Genetics and Molecular Research

Sequence variations in the *FAD2* gene in seeded pumpkins

Y. Ge[1,2*], Y. Chang[1*], W.L. Xu[1], C.S. Cui[1] and S.P. Qu[1]

[1]Key Laboratory of Biology and Genetic Improvement of Horticultural Crops (Northeast Region), Department of Horticulture, Northeast Agricultural University, Harbin, China
[2]Haikou Experimental Station, Chinese Academy of Tropical Agricultural Sciences, Haikou, China

*These authors contributed equally to this study.
Corresponding author: S.P. Qu
E-mail: spqu@neau.edu.cn

Genet. Mol. Res. 14 (4): 17482-17488 (2015)
Received August 15, 2015
Accepted October 29, 2015
Published December 21, 2015
DOI http://dx.doi.org/10.4238/2015.December.21.19

ABSTRACT. Seeded pumpkins are important economic crops; the seeds contain various unsaturated fatty acids, such as oleic acid and linoleic acid, which are crucial for human and animal nutrition. The fatty acid desaturase-2 (*FAD2*) gene encodes delta-12 desaturase, which converts oleic acid to linoleic acid. However, little is known about sequence variations in *FAD2* in seeded pumpkins. Twenty-seven *FAD2* clones from 27 accessions of *Cucurbita moschata*, *Cucurbita maxima*, *Cucurbita pepo*, and *Cucurbita ficifolia* were obtained (totally 1152 bp; a single gene without introns). More than 90% nucleotide identities were detected among the 27 *FAD2* clones. Nucleotide substitution, rather than nucleotide insertion and deletion, led to sequence polymorphism in the 27 *FAD2* clones. Furthermore, the 27 *FAD2* selected clones all encoded the FAD2 enzyme (delta-12 desaturase) with amino acid sequence identities from 91.7 to 100% for 384 amino acids. The same main-function domain between 47 and 329 amino acids was identified. The four species clustered separately based on differences in the sequences that were identified using the unweighted pair group

 ©FUNPEC-RP www.funpecrp.com.br

method with arithmetic mean. Geographic origin and species were found to be closely related to sequence variation in *FAD2*.

Key words: Fatty acid desaturases; Seeded pumpkin; Sequence variations; Nucleotide diversity

INTRODUCTION

The composition of fatty acids in oil determines its use as a food source, or its industrial applications (Harwood, 1997). A large number of fatty acids exist in nature, but only five fatty acids account for about 90% of the commercial vegetable oils produced: palmitic acid (16:0), stearic acid (18:0), oleic acid (18:1), linoleic acid (18:2), and linolenic acid (18:3) (Browse, 1991). In pumpkin seeds, approximately 60% of the oil content is composed of these five fatty acids (Qu et al., 2011). Among the five fatty acids, oleic acid and linoleic acid could effectively lower the total serum cholesterol level in humans (Qu et al., 2012). The relative proportions of oleic acid and linoleic acid determine the relevant physical and nutritional properties of edible oils (Cao et al., 2013). Oleic acid has higher oxidative stability than linoleic acid because it contains one less double bond, resulting in the extension of its shelf life; this property of oleic acid reduces the need for hydrogenation, a process that can generate undesirable trans-fatty acids that have adverse effects on human health (Mozaffarian et al., 2006). Therefore, increasing oleic acid content at the expense of linoleic acid content is an important research objective for the improvement of oilseed crops, including pumpkins, to produce highly stable cooking oils. In our previous study, we reported a variation of up to 35% in oleic acid content in pumpkin seeds (Qu et al., 2011).

Fatty acid biosynthesis is a complex process regulated at several points along the pathway. In this pathway, delta-12 desaturase, encoded by the fatty acid desaturase-2 (*FAD2*) gene, primarily converts oleic acid (18:1) to linoleic acid (18:2) in storage lipids in seeds by introducing a double bond at the 12th carbon in the fatty acid hydrocarbon chain (Mikkilineni and Rocheford, 2003). Since the first plant *FAD2* gene was cloned in *Arabidopsis thaliana* (Okuley et al., 1994), several *FAD2* genes have been identified, isolated, and characterized from other plant species, such as maize (Mikkilineni and Rocheford, 2003; Beló et al., 2008), safflower (Guan et al., 2012; Cao et al., 2013), *Camelina sativa* (Hutcheon et al., 2010), soybean (Pham et al., 2010; Chi et al., 2011), *Brassica napus* (Wang et al., 2010; Yang et al., 2012), peanut (Jung et al., 2000; Lopez et al., 2000), flax (Krasowska et al., 2007), sunflower (Rolletschek et al., 2007), and olive (Georgios et al., 2005). Only a single copy of *FAD2* was identified in *Arabidopsis* and maize (Okuley et al., 1994; Beló et al., 2008), while multiple copies of *FAD2* were found in soybean (Li et al., 2007; Schlueter et al., 2007), maize (Mikkilineni and Rocheford, 2003), cotton (Zhang et al., 2009), and safflower (Cao et al., 2013). Regulation of *FAD2* has been shown to be a successful strategy for modifications of the relative proportions of oleic acid and linoleic acid in plant lipids (Chen et al., 2010, 2011).

The aim of this study was to identify *FAD2* sequence polymorphisms in different pumpkin species and to determine the relationships between the sequence polymorphisms and the geographic origin of the seed cultivars and species.

MATERIAL AND METHODS

Twenty-seven accessions of the seeded pumpkins - five *C. moschata*; twelve *C. maxima*, eight *C. pepo*, and two *C. ficifolia* - from China and Russia were obtained from the Department of

Horticulture, Northeast Agricultural University, Harbin, China (Table 1). Young leaves, randomly chosen from five individuals of each accession, were collected from an open field and stored at 80°C after shock freezing with liquid nitrogen.

Table 1. Geographical origins and species of the 27 seeded pumpkin accessions used in this study.

Accession	Name	Area of origin	Species
1	L128	China	*C. moschata*
2	L127	China	*C. moschata*
3	063116	China	*C. moschata*
4	T4	Russia	*C. moschata*
5	T5	Russia	*C. moschata*
6	063114	Russia	*C. maxima*
7	06808	China	*C. maxima*
8	063120	Russia	*C. maxima*
9	086-1	China	*C. maxima*
10	H46-1261-21	China	*C. maxima*
11	0401-11-11	China	*C. maxima*
12	063113	Russia	*C. maxima*
13	L129	China	*C. maxima*
14	XieHe	China	*C. maxima*
15	HeiFeng	China	*C. maxima*
16	YinHuiYiHao	China	*C. maxima*
17	NO177	China	*C. maxima*
18	06311	Russia	*C. pepo*
19	2112	Russia	*C. pepo*
20	T6	Russia	*C. pepo*
21	JingHuiErHao	China	*C. pepo*
22	0516-2	China	*C. pepo*
23	0512-1	China	*C. pepo*
24	0510-1	China	*C. pepo*
25	0508-1	China	*C. pepo*
26	No1	China	*C. ficifolia*
27	No2	China	*C. ficifolia*

Genomic DNA was extracted using the cetyltrimethyl ammonium bromide (CTAB) protocol, as described by Doyle and Doyle (1990), with minor modifications. Polymerase chain reaction (PCR) for the sequences of the *FAD2* clones was performed using the forward primer, 5'-ctcgagatgacagtaaaaaag-3', and reverse primer, 5'-tctagactatatatactctgggaac-3', which corresponded to the *FAD2* core-conserved sequence obtained from GenBank (accession No. AY525163). The simple sequence repeat (SSR) amplification reaction mixtures (total volume = 50 µL) contained 20 ng DNA, 5 µL 10X PCR buffer (Mg^{2+}), 2 µL primers, 4 µL dNTPs, 0.5 µL Taq DNA polymerase (TaKaRa, Japan), and 36.5 µL ultrapure water. Amplification reactions were performed in a Bio-Rad iCycler (Bio-Rad Laboratories, Hercules, CA, USA) using the following PCR conditions: 5 min at 94°C, followed by 35 cycles with 1 min of DNA denaturation at 94°C, 1 min of annealing at 54℃, and 1 min of extension at 72°C, followed by a final extension at 72°C for 10 min. The PCR-amplified products were checked using 6% denaturing polyacrylamide gel electrophoresis before sequencing. All PCR products were sequenced in both directions using a capillary sequencer (ABI 3100; Applied Biosystems, Foster City, CA, USA). Sequences were aligned using Clustal X (Thompson et al., 1997) or Clustal W with MEGA 5.0 (Kumar et al., 2008).

Analysis of protein functional domains was performed using conserved domain database (CDD) (Marchler-Bauer et al., 2005). The distance matrix was subjected to cluster analysis by using the unweighted pair group method with arithmetic mean (UPGMA) (Sneath and Sokal, 1973), a SHAN (sequential hierarchical agglomerative nested) clustering technique (Sneath and Sokal, 1973) that compresses the patterns of variation into two-dimensional branch diagrams (dendrograms). A

dendrogram was constructed using the NTSYSpc 2.1 statistical package (Rohlf, 2000).

RESULTS

Sequencing analysis using known *FAD2* sequences from *Cucurbita pepo* showed 27 highly similar *FAD2* clones in *C. moschata*, *C. maxima*, *C. pepo*, and *C. ficifolia* (totally 1152 bp). Nucleotide insertion or deletion did not occur between the sequences of the 27 selected clones, and the polymorphisms were attributable only to nucleotide substitution.

The predicted amino acid sequences from the open reading frames of the 27 *FAD2* clones provided evidence that the 27 pumpkin *FAD2* genes all encoded FAD2 enzymes with 384 amino acids, and these enzymes showed an amino acid sequence identity between the four pumpkin species, ranging from 91.7 to 100% over a stretch of 352 amino acids. Analysis of protein functional domains of the FAD2 enzymes indicated that these enzymes had the same main-function domain: the functional domain of delta-12 fatty acid desaturase, from 47 to 329 amino acids, that was identified using the CDD.

The UPGMA was used to classify the 27 accessions into four major clusters with a similarity coefficient (0.7) (Figure 1). The dendrogram also showed the clear separation among *C. moschata*, *C. maxima*, *C. pepo*, and *C. ficifolia*. The four major clusters corresponded to the four seeded pumpkin species. In cluster I (*C. maxima*), three accessions (6, 8, and 12) from Russia were grouped into a sub-cluster, and two accessions (9 and 15) with a unique late-maturity trait shared the same sub-cluster. Similarly, in cluster II (*C. moschata*), three accessions (1, 2, and 3) from China and two accessions (4 and 5) from Russia were separately grouped into two different sub-clusters. In cluster III (*C. pepo*), three accessions (18, 19, and 20) with the same geographic origin (Russia) were closely clustered together.

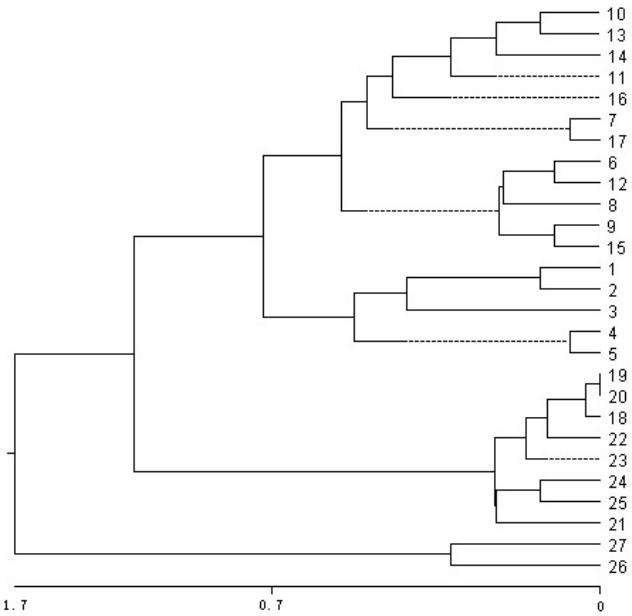

Figure 1. Dendrogram showing the relationships among 27 accessions of *Cucurbita moschata*, *Cucurbita maxima*, *Cucurbita pepo*, and *Cucurbita ficifolia* with the nucleotide sequences of the 27 *FAD2* clones.

DISCUSSION

The sequences of the 27 selected clones shared more than 90% similarity, which implied that *FAD2* was highly conserved, and the intraspecific and interspecific divergences of *FAD2* were all low for *C. moschata*, *C. maxima*, *C. pepo*, and *C. ficifolia*. Only one copy of *FAD2* was detected in each accession of the seeded pumpkins, which was distinguishable from other crops because they possessed several copies of *FAD2* (Hutcheon et al., 2010; Pham et al., 2010; Guan et al., 2012; Yang et al., 2012; Cao et al., 2013).

Analysis of protein functional domains of the FAD2 indicated that all 27 amino acid sequences included three conservative histidine clusters, namely, HECGH (105-109), HRRHH (141-145), and HVAHH (315-319). Previous studies have also suggested that the protein functional domain of the FAD2 enzyme is composed of three conservative histidine clusters; any mutation in this domain would lead to the inactivation of the FAD2 enzyme (Shanklin et al., 1994). Further analyses showed that four (Valine, Alanine, Leucine, Valine) of the eight amino acids (Threonine, Histidine, Valine, Alanine, Histidine, Histidine, Leucine, Valine) belonged to hydrophobic residues in the third histidine cluster and the neighboring amino acid sequences, which implied that these hydrophobic residues may be located in the interior of the active site on the spatial structure. The top ten amino acids were all found to be leucine in the second histidine cluster, which is consistent with the findings of a previous study (Tanhuanpää et al., 1995). This structure might be one of the active sites of the FAD2 enzyme. Moreover, threonine was identified as the last three residues in the second histidine cluster and as the last four residues in the third histidine cluster, which may play a significant role in the desaturation and hydroxylation of the FAD2 enzyme, and thus, be another active site (Broadwater et al., 2002).

The genetic diversity of seeded pumpkin landraces suggests that the species itself, as well as relatively isolated growing zones, could prevent *FAD2* from flowing, to some extent, and affect the mutation of *FAD2*. However, some *C. maxima* accessions from different geographic origins were still observed to group into the same sub-cluster. For example, two accessions (9 and 15) from China first grouped with three accessions (6, 8, and 12) from Russia, rather than with other accessions from China. Seed exchange among farmers and/or breeders and the out-crossing nature of *C. maxima* may have contributed to this confounding grouping phenomenon (Liu et al., 2013).

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Research supported by a grant from Research Fund for Foundation of Heilongjiang Educational Committee (#12531043) and the Public Welfare Industry (Agricultural) Research Special Foundation of the Agricultural Ministry of China (#201303112).

REFERENCES

Beló A, Zheng P, Luck S, Shen B, et al. (2008). Whole genome scan detects an allelic variant of fad2 associated with increased oleic acid levels in maize. *Mol. Genet. Genomics* 279: 1-10.

Broadwater JA, Whittle E and Shanklin J (2002). Desaturation and hydroxylation. Residues 148 and 324 of *Arabidopsis* FAD2, in addition to substrate chain length, exert a major influence in partitioning of catalytic specificity. *J. Biol. Chem.* 277: 15613-15620.

Browse J and Somerville C (1991). Glycerolipid synthesis: biochemistry and regulation. *Annu. Rev. Plant. Physiol. Plant Mol. Biol.* 42: 467-506.

Cao SJ, Zhou XR, Wood CC, Green AG, et al. (2013). A large and functionally diverse family of *FAD2* genes in safflower (*Carthamus tinctorius* L.). *BMC Plant Biol.* 13:5.

Chen W, Song K, Cai YR, Li WF, et al. (2011). Genetic modification of soybean with a novel grafting technique: downregulating the *FAD2-1* gene increases oleic acid content. *Plant Mol. Bio. Rep.* 29: 866-874.

Chen ZB, Wang ML, Barkley NA and Pittman RN (2010). A simple allele-specific PCR assay for detecting *FAD2* alleles in both A and B genomes of the cultivated peanut for high-oleate trait selection. *Plant Mol. Bio. Rep.* 28: 542-548.

Chi XY, Yang QL, Lu YD, Wang JY, et al. (2011). Genome-wide analysis of fatty acid desaturases in soybean (*Glycine max*). *Plant Mol. Bio. Rep.* 29: 769-783.

Doyle JJ and Doyle JL (1990). Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15.

Georgios B, Anastassios M, Nikos N and Polydefkis H (2005). Spatial and temporal expressions of two distinct oleate desaturase from olive (*Olea europaea* L.). *Plant Sci.* 168: 547-555.

Guan LL, Wang YB, Shen H, Hou K, et al. (2012). Molecular cloning and expression analysis of genes encoding two microsomal oleate desaturases (*FAD2*) from Safflower (*Carthamus tinctorius* L.). *Plant Mol. Biol. Rep.* 30: 139-148.

Harwood JL (1997). Plant lipid metabolism; Dey PM, Harborne JB, Eds.; Plant biochemistry: Academic Press: 237-272.

Hutcheon C, Ditt RF, Beilstein M, Comai L, et al. (2010). Polyploid genome of *Camelina sativa* revealed by isolation of fatty acid synthesis genes. *BMC Plant Bio.* 10: 233.

Jung S, Swift D, Sengoku E, Patel M, et al. (2000). The high oleate trait in the cultivated peanut [*Arachis hypogaea* L.]. I. Isolation and characterization of two genes encoding microsomal oleol-PC desaturases. *Mol. Gen. Genet.* 263: 796-805.

Krasowska A, Dziadkowiec D, Polinceusz A, Plonka A, et al. (2007). Cloning of flax oleic fatty acid desaturase and its expression in yeast. *J. Am. Oil. Chem. Soc.* 84: 809-816.

Kumar S, Nei M, Dudley J and Tamura K (2008). MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief Bioinform.* 9: 299-306.

Li LY, Wang XL, Gai JY and Yu DY (2007). Molecular cloning and characterization of a novel microsomal oleate desaturase gene from soybean. *J. Plant Physiol.*164: 1516-1526.

Liu C, Ge Y, Wang DJ, Li X, et al. (2013). Morphological and molecular diversity in a germplasm collection of seed pumpkin. *Sci. Hortic.* 154: 8-16.

Lopez Y, Nadaf HL, Smith OD, Connell JP, et al. (2000). Isolation and characterization of the Δ12-fatty acid desaturase in peanut (*Arachis hypogaea* L.) and search for polymorphisms for the high oleate trait in spanish market-type lines. *Theor. Appl. Genet.*101: 1131-1138.

Marchler-BauerA, Anderson JB, Cherukuri PF, DeWeese-Scott C, et al. (2005). CDD: a Conserved Domain Database for protein classification. *Nucleic Acids Res*. 33: 192-196.

Mikkilineni V and Rocheford TR (2003). Sequence variation and genomic organization of fatty acid desaturase-2 (*FAD2*) and fatty acid desaturase-6 (*FAD6*) cDNAs in maize. *Theor. Appl. Genet.* 106: 1326-1332.

Mozaffarian D, Katan MB, Ascherio A, Stampfer MJ, et al. (2006). Trans fatty acids and cardiovascular disease. *N. Engl. J. Med.*15: 1601-1613.

Okuley J, Lightner J, Feldmann K, Yadav N, et al. (1994). *Arabidopsis FAD2* gene encodes the enzyme that is essential for polyunsaturated lipid synthesis. *Plant Cell* 6: 147-158.

Pham AT, Lee JD, Shannon JG and Bilyeu KD (2010). Mutant alleles of *FAD2-1A* and *FAD2-1B* combine to produce soybeans with the high oleic acid seed oil trait. *BMC Plant Biol.* 10: 195.

Qu J, Mao HZ, Chen W, Gao SQ, et al. (2012). Development of marker-free transgenic Jatropha plants with increased levels of seed oleic acid. *Biotechnol. Biofuels* 5:10.

Qu SP, Chang Y, Qin JF, Xing W, et al. (2011). Accumulation pattern of fatty acids during pumpkin (*Cucurbita maxima* Duch) seed development. *Chin. J. Oil Crop Sci*. 33: 628-631.

Rohlf FJ (2000) NTSYSpc: Numerical Taxonomy and Multivariate Analysis System. Version 2.1. Exeter Software, New York.

Rolletschek H, Borisjuk L, Sánchez-García A, Gotor C, et al. (2007). Temperature-dependent endogenous oxygen concentration regulates microsomal oleate desaturase in developing sunflower seeds. *J. Exp. Bot.* 58: 3171-3181.

Schlueter JA, Vasylenko-Sanders IF, Deshpande S, Yi J, et al. (2007). The *FAD2* gene family of soybean: Insights into the structural and functional divergence of a paleopolyploid genome. *Crop Sci.* 47: S-14-26.

Shanklin J, Whittle E and Fox BG (1994). Eight histidine residues are catalytically essential in a membrane-associated ironenzyme, stearoyl-COA desaturase, and are conserved in alkane hydroxylaseand xylene monooxygenase. *Biochemistry* 33: 12787-12794.

Sneath PHA and Sokal RR (1973). Numerical taxonomy; Freeman, W.H.; San Francisco.

Tanhuanpää PK, Vilkki JP and Vilkki HJ (1995). Identification of a RAPD marker for palmitic-acid concentration in the seed oil of spring turnip rape (*Brassica rapa* ssp.*oleifera*). *Theor. Appl. Genet.* 91: 477-480.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, et al. (1997). The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25: 4876-4882.

Wang MH, Liu MX, Li DK, Wu J, et al. (2010). Overexpression of FAD2 promotes seed germination and hypocotyl elongation in *Brassica nap*us. *Plant Cell Tiss. Organ. Cult.* 102: 205-211.

Yang QY, Fan CC, Guo ZH, Qin J, et al. (2012). Identification of *FAD2* and *FAD3* genes in *Brassica napus* genome and development of allele-specific markers for high oleic and low linolenic acid contents. *Theor. Appl. Genet.* 125: 715-729.

Zhang DY, Pirtle IL, Park SJ, Nampaisansuk M, et al. (2009). 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. *Plant Physiol. Biochem.* 47: 462-471.

 ©FUNPEC-RP www.funpecrp.com.br