Jiang. Significant enhancement of fatty acid composition in seeds of the allohexaploid, *Camelina sativa*, using CRISPR/Cas9 gene editing.Plant Biotechnology Journal. 2017;15:648–657.
7. Li Yan. Creation of soybean GmWRI1a gene mutant using CRISPR/Cas9 technology [D]. Northeast Agricultural University, 2017.
8. Wilson R F, Schlueter. The role of genomics and biotechnology in achieving global food security for high-oleic vegetable oil [J]. Journal of Oleo Science. 2012; 61: 357-367.
9. Pirtle IL, Kongcharoensuntorn W, Nampaisansuk M, et al. Molecular cloning and functional expression of the gene for a cotton Δ -12 fatty acid desaturase (FAD2)[J]. Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression. 2001;1522:122-129.
10. Liu Q, Brubaker CL, Green AG, et al. Evolution of the FAD2-1 fatty acid desaturase 5' UTR intron and the molecular systematics of *Gossypium* (Malvaceae)[J]. American Journal of Botany. 2001;88:92-102.
11. Zhang DY, Pirtle IL, Park SJ, et al. Identification and expression of a new delta-12 fatty acid desaturase (FAD2-4) gene in upland cotton and its functional expression in yeast and *Arabidopsis thaliana* plants[J]. Plant Physiol Biochem.2009;47:462-471.
12. Wagner N, Mroczka A, Roberts P D, et al. RNAi trigger fragment truncation attenuates soybean FAD2-1 transcript suppression and yields intermediate oil phenotypes [J]. Plant Biotechnology Journal. 2011;9: 723-728.
13. Xingfen Miao, Wenping Xu, Candong Li , et al.Analysis of fatty acid composition and content of soybean varieties in Northeast China[J]Soybean Science.2011;30(3):529-531.
14. Qiu D , Vuong T , Valliyodan B , et al. Identification and characterization of a stachyose synthase gene controlling reduced stachyose content in soybean[J]. Theoretical and Applied Genetics. 2015;128:2167-2176.
15. La T C , Pathan S M , Vuong T , et al. Effect of High-Oleic Acid Soybean on Seed Oil, Protein Concentration, and Yield[J]. Crop Science. 2014;54:2054-2057.
16. Cheng Shi, Zhenfeng Jiang, Wenbin Li. Correlation and path analysis of soybean oil content and yield traits under different genetic backgrounds and environmental conditions. Soybean Science.2012; 31: 744-748.
17. Yongzhan Zheng, Wei Gai, Tuanjie Zhao , et al. Study on Variation Characteristics of Fat Traits in Soybean Cultivation and Wild Resources in China[J]. Chinese Journal of Agricultural Sciences.2008; 41: 1283-1290.

17. Xiaofeng Dai. Cloning of key genes for fatty acid synthesis in rapeseed and accumulation of fatty acids [D]. Chinese Academy of Agricultural Sciences, 2014.
18. Xiu-qin Luo, Wen-jun Ou, Kai-mian Li, Song-bi Chen. Structure and function prediction of cold-resistant protein stearyl-ACP desaturase[J].Journal of Fujian Agricultural and Forestry University.2014;43:484- 489.
19. Rui Zhang, Yihui Zhang, Dan Shao, Laijun Wang, Daoqing Gong.The function and regulation of stearyl-CoA desaturase gene[J].Chinese Journal of Life Sciences.2013;25:378-382.
20. Junming Sun, Pinkia Han, Shurong Yan, et al. Rapid determination of fatty acid content in soybean by Fourier near-infrared reflectance spectroscopy [J]. Spectroscopy and Spectral Analysis.2010;28: 1290-1295.
21. Fukuan Yu, Junming Sun, Pinkia Han, et al. Comparative Analysis of Rapid Determination Methods of Fatty Acid Composition in Soybean Seeds[J].Soybean Science.2011;30: 626-631.
22. Shi Cheng, Xin Wang, Baolin Liu. Study on Low Field Nuclear Magnetic Resonance Relaxation of Oleic Acid Glyceride System[J].Journal of Spectroscopy.2018;35:243-254.
23. Yingjun Zhang, Huimin Gao, Chunzhi Jiang, et al. Rapid determination of fatty acid content in soybean seeds [J]. Soybean Science. 2008;27: 859-862.
24. Chai Y H, Tan K Z. Measurement of soybean fatty acid content by near infrared spectroscopy [J]. Transactions of the CSAE. 2008; 23:238-241.
25. Wei Wang, Li Liu, Sujie Fan, Xiao Han, Wei Wu, Yiwu Wang, Dan Yao, Jun Zhang.The establishment of near-infrared model of soybean fatty acid content[J].Soybean Science.2017;36:295-299.
26. Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J.Genome sequence of the palaeopolyploid soybean. Nature.2010; 463: 178~183.
27. Xie K, Zhang J, Yang Y. Genome-wide prediction of highly specific guide RNA spacers for CRISPR-Cas9-Mediated genome editing in model plants and major crops. Mol Plant.2014;7: 923~926.
28. Gao Y, Zhao Y. Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. J Integr Plant Biol. 2014; 56: 343~349.
29. Nissim L, Perli SD, Fridkin A, Perez-Pinera P, Lu TK.Multiplexed and programmable regulation of gene networks with an integrated RNA and CRISPR/Cas toolkit in human cells. Mol Cell. 2014; 54: 698~710.
30. Collier R, Fuchs B, Walter N, Kevin Lutke W, Taylor CG.Ex vitro composite plants: an inexpensive, rapid method for root biology. Plant J.2005;43: 449~457.
31. Kereszt A, Li D, Indrasumunar A, Nguyen CD, Nontachaiyapoom S, Kinkema M, Gresshoff PM. Agrobacterium rhizogenes-mediated transformation of soybean to study root biology. Nat Protoc.2007; 2: 948~952.
32. Ron M, Kajala K, Pauluzzi G, Wang D, Reynoso MA, Zumstein K, Garcha J, Winte S, Masson H, Inagaki S Hairy root transformation using Agrobacterium rhizogenes as a tool for exploring cell

- type-specific gene expression and function using tomato as a model. *Plant Physiol.* 2014;166: 455~469.
33. Dolatabadian A, Sanavy SAMM, Ghanati F, Gresshoff PM. *Agrobacterium rhizogenes* transformed soybean roots differ in their nodulation and nitrogen fixation response to genistein and salt stress. *World J Microb Biot.* 2013; 29: 1327~1339.
34. Li H, Deng Y, Wu T, Subramanian S, Yu O. Misexpression of mi R482, mi R1512, and mi R1515 increases soybean nodulation. *Plant Physiol.* 2010;153: 1759~1770.
35. Sun X, Hu Z, Chen R, et al. Targeted mutagenesis in soybean using the CRISPR/Cas9 system[J]. *Scientific reports.* 2015;5:10342.

Figures

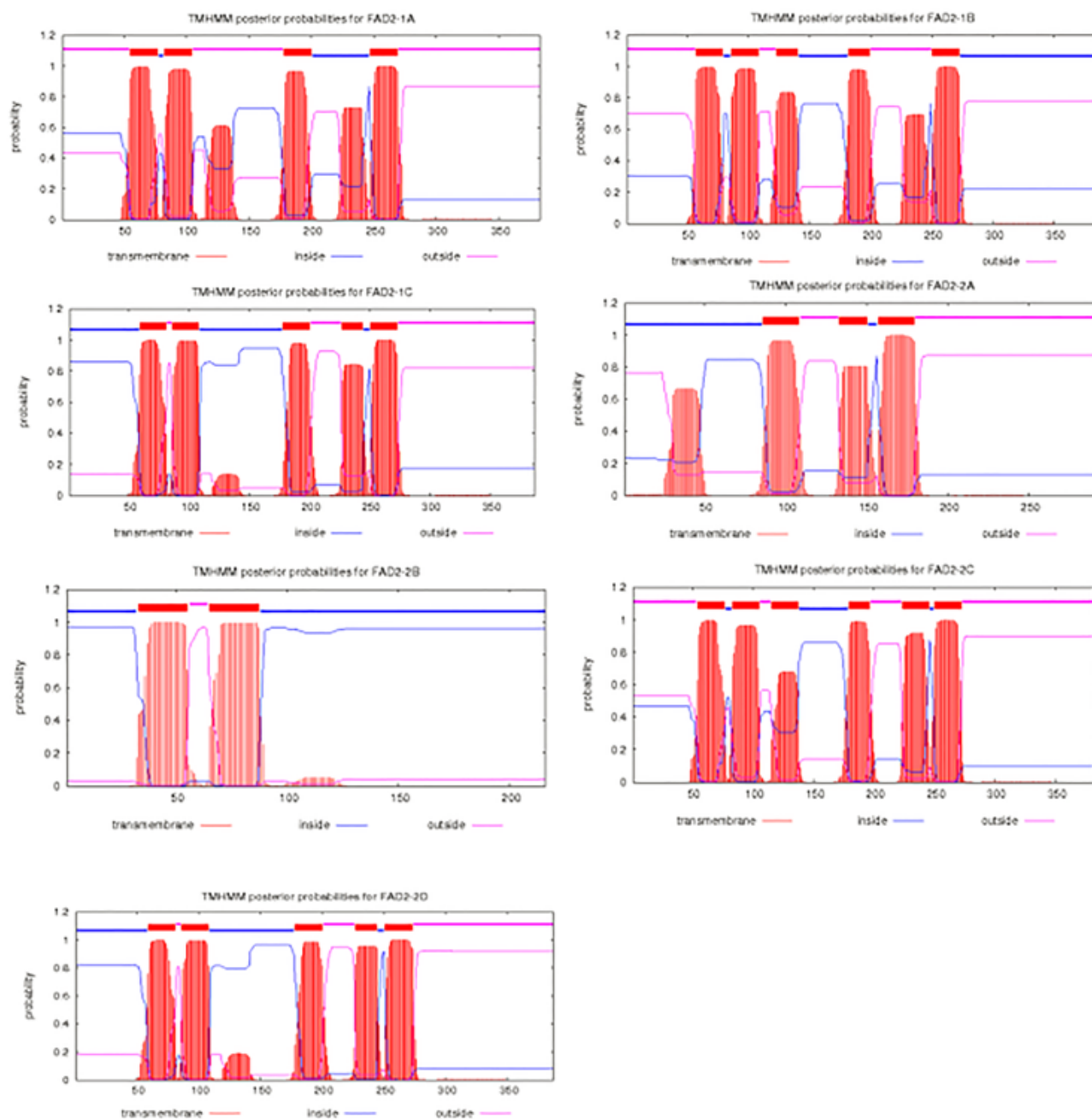


Figure 1

Analysis of transmembrane structure of gene proteins.

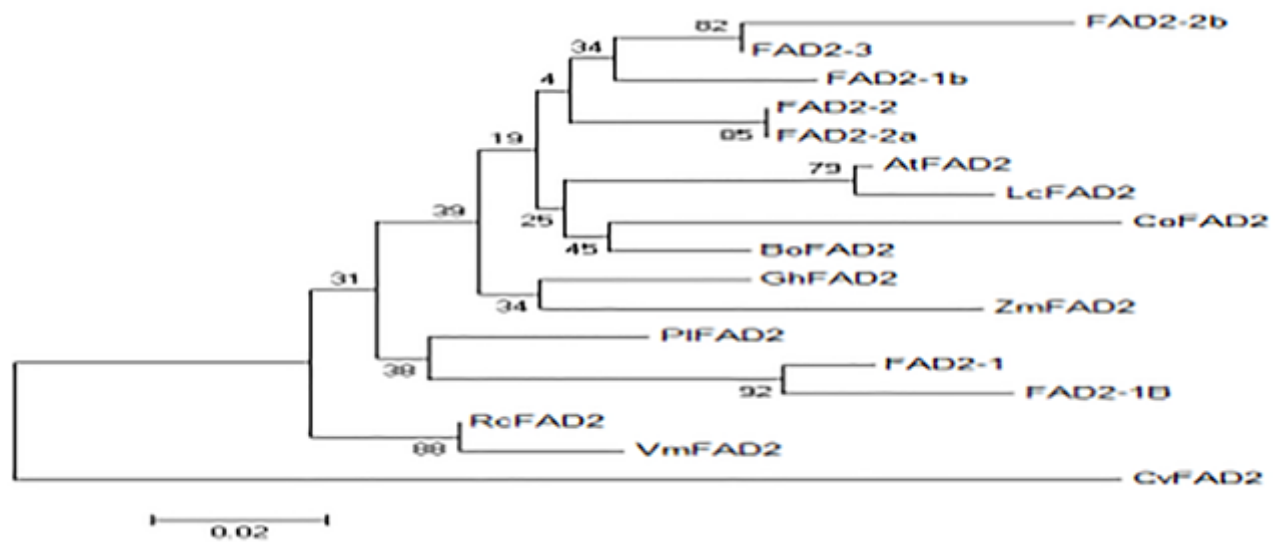


Figure 2

Phylogenetic analysis of FAD2 genes among species



Figure 3

Structural comparison of introns and exons in different genes

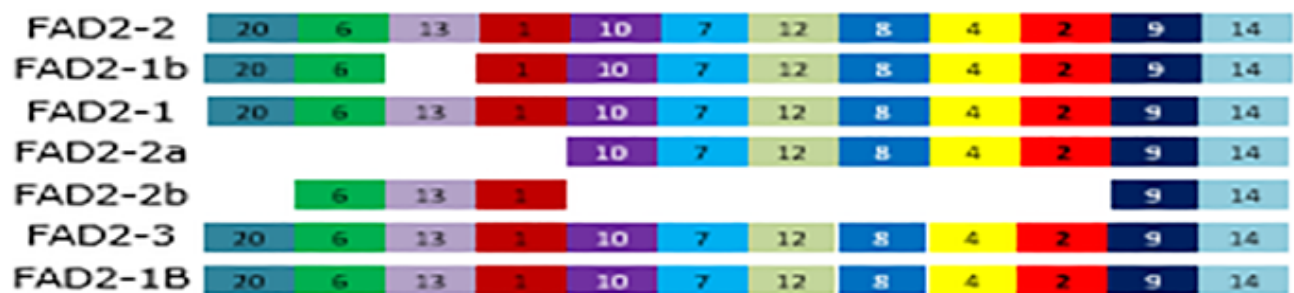


Figure 4

Analysis of protein sequences encoded by the genes

FAD2-2 1 ----MGAGC^RRTDVPPANRKSEVD^PLKRV^FFK^PPF^TLSQ^KKK^VIPPHCFORS^VLRFSFY

FAD2-2a 1 -----

FAD2-3 1 ----MGAGC^RRTDVPPANRKSEAD^PLKRV^FFK^PPF^TLSQ^KKK^VIPPHCFORS^VLRFSFY

FAD2-1b 1 ---MGGGGRSSATLKHQNS^KKNHKK^RRVPHAK^PPF^TLSQ^KKK^VIPPHCFORS^VLRFSFY

FAD2-2b 1 -----MGDTM^KRV^FFK^PPF^TLSQ^KKK^VIPPHCFORS^VLRFSFY

AtFAD2 1 -----M^GAGGRMD^PAPKSTD^RVP^SAK^PPF^TLSQ^KKK^VIPPHCFORS^VLRFSFY

FAD2-1 1 MGLAKETTM^GGR^RRVAKVE^VQ^GKK^PLSRV^FNT^KPP^FTV^GQL^KKAIPPHCFORS^VLRFSFY

FAD2-1B 1 MGLAKETIM^GGG^RRVAKVE^VQ^GKK^PLSRV^FNT^KPP^FTV^GQL^KKAIPPHCFORS^VLRFSFY

FAD2-2 56 VVYDLTIAFCL^VY^ATH^VFHLL^PQ^FLS^FAWPIYWAVQG^CIL^TGVWVIA^HECGHH^AFSD

FAD2-2a 1 -----AVOGC^IL^TGVWVIA^HECGHH^AFSD

FAD2-3 56 VVYDLTIAFCL^VY^ATH^VFHLL^PQ^FLS^FAWPIYWAVQG^CIL^TGVWVIA^HECGHH^AFSD

FAD2-1b 58 VVYDLTIAFCL^VY^ATH^VFHLL^PQ^FLS^FAWPIYWAVQG^CIL^TGVWVIA^HECGHH^AFSD

FAD2-2b 40 VVYDLTIAFCL^VY^ATH^VFHLL^PQ^FLS^FAWPIYWAVQG^CIL^TGVWVIA^HECGHH^AFSD

AtFAD2 50 VVMDL^SFS^SLL^VIT^TQ^HIL^LFLARE^MAWAL^VYW^VQG^CIL^TGVWVIA^HECGHH^AFSD

FAD2-1 61 VVYDL^SFAH^IY^ATH^VFHLL^PQ^FLS^FAWPIYWAVQG^CIL^TGVWVIA^HECGHH^ASK

FAD2-1B 61 VVYDL^SFAH^IY^ATH^VFHLL^PQ^FLS^FAWPIYWAVQG^CIL^TGVWVIA^HECGHH^ASK

FAD2-2 115 YQL^LDD^VGL^VLHS^LLLVPYFSWKY^SSHRRH^SNTGSLERDEVFVPKQ^SCK^VWYSKYLN^N

FAD2-2a 25 YQL^LDD^VGL^VLHS^LLLVPYFSWKY^SSHRRH^SNTGSLERDEVFVPKQ^SCK^VWYSKYLN^N

FAD2-3 115 YQL^LDD^VGL^VLHS^LLLVPYFSWKY^SSHRRH^SNTGSLERDEVFVPKQ^SCK^VWYSKYLN^N

FAD2-1b 118 H^CW^LDD^VGL^VLHS^LLLVPYFSWKY^SSHRRH^SNTGSLERDEVFVPK^SCK^VWYSKYLN^N

FAD2-2b 99 YQL^LDD^VGL^VLHS^LLLVPYFSWKY^SSHRRH^SNTGSLERDEVFVPKQ^S-----

AtFAD2 109 Y^SW^LDD^VGL^VLHS^LLLVPYFSWKY^SSHRRH^SNTGSLERDEVFVPK^SCK^VWYSKYLN^N

FAD2-1 119 Y^CW^LDD^VGL^VLHS^LLLVPYFSWKY^SSHRRH^SNTGSLERDEVFVPK^SCK^VWYSKYLN^N

FAD2-1B 119 Y^PW^LDD^VGL^VLHS^LLLVPYFSWKY^SSHRRH^SNTGSLERDEVFVPK^SCK^VWYSKYLN^N

FAD2-2 175 PPGRVLTAVTLTLGWPLY^LALNVSGRPYDRFACHYDPYGPIYS^DRERLQIYISDAG^VLA

FAD2-2a 85 PPGRVLTAVTLTLGWPLY^LALNVSGRPYDRFACHYDPYGPIYS^DRERLQIYISDAG^VLA

FAD2-3 175 PPGRVLTAVTLTLGWPLY^LALNVSGRPYDRFACHYDPYGPIYS^DRERLQIYISDAG^VLA

FAD2-1b 178 SPGRVLTALTLTLGWPLY^LALNVSGR^SY^RFACHYDPYGPIYS^DRERLQIYISDAG^VLA

FAD2-2b 148 -----

AtFAD2 169 PPGRVLTAVTLTLGWPLY^LALNVSGR^AY^RFACHYDPYGPIYS^DRERLQIYISDAG^VLA

FAD2-1 179 PLGRAASTLTLTLGWPLY^LALNVSGRPYDRFACHYDPYGPIYS^DRERLQIYISDAG^VLA

FAD2-1B 179 PLGRAASTLTLTLGWPLY^LALNVSGRPYDRFACHYDPYGPIYS^DRERLQIYISDAG^VLA

FAD2-2 235 VCYGL^RLA^MAKGL^AWVVCYGVPLLVNGLVLTITLQHTHPALPHY^TSEWDWL^RGAL

FAD2-2a 145 VCYGL^RLA^MAKGL^AWVVCYGVPLLVNGLVLTITLQHTHPALPHY^TSEWDWL^RGAL

FAD2-3 235 VCYGL^RLA^MAKGL^AWVVCYGVPLLVNGLVLTITLQHTHPALPHY^TSEWDWL^RGAL

FAD2-1b 238 VCYGL^RLA^MAKGL^AWVVCYGVPLLVNGLVLTITLQHTHPALPHY^TSEWDWL^RGAL

FAD2-2b 148 -----

AtFAD2 229 ALY^AY^RASH^AY^CTH^VVVCYGVPLLVNGLVLTITLQHTHPALPHY^TSEWDWL^RGAL

FAD2-1 239 VY^SLY^RLA^MAKGL^AWVVCYGVPLLVNGLVLTITLQHTHPALPHY^TSEWDWL^RGAL

FAD2-1B 239 VY^SLY^RLA^MAKGL^AWVVCYGVPLLVNGLVLTITLQHTHPALPHY^TSEWDWL^RGAL

FAD2-2 295 ATVDRDYGLNKVFHNITD^THVAHH^LFSTMPHYHAMEATKAIPILGEYYRFD^ETP^VKA

FAD2-2a 205 ATVDRDYGLNKVFHNITD^THVAHH^LFSTMPHYHAMEATKAIPILGEYYRFD^ETP^VKA

FAD2-3 295 ATVDRDYGLNKVFHNITD^THVAHH^LFSTMPHYHAMEATKAIPILGEYYRFD^ETP^VKA

FAD2-1b 298 ATVDRDYGLNKVFHNITD^THVAHH^LFSTMPHYHAMEATKAIPILGEYYRFD^ETP^VKA

FAD2-2b 148 -----STH^VVHLL^FSTMPHYHAMEATKAIPILGEYYRFD^ETP^VKA

AtFAD2 289 ATVDRDYGLNKVFHNITD^THVAHH^LFSTMPHYHAMEATKAIPILGEYYRFD^ETP^VKA

FAD2-1 299 ATVDRDYGLNKVFHNITD^THVAHH^LFSTMPHYHAMEATKAIPILGEYYRFD^ETP^VKA

FAD2-1B 299 ATVDRDYGLNKVFHNITD^THVAHH^LFSTMPHYHAMEATKAIPILGEYYRFD^ETP^VKA

FAD2-2 355 MWREARECIYVEPD^QSE^SKG^VFWYNNK^L

FAD2-2a 265 MWREARECIYVEPD^QSE^SKG^VFWYNNK^L

FAD2-3 355 MWREARECIYVEPD^QSE^SKG^VFWYNNK^L

FAD2-1b 358 MWREARECIYVEPD^QSE^SKG^VFWYNNK^L

FAD2-2b 190 MWREARECIYVEPD^QTEN⁻KG^VFWYNNK^L

AtFAD2 349 MWREARECIYVEPD^QKKG⁻EG^VFWYNNK^L

FAD2-1 359 MWREARECIYVEPD^QSE^SKG^VFWYNNK^L

FAD2-1B 359 MWREARECIYVEPD^QSE^SKG^VFWYNNK^L

Figure 5

FAD2 amino acid sequence in *Arabidopsis thaliana* and soybean



Figure 6

Relative expression of FAD2 genes in different plant parts under low temperature stress

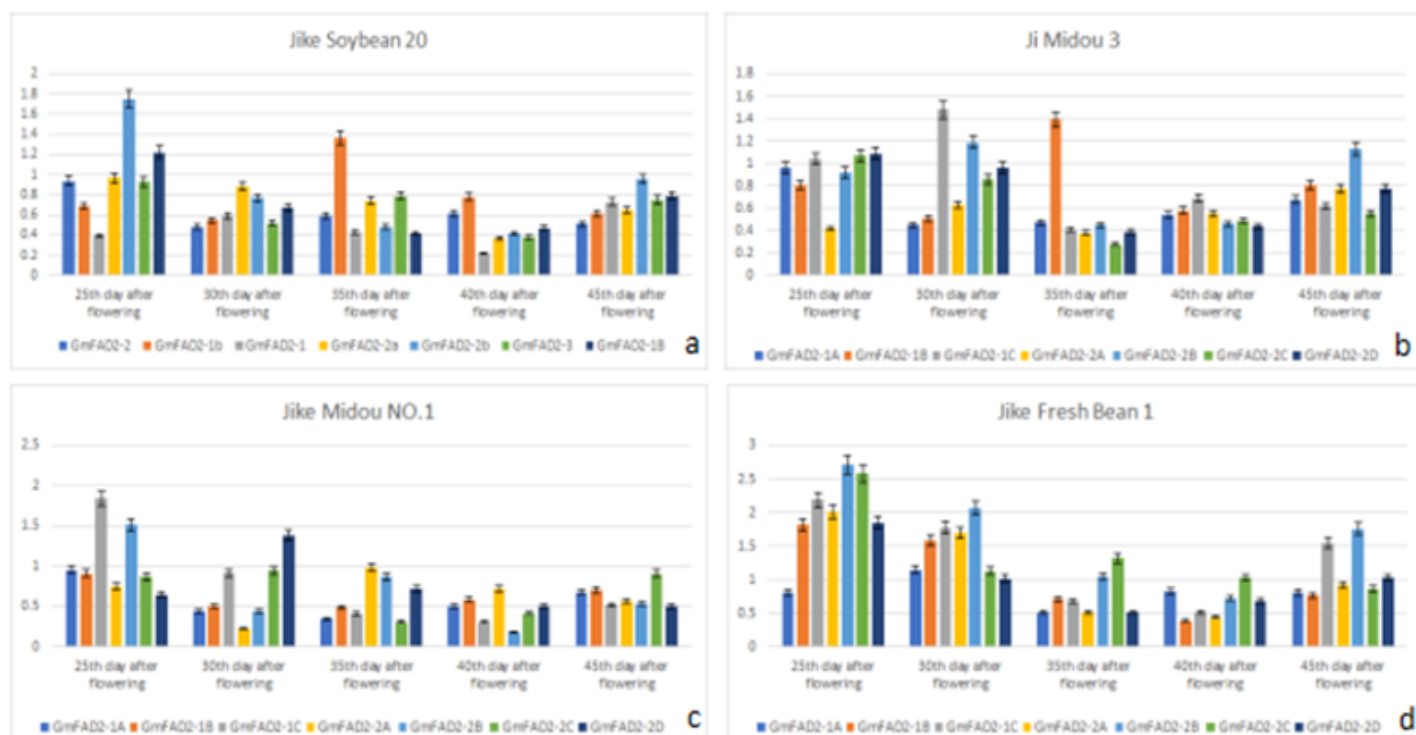


Figure 7

Variation in gene expression in different cultivars at different post-anthesis stages

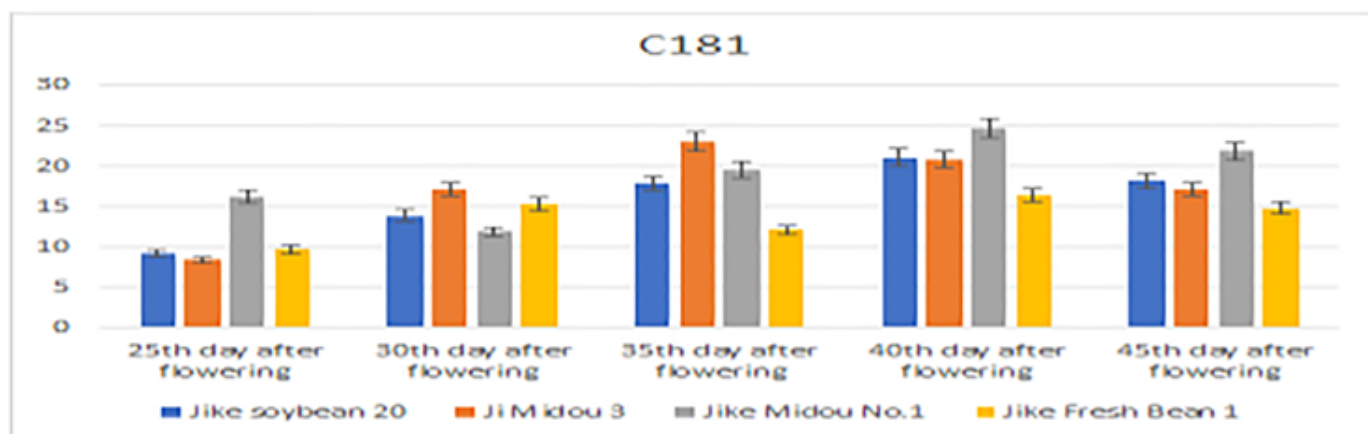


Figure 8

Changes in oleic acid content in different soybean varieties at different stages after flowering