



ARTICLE

Correlation Analysis of New Soybean [*Glycine max* (L.) Merr] Gene Gm15G117700 with Oleic Acid

Shuo Qu¹, Yaolei Jiao², Lamboro Abraham² and Piwu Wang^{1,*}

¹Plant Biotechnology Center, Jilin Agricultural University, Changchun, 130018, China

²College of Agronomy, Jilin Agricultural University, Changchun, 130018, China

*Corresponding Author: Piwu Wang. Email: peiwuw@163.com

Received: 01 December 2020 Accepted: 26 January 2021

ABSTRACT

The fatty acid dehydrogenase gene plays an important role in regulating the oleic acid content in soybean. Genome-wide association study screened out soybean oleic acid related gene Gm15G117700. A fragment size of 693bp was obtained by PCR amplification of the gene and, it was connected by seamless cloning technology to the pMD18T cloning vector. Based on the gene sequence cloned, bioinformatic analysis of gene protein was performed. The overexpression vector of Gm15G117700 and the CRISPR/Cas9 gene editing vector were constructed. The positive plants were obtained by *Agrobacterium*-mediated transformation of soybean cotyledon nodes and T₂ plants were identified by conventional PCR, QT-PCR and Southern blot hybridization. 10 copies of high and low oleic acid seeds were selected for QT-PCR to identify the expression content of Gm15G117700 gene in different soybeans, and finally near-infrared spectroscopy analyzer was used to identify the oleic acid quality of soybeans. T₂ RT-PCR identification showed that overexpression was reduced by 3.94%, and gene editing was increased by 3.49%. It is determined that the Gm15G117700 gene may belong to a regulatory gene, a minor gene that can promote the conversion to linoleic acid content in soybean oleic acid synthesis. The gene cloning and its functional verification was not reported yet. This is the first report by PCR amplification of soybean Gm15G117700 genes and gene expression vector. Improving the content of oleic acid in soybean lay a foundation for researchers. Therefore, this study clearly identified the function of soybean Gm15G117700 gene and its role played in oleic acid synthesis and metabolism.

KEYWORDS

Soybean; Gm15G117700 genes; cloning; expression vector

1 Introduction

Soybean [*Glycine max* (L.) Merr] is one of the economically important oil crop with high quality vegetable oil. China is considered to be the origin and diversity of soybeans [1]. Many of the vegetable oils and protein ingested by the human body are derived from soybeans. With the improvement of the economy and quality of life, people are not only pursuing enough food, but also focusing on quality and health. Therefore, people's pursuit and demand for high quality soybean are increasing. So, cultivating high quality soybean in soybean breeding is one of the important goal. Soybean oil is one of the main edible oils in the world, accounting for 1/3 of the global total [2]. Soybean mainly contains five major



fatty acids, almost constituting all fats and oils. Its composition consists of 12% palmitic acid (16:0), 4% stearic acid (18:0), 21% oleic acid (18:1), 53% linoleic acid (18:2) and 8% linolenic acid (18:3) [3]. Unsaturated fatty acids mainly contain linoleic acid, linolenic acid, and oleic acid whereas saturated fatty acids mainly contain stearic acid and palmitic acid. Saturated fatty acids are less healthy for the human body and are not easily absorbed by the human body, causing obesity and cardiovascular disease [4]. Linolenic acid and linoleic acid have poor stability and weak oxidation resistance. They are easy to be oxidized when exposed to air and high temperature processing. Their nutritional value is affected, which reduces the quality of oleic acid. In industry, hydrogenation can be used to avoid this phenomenon. The side effects are trans fatty acids, which are harmful to health [5]. Oleic acid is stable in nature and has strong antioxidant capacity. The proportion of fatty acids was adjusted to increase unsaturated fatty acids and decrease saturated fatty acids. In unsaturated fatty acid, the content of oleic acid increases, and the content of linoleic acid and linolenic acid decreases, which will improve the quality of soybean fatty acid. At present, the improvement of soybean oleic acid content has become one of the most important goals in breeding. A large number of studies have shown that oleic acid can effectively reduce blood pressure and cholesterol content in plasma [5–7]. As Mediterranean people have more foods rich in oleic acid in their diet, the incidence of cerebrovascular disease (CVD) is much lower than that of people in European countries [8,9], so it can be seen that long-term consumption of oleic acid can improve the incidence of CVD.

The biosynthesis of fatty acids takes place jointly in the cytoplasm and endoplasmic reticulum in plants, and mainly in the cytoplasm in animals [10]. Sucrose as a carbon source for biosynthesis, through EMP (Embden-Meyerhof-Parnas pathway) approach breaks down sugar into important intermediate metabolites. Pyruvic acid, the process takes place in the cytoplasm. After the catalysis of pyruvate dehydrogenase form the three important energy carbohydrate, protein, lipid intermediate metabolite acetyl CoA. Acetyl CoA as a precursor of the synthesis of fatty acids, transported to the cytoplasm corresponding position. Acetyl CoA carboxylase—catalytic formation of Malonyl-CoA under the effect of fatty acid synthase, catalytic synthesis of acyl chain, is a the process every time increase 2 carbon. The process is continuous, until 16 to 18 carbon (16:0/18:0) catalyzed by desaturase. Saturated fatty acids generate unsaturated fatty acids like oleic acid, linoleic acid and linolenic acid. Under the effect of Δ^12 -fatty acid dehydrogenase catalyzed, oleic acid (18:1 ^{$\Delta^{9,12}$}) dehydrogenase polyunsaturated fatty acid linoleic acid, linoleic acid in Δ^{15} -fatty acid dehydrogenase generated linolenic acid under catalysis (18:3 ^{$\Delta^{9,12,15}$}) [11]. Stearic acid in soybean fatty acids is the precursor of oleic acid, which is regulated by the only soluble desaturase, and it is stearoyl-ACP-desaturase (SACPD) gene family. SACPD plays a major role in metabolic regulation [12] and can increase the content of oleic acid through overexpression of SACPD gene family. In fatty acid biosynthesis, the catalytic conversion of stearic acid (18:0) to oleic acid (18:1) by the protein encoding stearoyl (SACPD), the defects of *npr1–5 (ssi2)* mutant and the ASACPD gene encoding the salicylate insensitive inhibitor have been confirmed in a variety of plant genomes [13,14]. Linoleic acid is the downstream of oleic acid biosynthesis, which is controlled by the main regulatory gene fatty acid dehydrogenase (FAD2) gene family. The inhibition of the gene of FAD2 gene family can achieve the purpose of increasing oleic acid content. The variation of oil content in soybean is subject to genetic additive effect, and there are both superior and dominant effects [15].

In the laboratory early in the spring, 260 soybean germplasm resources for whole genome sequencing association analysis (GWAS) [16] were used to select Gm15G117700 genes involved in soybean oleic acid. Oleic acid biological metabolism of the gene of the mechanism is not clear. By using agrobacterium mediated transformation of soybean cotyledon section was restrain the carrier, and the expression vector into soybean to obtain T₂ seeds and verified by the conventional PCR, Southern blot hybridization and QT-PCR. 10 copies of high and low oleic acid seeds each was used for QT-PCR to identify the level of Gm15G117700 gene expression in different soybeans. Near infrared spectrum analyzer was used to measure soybean oleic acid content change, to explore the soybean Gm15G117700 gene function. The effect on the quality of the oleic acid modified laid a certain foundation.

2 Materials and Methods

2.1 Plant Materials and Carriers

In this experiment, plant materials such as soybean JN28, JN38, *Escherichia coli* DH5, Agrobacterium EHA105 and overexpressing pCAMBIA-3301 were used which were provided by the Biotechnology Center of Jilin Agricultural University.

2.2 Experimental Reagents

The main reagents and their commercial sources were listed in [Tab. 1](#). Gene primers for synthesis and sequencing were provided by Kumei Biotechnology Company (Changchun, China).

Table 1: Main reagents and sources

Reagent	Source
Plant DNA extraction kit	Kang Wei Reagent Company (Jiangsu, China)
High purity plasmid extraction kit	Wegrath Biology Company (Beijing, China)
Agarose gel extraction kit	Changchun Weijie Company (Changchun, China)
Taq polymerase in PCR, dNTP, Buffering	Fermentas Company (Beijing, China)
Bgl II	Seymour Flight Company (Shanghai, China)
Bste II	Seymour Flight Company (Shanghai, China)
Seamless assembly cloning kit	Zhongmei Taihe Company (Beijing, China)
10× Loding buffer, DL2000 Marker	TaKara Company (Beijing, China)
CRISPR/Cas vector construction kit	Baige Gene Technology Company (Jiangsu, China)
Other biological reagents were analyzed for pure classes	Beijing Dingguo Company (Beijing, China)

2.3 Genomic DNA Extraction

The young leaves of soybean JN28 after three leaf stage were extracted by DNA kit produced by Kang Wei Reagent Company (Jiangsu, China). The quality (purity and concentration) of the extracted genome was measured by nucleic acid protein detector and used as the template for subsequent tests. The leaves were stored at -20°C for later use.

2.4 Cloning of Soybean Gm15G117700 Gene

Nucleotide sequences of oleic acid regulatory genes screened in the laboratory in the early stage were used to design upstream and downstream primers 8A/8As ([Tab. 2](#)). The PCR system was 25 μL : template 2 μL , upstream primers 1 μL , down-stream primers 1 μL , ddH₂O 7.5 μL and PCR master mix 12.5 μL . PCR conditions were: pre-denaturation at 95°C , denaturation at 95°C , annealing at 58°C , extension at 72°C , and extension at 72°C for 5 min, 35 s, 40 s, 30 s, 8 min respectively about 35 cycles. After PCR amplification, 1.0% agarose gel electrophoresis was used to identify the position and purpose stripe bands of DNA. Gel recovery on purpose using cloning technology, seamlessly connected PMD18-T-Gm15G117700. Heat shock method was used to connect the product into DH5 alpha coated board, at 37°C for 15 h. Extraction of plasmid sent to Changchun kumei of Jilin province biotechnology laboratory company for DNA sequencing.

2.5 Sequence Analysis of Target Genes

To conduct sequence analysis of target genes, different softwares like, SignalP analyzing protein signal peptide and PSORT II analysis of protein subcellular localization were used. In addition, ProtScale,

TMHMM Server 2.0, NetPhos, SOPMA and Swiss-Model was used to analyze hydrophobicity and hydrophilicity of protein, the protein transmembrane, the protein phosphorylation site, the secondary structure of the protein and the tertiary structure of the protein, respectively.

Table 2: Primer sequences used in this experiment

Name	Primer sequence (5'→3')	Target fragment/bp
8A	ATGGATATTCAAAATCAAAG	693 bp
8As	CTAAACTGCAATAAGCTGT	693bp
8AOE	actcttgaccatggtagatctGATGGATATTCAAAATCAAAG	
8ASOE	ggggaaattcgagctggtcaccCTAAACTGCAATAAGCTGT	
barA	TCAAATCTCGGTGACGGGC	552bp
barAs	ATGAGCCCAGAACGACGCC	
35SA	TAGAGGACCTAACAGAAC	500bp
35sAS	CCGTGTTCTCTCCAAATG	
NoSA	TTTCTTAAGATTGAATCCTGTTGCC	207bp
NosAS	ACCGCGCGCGATAATTT	
Q8A1	CTGGCCTTTCCATACATCCTC	
Q8As2	TGGCTACAAATGCAAGGATTCCA	
Qb1	ATCTTGACTGAGCGTGGTTATTCC	
Qb2	GCTGGTCCTGGCTGTCTCC	

2.6 Construction and Identification of Overexpression Vectors

Using the *Bgl* II and *Bst* E II pCambia-3301 empty carrier for double enzyme linearization pCambia-3301 empty carrier was used. PMD18-T-Gm15G117700 was used as a template to amplify the target fragment specifically by PCR. After 1.5% agar gel electrophoresis, the linearized target fragment (693 bp) was purified by DNA purification kit and connected by seamless cloning kit to obtain the overexpression vector of soybeanGM15G117700 gene. The overexpression vector was transformed into the receptive state of *E. coli* by heat shock method, which was screened using LB petri dishes with Kanamycin. After screening, single colonies were selected and shaken by LB petri dishes with Kanamycin liquid medium. DNA plasmids were extracted, and double enzyme digestion was performed for identification and in order to confirm whether the target gene was successfully connected to the pcambia-3301 vector (Fig. 1).

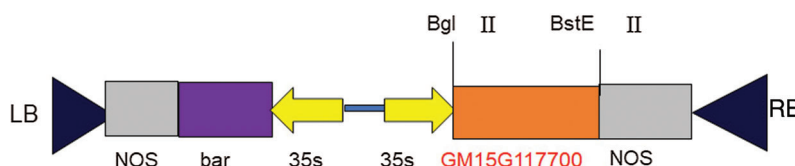


Figure 1: Soybean GM15G117700 gene overexpression vector structure

2.7 CRISPR/Cas9 Gene Editing Vector Target Information

The target gene GM15G117700 for CRISPR/Cas9 gene editing uses single-gene double-target knockout. The two target sequences are target 1: TCTCTGACAGATTTTGAAGGTGG, target 2: TATGGTTGAAGAACAACAAGGGG, where target 1 is on the sense strand and target 2 Targeted editing of the target gene on the antisense strand (Fig. 2).

```

CAGAAGTACACAGCAAGGCACAACTATAGTAACACAATACTTATCATTTACACCAAACCATCACCTCACACTCAC
GAATTTTCAATACTCAATATTTATTCTTTGTGCTTACTTTTGTAGAGAAGAGTTAGCTAGGAAAGGTAGCATGGATAT
TCAAAATCAAAGTGAGGATCATGTCCTTAGGGAATATTATTTAAACAACAAACAGTTTTTGAAAAGAACTTGGCAGT
TTGTTCTCTCAGTGCTGTGTTTTCTGTTTTGTTTGGTACTCTTCTGGCCTTCCATACATCCTCAGTCCTTCAAT
GCCTACTTCTCCACGTGCCTTTTCTACTCCATGATCACTCACACCCTTGAGAGGAAATACATGTTTCTCATTGTAA
TGGAATCCTTGCATTTGTAGCCAAACCTCACTCAGGACGTGCAACTCTTCAGATTCTGCTGACAGCTTTGACCACC
TTCAAAATCTGTGAGAGCAAAAGACAACAGCAGTTTCTGACATGGTTGTTGTCCCTTGGGAAGTTTGAATCCCCA
GAGAATGTTCTCTTATGGTTGAAGAACAACAAGGGAGAACAAATGAAGAAGTTTCAGAAGCTGATGAAGATCAAGA
AGAATCTCACATTACAGAAGAAGAGGATATCAAAGAGGAGGATGAAGAAGGAGAAGCAGAAAGTGAGGCTGAGGTTG
CACAAGATTATGAAGTAGAAGAAACAACAACAACAAATGAAGAGTTAGTAAATACAGAAGAACTGAACAGAAAGTTT
GAAGAGTTCATAAGGAAAATGAAGGAAGAGATCAGAATTGAAGCTCAAAGACAGCTTATTGCAGTTTAGCTAGTGCA
TCTGATATGATAAAAGATGATGACAAGGAATAAAGTGCTTTAGTCTGTTTTCAAATAATAAACCAAGCCTTATGGT
TCTCTTTTGTGTAATTTGCTTCCTTTAATTTGTTTGCATGTGTAGAGCTATCTCATTCCATGATTCCATCCCTCT
AGCTTCTTTCCTTTTTCAGTCTGCAAAATTTATGTCATGTCATTTTATAGTGA

```

Figure 2: Target information of CRISPR/Cas9 gene editing vector

3 Results

3.1 Objective Gene Cloning

The nucleotide sequence of soybean oleic acid related gene Gm15G117700 was obtained by designing with specific Primer 5.0 and the target fragment was amplified by PCR. The fragment was obtained by 210 V, 100 mA and 1% agar gel electrophoresis, and 693 bp fragment was found (Fig. 3). Which was consistent with the expected fragment. It was preliminarily considered that fragment was the target fragment.

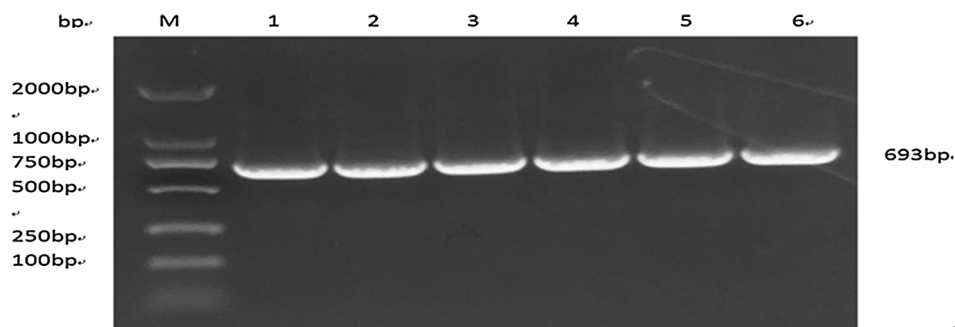


Figure 3: PCR products of GM15G117700 gene in soybean. Note: M: DL2000 DNA Marker 1–6:PCR products

The specific target band was recovered by DNA purification kit, and the method of seamless cloning was used to connect the purified target fragment DNA with pmD-18T vector through the method of seamless cloning by using seamless cloning kit. Carrier of the heat shock method connect seamlessly cloned into DH5 alpha, get into the bacterium fluid, containing Ampicillin with LB medium conditions, extraction of plasmid DNA in bacteria liquid, specific primers biological sequencing. Using DNAMNA sequence alignment analysis software with the purpose gene the result shown homology comparison analysis of homologous rate of 100% (Fig. 4). The gene was successfully cloned by gene sequence alignment analysis.

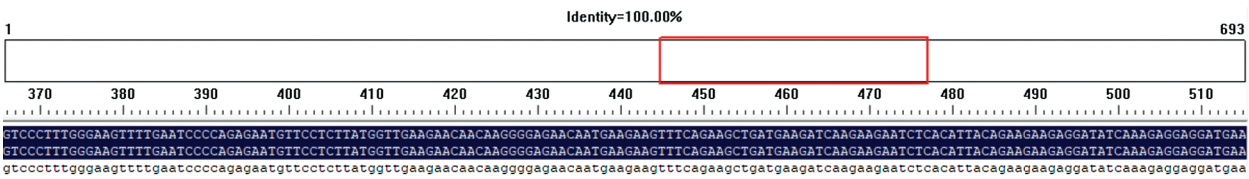


Figure 4: Comparison of Gm15G117700 gene sequencing results

3.2 Nucleotide Sequence Analysis of Target Genes

By Blast analysis and search on NCBI website, the nucleotide sequence of soybean gene fragment was obtained. It was found that the homology of this gene with Gm15G117700 gene was up to 100%, so it can be inferred that Gm15G117700 gene was successfully cloned (Fig. 5).

	Description	Score	Score	Cover	value	Ident	
✓	PREDICTED: Glycine soja uncharacterized LOC114367706 (LOC114367706).mRNA	1280	1280	100%	0.0	100.00%	XM_028324903.1
✓	PREDICTED: Glycine max uncharacterized LOC102868654 (LOC102868654).mRNA	1280	1280	100%	0.0	100.00%	XM_006586713.2
✓	PREDICTED: Vigna unguiculata uncharacterized LOC114188587 (LOC114188587).mRNA	606	606	98%	5e-169	83.17%	XM_028077158.1
✓	Vigna unguiculata cultivar Xiabao 2 chromosome Vu06	601	601	98%	2e-167	83.03%	CP039345.1

Figure 5: The predicted result of target gene

3.3 Signal Peptide Prediction and Subcellular Localization of Soybean Gm15G117700 Protein

The protein signal peptide can be predicted by online analysis of Signalp (Fig. 6). The protein encoded by the gene Gm15G117700 is a non-secretory protein, and the localization of the protein Gm15G117700 subcell is studied by using the online website PSORT II (Fig. 7), which contains 65.2% in the cytoplasm, indicating that the protein in the gene Gm15G117700 may play a major role in the cytoplasm.

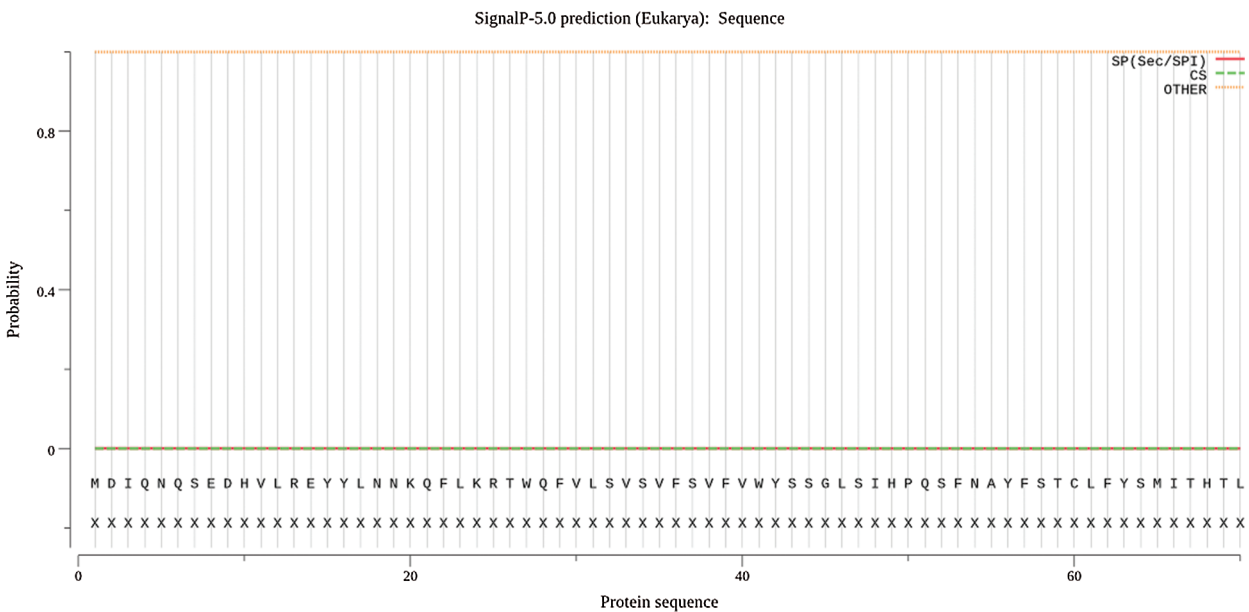


Figure 6: Gm15G117700 protein signal peptide prediction

```

k = 9/23

65.2 %: cytoplasmic
13.0 %: mitochondrial
13.0 %: nuclear
4.3 %: cytoskeletal
4.3 %: Golgi

>> prediction for QUERY is cyt (k=23)

```

Figure 7: Subcellular localization of protein of soybean Gm15G117700 gene

3.4 Soybean Protein Hydrophobicity, Hydrophilic Research and Transmembrane Domain Structure Prediction

The hydrophobicity and hydrophilicity of proteins were studied and analyzed with ExPASy-ProtScale software, and Hphob/Kyte&Doolittle were used as the prediction criteria (Fig. 8). There were many high peaks, among which the amino acids at positions 31–38, 78–87, 119–126 had obvious hydrophobic regions, and the value of amino acid at the 34th position was the largest, 2.333, where the hydrophobicity was the strongest. The value of amino acid at 143,144 was the lowest, –3.156, where hydrophilicity is the highest. The whole polypeptide chain was composed of 28% hydrophobic amino acids and 72% hydrophilic amino acids. The overall expression was hydrophilic. It was found that the protein of Gm15G117700 gene was a soluble protein. According to the online study of TMHMM Server 2.0, the protein transmembrane region was predicted (Fig. 9). The protein of soybean Gm15G11770 has a transmembrane structure, which is a transmembrane protein and migrates in the cell. TMpred Server software was used for the prediction analysis of transmembrane region (Fig. 9). According to the ordinate value in the figure, when the value exceeds 500, it can be considered as a structural domain containing transmembrane. It is shown in Fig. 10 that this gene protein has a transmembrane structure, which is a transmembrane protein and migrates in the cell.

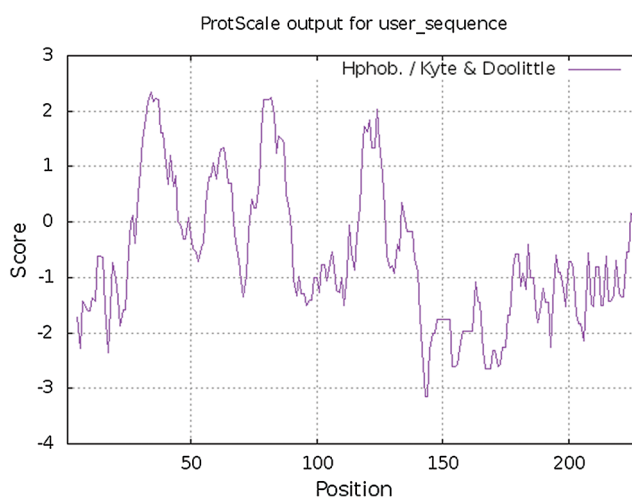


Figure 8: Hydrophobic and hydrophilic analysis of Gm15G117700 protein

3.5 Study on Protein Phosphorylation Sites of Soybean Gm15G117700 Gene

Using NetPhos software, protein sequence phosphorylation sites (Fig. 11) were analyzed. Gm15G117700 protein 29 potential phosphorylation sites, including two tyrosine (Tyr), 12 threonine (Thr), 15 serine phosphorylation site (Ser) were identified.

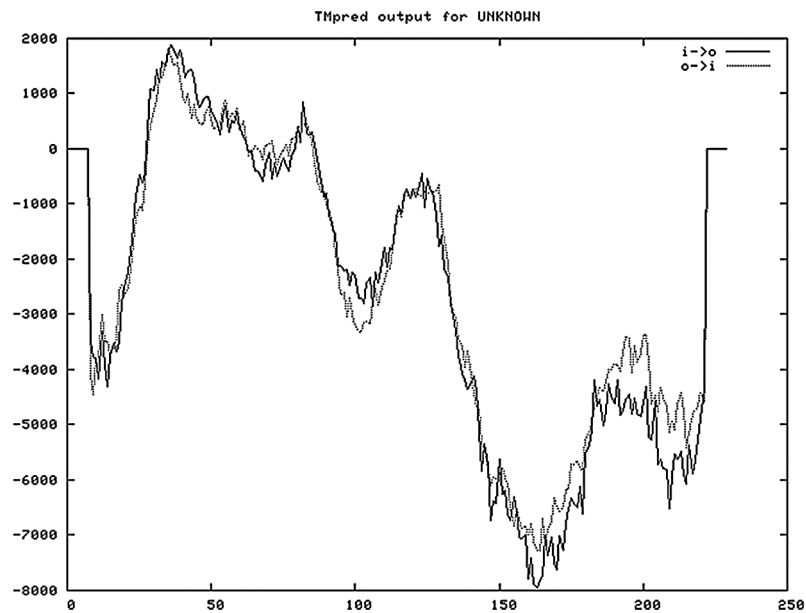


Figure 9: Prediction of protein transmembrane structure domain (TMpred Server)

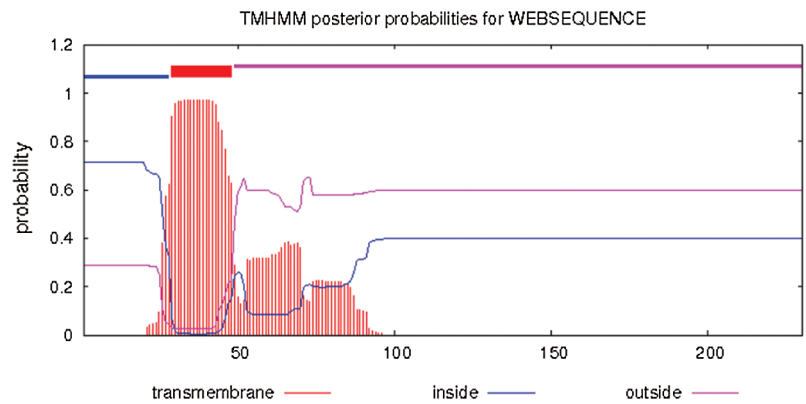


Figure 10: Transmembrane region prediction of Gm15G117700 protein (TMHMM Server)

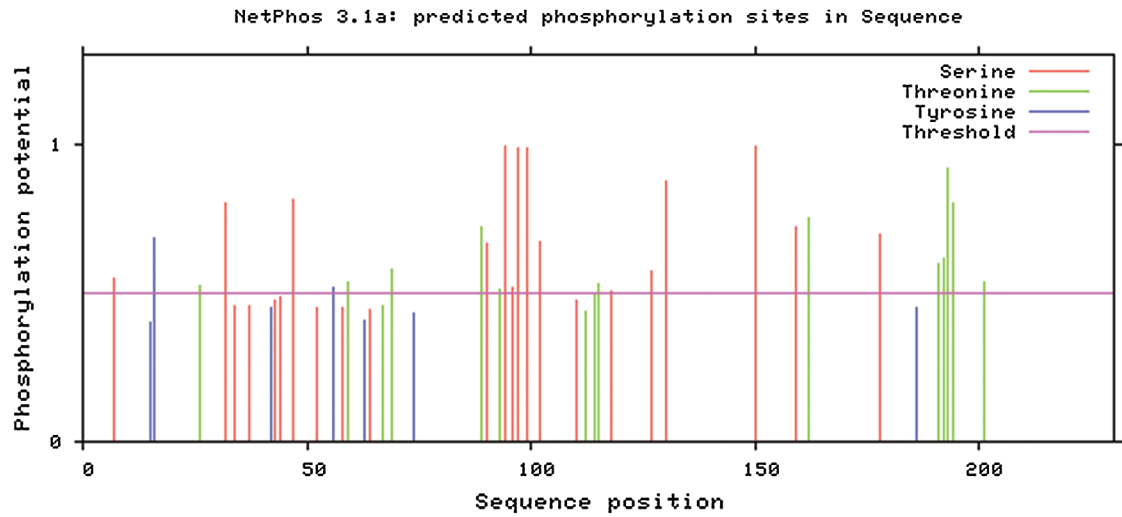


Figure 11: Gm15G117700 gene protein phosphorylation site research (NetPhos)

3.6 Protein Structure Analysis of Soybean Gm15G117700 Gene

Using SOPMA online analysis, Gm15G117700 gene encoding amino acid sequences and protein secondary structure prediction were analyzed (Fig. 12). Gm15G117700 gene encoding protein composition consists of four parts, the alpha helix (59.57%), extend chain (14.78%), beta angle (3.04%), and no rules curly (22.61%). Alpha helix of ratio was the largest proportion of the second random curl whereas the least percentage was beta around the corner. Thus alpha helix and random curl play an important guiding role in gene function.

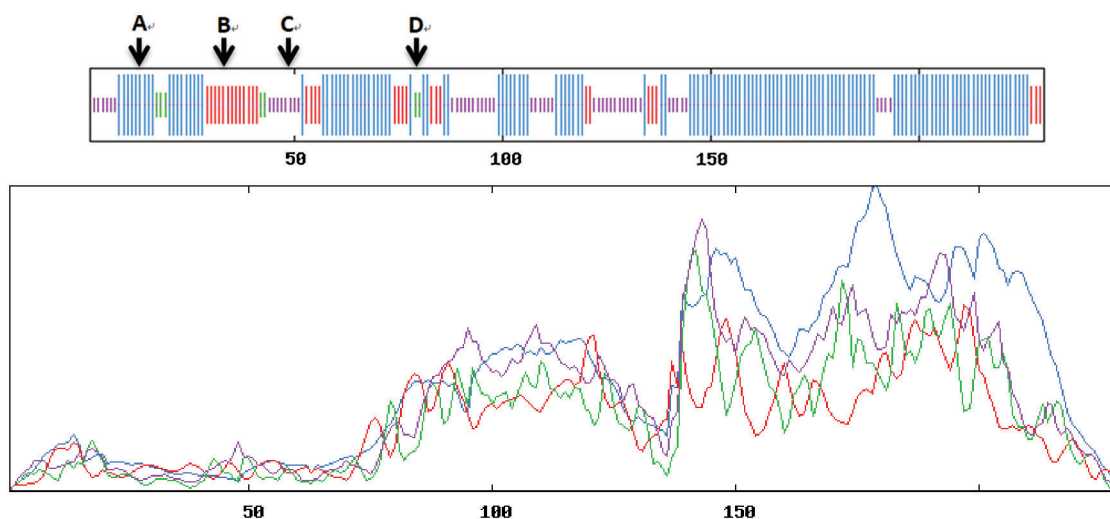


Figure 12: Secondary structure for protein encoded by soybean Gm15G117700 (A) α -helix (B) extended chain (C) random coil (D) β -turn

SWISS-MODEL software was used for Gm15G117700 gene to sequence amino acids and protein tertiary structure (Fig. 13) which contains alpha helix, extending chain and beta curly angle, but no rules. Consistent with the protein secondary structure prediction, and Gm15G117700 gene encoding protein and the protein three-dimensional structure of the database in Arabidopsis homology as high as 55%.

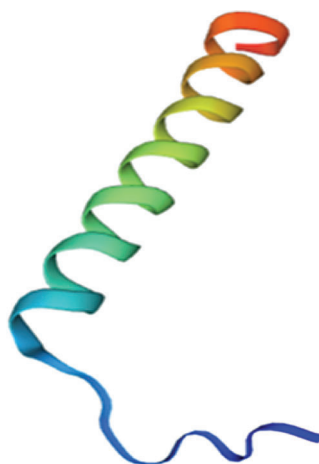


Figure 13: Tertiary structure for protein encoded by soybean Gm15G117700

3.7 Construction and Verification of Overexpression Vector

Through CE design software for pCambia3301 empty carrier gene Gm15G117700 sequence designed and purpose, find out the appropriate enzyme loci *Bgl* II and *Bst*E II, pCambia3301 double enzyme linearization. The Gm15G117700 gene overexpression vector was obtained by connecting the two linearized fragments with the seamless cloning kit, and the double enzyme digestion was used to verify whether the carrier was successfully connected. The target fragment with a fragment size of 693 bp and the vector fragment with an action size of about 9500 bp were obtained (Fig. 14), which were consistent with the prediction, indicating that the soybean Gm15G117700 overexpression vector was successfully constructed.

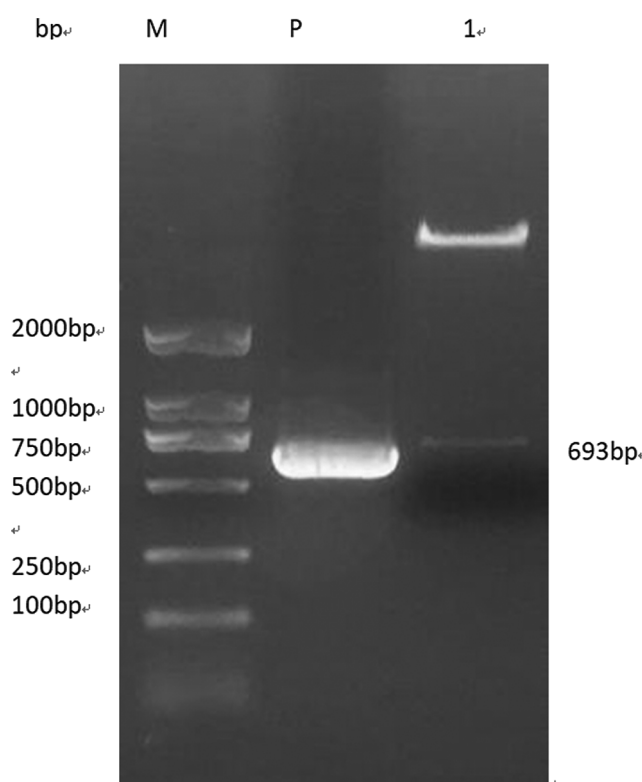


Figure 14: Double restriction digestion verification of overexpression vector. Note: M: DL2000 DNA Marker P: Plasmid 1: Enzyme digestion product

3.8 CRISPR/cas9 Gene Editing Carrier to Build, Test and Verify

Using BGK053 as the basic vector skeleton, Oligo sequence was designed and Oligo dimer was prepared using online software for the target site. The Gm15G117700 gene editing vector was constructed using kit method. Due to the presence of anti-glyphosate screening sites on the vector, PCR specific amplification was performed as shown in Fig. 15, and the soybean Gm15G117700 expression inhibiting vector was successfully constructed.

3.9 CRISPR/Cas9 Gene Editing Mutation Sites

CRISPR/Cas9 gene-edited soybean mRNA was extracted and reverse transcribed cDNA. Specifically the band of interest amplified, then, gel-purified, and sent it to Kumei Biotech Company for sequencing. DNA man software was used to compare sequences. On target 1, one mutation site (T changed to C) and 2 mutation sites on target 2 (G changed to A) were found (Fig. 16). Suggesting, new mutation were found.

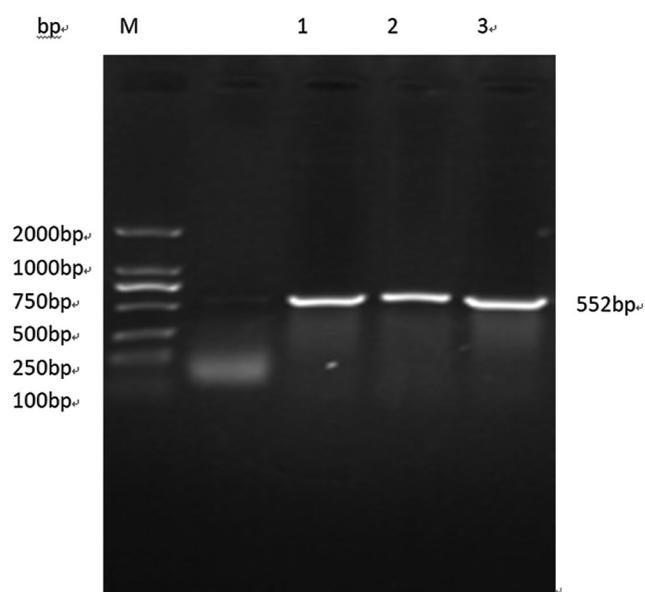


Figure 15: Verification of interference expression vector. Note: M: DL2000 DNA Marker 1–3: bar gene

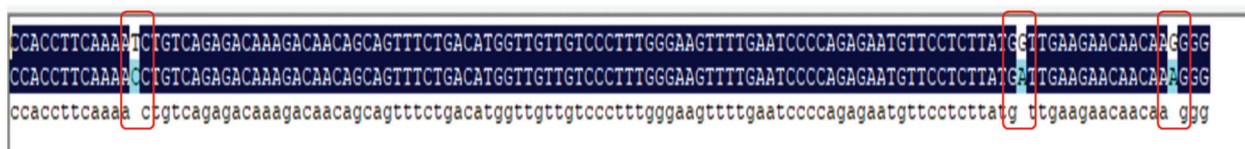


Figure 16: CRISPR/Cas9 gene editing mutation site

3.10 PCR Identification of Transgenic *T₂* Plants

Plant DNA extraction kit for *T₂* positive soybean was used. DNA genome was extracted from untransformed plants as CK control group and CK ddH₂O as negative control. The *Escherichia coli* containing the target gene was stored in the –80°C refrigerator. The constant temperature was 37°C, the bacteria were propagated and the plasmid was extracted as the positive control. Marker genes of 35S promoter, NOS terminator, and Bar were screened for PCR identification (Fig. 17), and 6 overexpressed positive plants and 5 suppressed expressed positive plants were obtained.

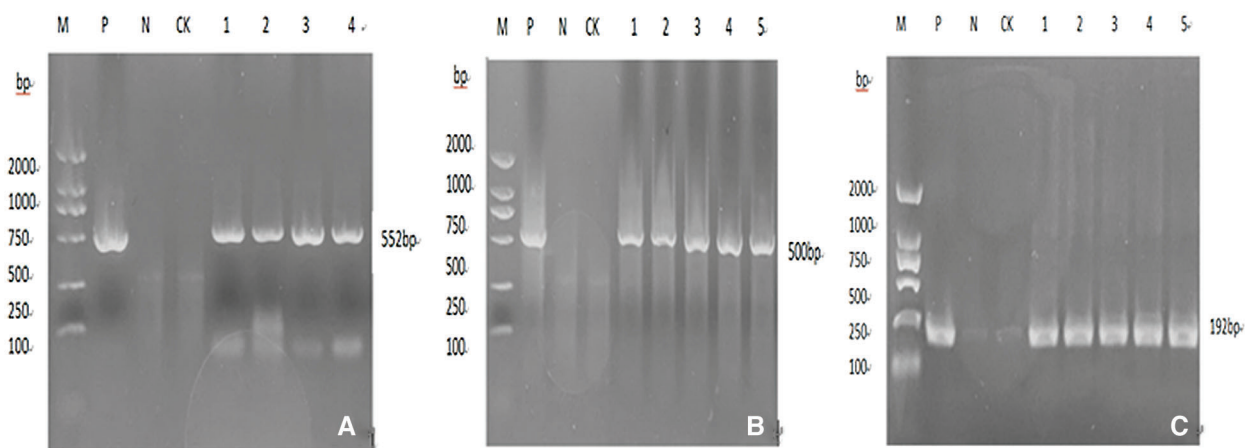


Figure 17: A, B, and C are the over-expression bar, 35 s, and Nos gene. Note: M: DL2000 DNA Marker; P: Plasmid (positive control); N: H2O; CK: untransformed plants 1–5 transformed plants

3.11 Southern Blot Hybridization Identification of T2 Transgenic Plants

Plasmid containing the purpose gene as the positive control group was constructed. CK in the control group, ddH₂O negative control group, *Bgl* II restriction enzymes, were used using filter Bar marker genes as molecular probes to Southern Blot hybridization detection. The result indicates that there was no hybridization signal in the control group. 4 of the 6 strains showed significant hybridization signal in the overexpression and all of which were single copies (Fig. 18(A)), and 4 of the 5 strains with inhibited expression vectors showed hybridization signal, all of which were single copies (Fig. 18(B)), indicating different integration sites of the introduced genes in the genome.

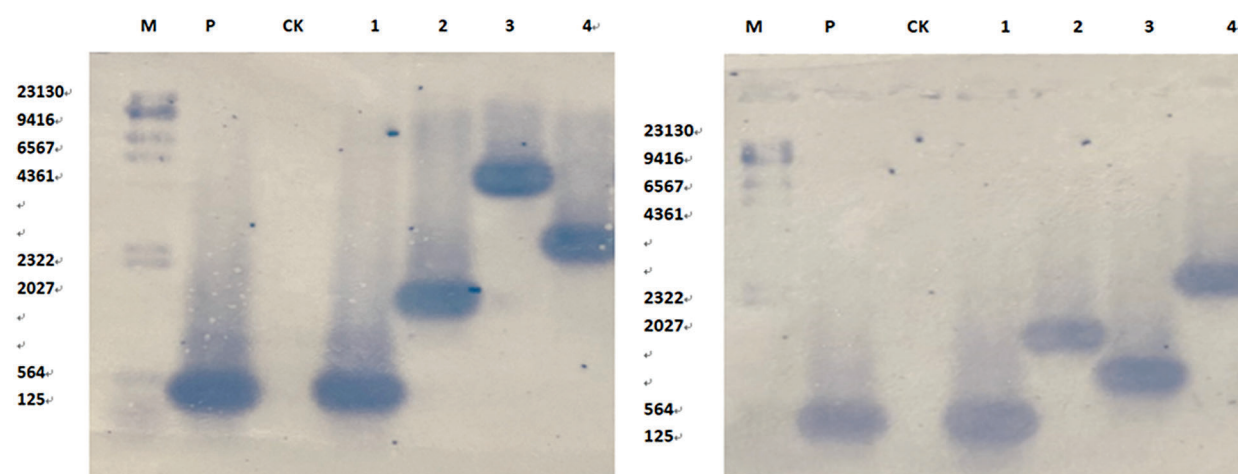


Figure 18: Southern Blot hybridization of T2 transgenic plants. Note: M: Southern dedicated Marker; P positive control; CK untransformed plants; A1-4 overexpression plants; B1-4 suppression expression plants

3.12 T₂-generation Plants QT-PCR Identification

T₂ transgenic soybean seed mRNA was extracted. Oligo (dT), reverse transcription cDNA, cDNA template for PCR amplification were used. The result showed that the expression level of Gm15G117700 overexpressed plants were about 2.4 times that of CK (untransformed plants), and the expression level of gene edited plants was about 0.5 times that of CK (Fig. 19). The results were the same as expected, and the transformation was successful.

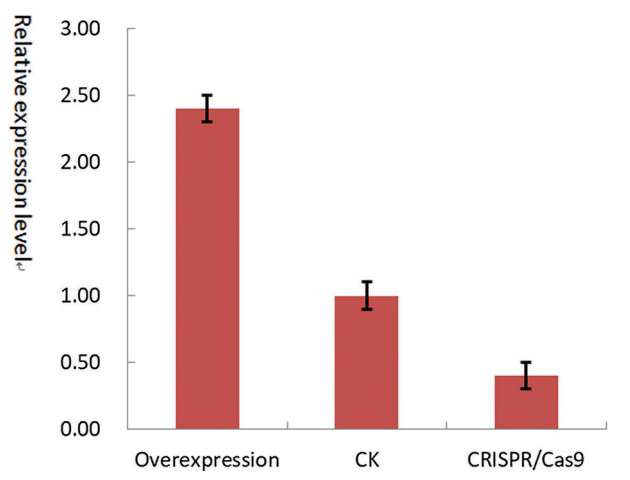


Figure 19: QT-PCR identification of T2 generation plants

3.13 QT-PCR Identification of Different Tissues of T2 Generation Plants

DNA was extracted from the roots, stems, leaves and seeds of T2 transgenic plants. DNA was used as a template to specifically amplify gene and β -actin was used as an internal reference gene. The over-expression in the roots, stems, leaves and seeds of the plant were found. The relative expression levels in the roots, stems, leaves and seeds were found to be 1.60, 1.60, 1.32, 2.40, respectively. The relative expression levels of CRISPR/Cas9 in plant roots, stems, leaves and seeds were 0.40, 0.30, 0.20, and 0.50, respectively (Fig. 20).

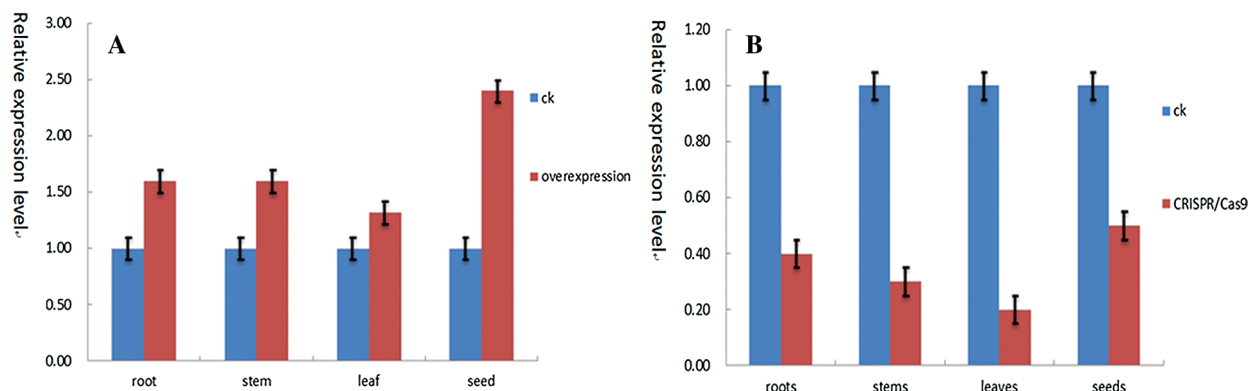


Figure 20: (A) Overexpression in roots, stems, leaves and seeds. (B) CRISPR/Cas9 expression content in roots, stems, leaves and seeds

3.14 Identification of High and Low Oleic Acid Soybean and New Gene Soybean QT-PCR

The whole genome sequencing (GWAS) of 260 spring soybean germplasm resources was performed to select 10 soybeans with high oleic acid content and 10 soybeans with low oleic acid content. Fluorescence quantitative analysis (QT-PCR) was performed on 10 soybeans with high and low oleic acid content and overexpressed new genes. High oleic acid in soybean relative expression of the gene is low. The relative expression quantity was below 1.50 (Fig. 21(A)). Low oleic acid soybean expression of the genes with high content in the highest relative expression is 1.80 (Fig. 21(B)), which can presumably the genes may promote oleic acid in the acid metabolism to linoleic acid to synthesize a kind of regulatory gene. The gene expression in high oleic acid soybean volume is relatively low. Low oleic acid in soybean has shown relatively high amount of gene expression.

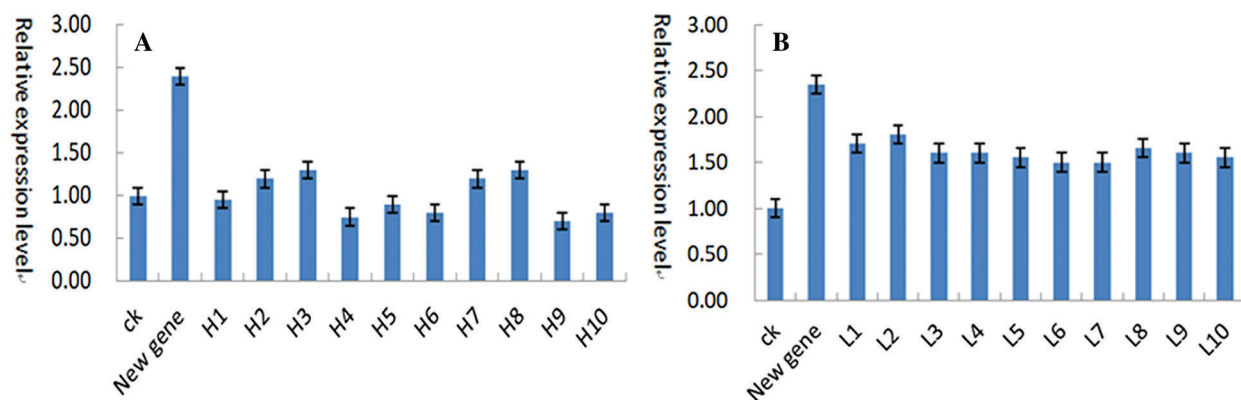


Figure 21: (A) High oleic acid germplasm and new gene QT-PCR (B) Low oleic acid germplasm and new gene QT-PCR

3.15 *Gm15G117700* Gene Identification of Oleic Acid Content

The quality and characters of soybean T₂ seeds were identified by near-infrared spectrum analyzer. Palmitic acid, stearic acid, oleic acid, linoleic acid and linolenic acid were measured in 6 overexpressed plants and 5 CRISPR-Cas9 gene edited and expressed plants. Each seed was measured three times and the average value was taken to ensure the accuracy of the data (Tab. 3). There is no change in palmitic acid content from the data. There was no significant change between the overexpression vector and CK in the content of stearic acid, while the content of gene editing vector was increased by 2.51%. The overexpression vector on oleic acid content was 3.94% lower than on average CK. The gene editing vector was 3.49% higher on average than CK. The content of linoleic acid in the overexpression vector increased by 1.7% and the gene editing vector decreased by 5.61%. The change of linolenic acid content is not obvious. Thus it can be seen that *Gm15G117700* gene may play a positive regulation role in the synthesis of linoleic acid. In the FAD2 family, this gene acts as an enhancer, and the content of oleic acid is increased only when the expression of *Gm15G117700* gene is inhibited.

Table 3: Determination of soybean fatty acid content

Type	Palmitic acid (%) C16:0	Stearic acid (%) C18:0	Oleic acid (%) C18:1	Linoleic acid (%) C18:2	Linolenic acid (%) C18:3
O-8	10.88	1.63	14.04	60.57	11.06
O-23	10.53	1.02	12.59	64.52	10.17
O-5	10.37	2.61	16.28	57.30	10.59
O-10	10.31	2.29	15.93	57.70	11.57
O-21	10.96	2.03	15.65	60.16	10.19
O-6	11.68	1.23	15.34	61.75	9.84
C-2	12.90	4.29	24.01	51.97	13.26
C-7	11.04	4.35	21.61	53.75	8.41
C-14	11.58	5.60	23.00	50.74	14.40
C-16	12.16	5.38	23.24	51.67	13.91
C-18	13.16	4.59	24.03	51.40	14.62
CK	10.83	1.79	18.91	58.63	9.65
G Ave	10.79	1.80	14.97	60.33	10.57
C Ave	11.94	4.30	22.40	53.02	12.38

4 Discussion

Oleic acid is listed as a safe fatty acid in nutrition. Studies have shown that ingestion of large amounts of polyunsaturated fatty acids (linoleic acid, linolenic acid) and saturated fatty acids (stearic acid, palmitic acid, etc.) may have negative effects on the body [17]. Excessive intake of saturated fatty acids is prone to hypertension, coronary heart disease and cardiovascular diseases [18]. Both linoleic and linolenic acids are essential fatty acids. Can produce both beneficial and harmful cholesterol levels, which can accelerate atherosclerosis [19]. As for the treatment of polyunsaturated fatty acids, hydrogen reaction can be adopted in industry to make them more stable, which is not easy to be oxidized by air and has excellent stability. However, trans-fatty acids are produced, which will increase the content of cholesterol harmful to health [20]. Trans fatty acids may cause the occurrence of cardiovascular diseases (coronary heart disease, etc.) [21]. Oleic acid can reduce low-density lipoprotein in human blood, while high-density lipoprotein will not be affected. It can slow atherosclerosis and prevent cardiovascular diseases

[22]. So oleic acid is very good for health, the body can't absorb trans fatty acids, it can absorb the cis-fatty acid structure. Natural oleic acid is a fatty acid that contains only cis-structured fatty acids and is important in metabolism and in improving the elasticity of blood vessels. The human body can conduct its own biosynthetic oleic acid, but sometimes it cannot meet the demand, so it needs to absorb from the outside food. The high quality oil in food oil contains food oil which is easy to process, is not easy to be oxidized under high temperature, and avoids industrial hydrogenation, such as monounsaturated fatty acid and high proportion of oleic acid.

There were no reports about soybean oleic acid Gm15G117700 gene cloning and functional analysis. This experiment was successfully cloned and early genetic correlation analysis of oleic acid soybean gene Gm15G117700 identified by NCBI database. To identify gene Gm15G117700, the purpose gene was cloned. The gene of signal peptide and protein subcellular localization, hydrophobicity, hydrophilicity, protein phosphorylation sites, the secondary and tertiary structure prediction analysis were performed. Presumably the genes may be associated with soybean oleic acid. For the construction of gene vectors (the expression vector and gene editing carrier construction), agrobacterium mediated transformation method were applied and the transgenic plants were generated. The conventional PCR, Southern blot hybridization and RT-PCR were used for T₂ plant identification. Ten varieties of high and low oleic acid soybean were selected from spring soybean. RT-PCR was performed to identify the new gene expression quantity of Gm15G117700, and then the function of this gene was preliminarily determined. Conventional PCR preliminarily verified and the gene was transferred into the soybean genome. Southern blot hybridization result showed that the gene was a single copy in both overexpressed plants and disturbed the expression plants. The overexpressed vector plants in RT-PCR seeds were about 2.4 times as much as those in CK control group, and the gene editing vector plants were about 0.5 times CK. Ten varieties of soybeans with high and low oleic acid were selected and RT-PCR was conducted. The results showed that Gm15G117700 gene expression was very low in soybeans with high oleic acid, while it was high in soybeans with low oleic acid. This result proving that this gene had a synergistic effect with FAD-2 gene. Then by using near infrared spectrum analyzer the content of fatty acid in soybean change was determined. Overexpression of oleic acid content in seeds were 3.94% while gene editing oleic acid content in seeds increased by 3.49%. This result suggests that Gm15G117700 gene may be a regulatory gene, which can promote the conversion of linoleic acid content in the synthesis of oleic acid in soybean. This finding lay a certain foundation for further analysis of the function of Gm15G117700 gene.

Authors Contribution: Shuo Qu and Yaolei Jiao, conduct the experiment, analyzed data and wrote first draft. Abraham Lamboro analyzed data and edit the paper. Piwu Wang was the project leader of the experiment and guided the writing of this article. All the authors read and approved the final manuscript.

Funding Statement: This study was supported by Chinese National Natural Science Foundation (31571689) and Major National Science and Technology Project for New Varieties of Genetically Modified Organisms (2016ZX08004-004-003).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Cui, D. Q. (2017). The re-discussion of the Chinese Soybeans and Rong Shu: The origin of Chinese soybeans. *Chinese Agricultural History*, 36(4), 15–33.
2. Won-Seok, K., Hari, B. K. (2004). Expression of an 11 kDa methionine-rich delta-zein in transgenic soybean results in the formation of two types of novel protein bodies in transitional cells situated between the vascular tissue and storage parenchyma cells. *Plant Biotechnology Journal*, 2(3), 199–210.

3. Santiago, T., Jose, A., Aznar-Moreno, T. P., Durrett, P. V. V., José, L. R. et al. (2020). Dynamics of oil and fatty acid accumulation during seed development in historical soybean varieties. *Field Crops Research*, 248(12), 107719. DOI 10.1016/j.fcr.2020.107719.
4. Chen, L. H. (2020). Analysis of fatty acid composition of genetically modified soybean oil. *Journal of Jinggangshan University (Natural Science Edition)*, 41(5), 26–30.
5. Zhong, X. F., Qian, X. Y., Niu, L., Guo, D. Q., Yang, X. D. (2019). The effect of genetically modified technology on increasing soybean oil and oleic acid content. *Soybean Science and Technology*, 2019(6), 27–29.
6. Cao, Y. Q., Xie, F. J., Dong, L. J., Wang, Y. Z., Song, S. H. et al. (2015). Research progress on oleic acid content of soybean seeds. *Soybean Science*, 34(2), 329–334.
7. Chen, X., Liang, K. H., Wang, J., Zhu, H. (2020). Research progress in the prevention and treatment of cardiovascular diseases by polyunsaturated fatty acids in diet. *China Oils and Fats*, 45(10), 87–94.
8. Nerea, B. T., Sonia, B. M., Effie, V., Tauseef, K., Cyril, W. C. K. et al. (2020). Mediterranean diet, cardiovascular disease and mortality in diabetes: A systematic review and meta-analysis of prospective cohort studies and randomized clinical trials. *Critical Reviews in Food Science and Nutrition*, 60(7), 1207–1227. DOI 10.1080/10408398.2019.1565281.
9. Critselis, E., Kontogianni, M. D., Georgousopoulou, E., Chrysoshoou, C., Tousoulis, D. et al. (2020). Comparison of the Mediterranean diet and the dietary approach stop hypertension in reducing the risk of 10-year fatal and non-fatal CVD events in healthy adults: The ATTICA Study (2002–2012). *Public Health Nutrition*, 23, 11–12.
10. Zhang, Y. H., Zhang, Y. J., Su, P., Zhang, T., Zhou, R. M. et al. (2019). Metabolic pathways of oils and breeding strategies to increase the oil content of oil crops. *Molecular Plant Breeding*, 17(9), 3084–3089.
11. Topfer, R., Martini, N., Schell, J. (1995). Modification of plant lipid synthesis. *Science*, 268(5211), 681–686. DOI 10.1126/science.268.5211.681.
12. Zhang, J. T., Li, J. H., Hernan, G. R. (2014). A stearyl-acyl carrier protein desaturase, NbSACPD-C, is critical for ovule development in *Nicotiana benthamiana*. *Plant Journal*, 80(3), 489–502. DOI 10.1111/tpj.12649.
13. Jiang, C. J., Shimono, M., Maeda, S. (2009). Suppression of the rice fatty-acid desaturase gene OsSSI2 enhances resistance to blast and leaf blight diseases in rice. *Molecular Plant Microbe Interactions*, 22(7), 820–829. DOI 10.1094/MPMI-22-7-0820.
14. Aardra, K. H., John, S. K., Edward, W. (2007). The Arabidopsis stearyl-acyl carrier protein-desaturase family and the contribution of leaf isoforms to oleic acid synthesis. *Plant Molecular Biology*, 63(2), 257–271.
15. Hou, Z. H., Wu, Y., Cheng, Q., Dong, L. D., Lu, S. J. et al. (2019). Using CRISPR/Cas9 technology to create soybean high oleic acid mutant lines. *Acta Agronomica Sinica*, 45(6), 839–847.
16. Gupta, P. K., Kulwal, P. L., Jaiswal, V. (2019). Association mapping in plants in the post-GWAS genomics era. *Advances in Genetics*, 104, 75–154.
17. Raatz, S. K., Conrad, Z., Jahns, L., Belury, M. A., Picklo, M. J. (2018). Modeled replacement of traditional soybean and canola oil with high-oleic varieties increases monounsaturated fatty acid and reduces both saturated fatty acid and polyunsaturated fatty acid intake in the US adult population. *American Journal of Clinical Nutrition*, 108(3), 594–602. DOI 10.1093/ajcn/nqy127.
18. Li, J., Qi, S. (2019). Consumption of saturated fatty acids and coronary heart disease risk. *International Journal of Cardiology*, 45(2), 279.
19. Michael, J. A. W., Wayne, H. F. S., Maree, P. M. C. (1999). Impaired endothelial function following a meal rich in used cooking fat. *Journal of American College of Cardiology*, 33(4), 1050–1055. DOI 10.1016/S0735-1097(98)00681-0.
20. Xin, Z. W., Gao, L., Zhang, L. Y., Liu, L., Wei, G. T. et al. (2020). Progress in the selective hydrogenation of polyunsaturated fatty acids and their esters. *Fine Chemicals*, 37(6), 1136–1144+1198.
21. Anupam, C., Magnus, N. L., Ivar, A. E., Helge, R., Thea, V. et al. (2020). Plasma trans fatty acid levels, cardiovascular risk factors and lifestyle: Results from the Akershus Cardiac Examination 1950 Study. *Nutrients*, 12(5), 1419. DOI 10.3390/nu12051419.
22. Michihiro, S., Takayoshi, O., Kei, A., Yoshitaka, M. K., Daisuke, S. Y. et al. (2021). A combination of blood pressure and total cholesterol increases the lifetime risk of coronary heart disease mortality: EPOCH-JAPAN. *Journal of Atherosclerosis and Thrombosis*, 28(1), 6–24. DOI 10.5551/jat.52613.