JPMB | Journal of Plant Molecular Breeding
Published by Genetics and Agricultural Biotechnology Institute of Tabarestan
2023 | Volume 11 | Issue 1

Comparison of predicted protein sequences of the omega-3 fatty acid desaturase gene family in some of the oil seed crops

Original paper

Ammar Afkhami Ghadi *[1], Seyyed Jaber Hosseini [2] and Ali Ghanbari [3]

[1.] Department of Genetic Engineering and Plant Breeding, Imam Khomeini International University, Qazvin, Iran
[2.] Department of Agronomy, Tarbiat Modares University, Tehran, Iran
[3.] Department of Plant Production Engineering and Genetics, Zanjan University, Zanjan, Iran

Abstract: This study aimed to conduct a comparative analysis of protein-predicted sequences within the omega-3 fatty acid desaturase (FAD) gene family across various oilseed crops, such as cotton, soybean, rapeseed, and corn. Twenty-five omega-3 FAD genes were distinguished by the removal of similar sequences. Phylogenetic analysis of the omega-3 FAD gene family was detected. Omega-3 FAD gene information, and physical parameters including the number of amino acids, chromosome locations, exon count and other features were obtained. The results showed that the average length of the proteins encoded by the omega-3 fatty acid desaturase proteins was 388.58 amino acid. The molecular weights of these proteins ranged from 22.2 to 51.3 kDa. The phylogenetic tree divided the omega-3 FAD proteins into five clades. Clade 2 comprises the smallest number of omega-3 fatty acid desaturase gene members, whereas clade 1 encompasses the highest number of members. We identified five pairs of orthologous genes among the omega-3 FAD genes and identified a total of twenty distinct conserved motifs. Four of these motifs were associated with encoding the FAD domain, while an additional four were related to the DUF3474 domain. Undoubtedly, characterizing these FADs could offer valuable candidate genes for enhancing new oilseed varieties.

Keywords: conserved motifs, gene family, oil seed plants, omega-3 fatty acid

OPEN ACCESS

Edited by
Dr. Parviz Heidari,
Shahrood University of Technology,
Iran

Date
Received: 1 August 2023
Accepted: 19 October 2023
Published: 28 January 2024

Correspondence
Dr. Ammar Afkhami Ghadi
a.afkhami@sanru.ac.ir

Citation
Afkhami Ghadi, A., Hosseini, S. J., and Ghanbari, A. (2023). Comparison of protein predicted sequences of the omega-3 fatty acid desaturase gene family in some of the oil seed crops. *J. Plant Mol. Breed 11* (1): 28-40. doi: 10.22058/JPMB.2023.2008246.1281.

Introduction

A gene family consists of a group of genes that share significant characteristics. In numerous instances, genes within a family share identical sequences and may possess indistinguishable functions. These genes offer instructions for generating products, such as proteins, that share common components or biochemical functions. Alternatively, in some cases, heterologous genes are categorized into a family because the proteins they produce function together as a cohesive unit or participate in analogous biological processes. Researchers can employ gene families to forecast the functions of newly discovered genes by assessing their resemblance to established genes. Additionally, the similarities among genes within a family can aid to predict the locations and temporal patterns of expression of a particular gene.

Oil seed plants are important crops worldwide, and provide oilseeds for food, feed, and biofuel. Recently, FAD genes have also been identified in several oilseed crops, including olives (Hernández et al., 2016) cotton (Chi et al., 2011; Yurchenko et al., 2014; Liu et al., 2015) and soybean (Chi et al., 2011; Román et al., 2012). Fatty acid desaturase in oil seed plants is an enzyme that removes two hydrogen atoms from a fatty acid, creating a carbon/carbon double bond (Shanklin and Cahoon, 1998). These desaturases are classified as delta indicating that a double bond is created between the third and fourth carbon atoms of the methyl end of the fatty acid. Methyl end desaturases are responsible for produce FAs with two and three double bonds, establishing the n-6 and n-3 series of PUFAs (Alloatti and Uttaro, 2011) In contrast a front-end desaturase inserts a double bond between the pre-existing double bond and the carboxyl end of a fatty acid (Meesapyodsuk and Qiu, 2012).

Information about the fatty acid desaturase genes encoding these enzymes has permitted more detailed analyses of the role of these genes, and their fatty acid products, in plant lipid metabolism and abiotic stress responses. For example, omega-3 fatty acids are known to increase in plants in response to both drought (Martin et al., 1986) and cold temperature (Williams et al., 1988; Guschina and Harwood, 2006) and overexpression of omega-3 desaturases in several transgenic plants has been shown to increase both drought and chilling tolerance (Zhang et al., 2005; Domínguez et al., 2010). The endoplasmic reticulum (ER-localized) desaturases FAD2 and FAD3 are also involved in the production of PUFA components of seed oils (Ohlrogge and Browse, 1995). Given the value of these fatty acids to human nutrition, and to determine the stability of oils during cooking or other food applications, molecular markers for these genes have been developed to evaluate germplasm and identifying oilseed varieties with improved oil compositions (Bocianowski et al., 2012; Yang et al., 2012). a genome wide analysis of Gossypium's omega-3 fatty acid desaturase gene family indicated that 11 omega-3 FAD genes were identified (Ohlrogge and Browse, 1995; Yurchenko et al., 2014). They found that the omega-3 FAD family of cotton includes five distinct genes, two of which encode endoplasmic reticulum-type enzymes and three encode chloroplast-type enzymes. In soybean, the expression of FAD3 and FAD7 is tightly regulated in response to cold temperature (Román et al., 2012). This study aimed to compare the predicted protein sequences of the omega-3 fatty acid desaturase (FAD) gene family in oil seed crops.

Materials and Methods

Database search for properties of omega-3 fatty acid desaturase genes

Six species' omega-3 fatty acid desaturase genes (*Gossypium raimondii, Gossypium arboretum, Glycine soja, Brassica napus, Brassica juncea,* and *Zea mays*) were obtained from the National Centre for Biotechnology Information (NCBI) databases (**http://www.ncbi.nlm.nih.gov/**). Twenty-five Omega-3 fatty acid desaturase genes (5 *Gossypium raimondii*, 2 *Gossypium arboreum*, 9 *Glycine soja*, 3 *Zea mays*, 5 *Brassica napus*, and 1 *Brassica juncea*) were identified. The differentiation was achieved by removing similar sequences. The accession numbers of published omega-3 fatty acid desaturase genes from oil seed plants in this study are listed in Table 1. Phylogenetic analysis of comparative and functional genomics of the Omega-3 fatty acid desaturase gene family was detected of GreenPhylDB. This database contains a catalog of gene families based on gene predictions

of genomes, covering a broad taxonomy of green plants. Omega-3 fatty acid desaturase gene information, including the number of amino acids, chromosome locations, exon count and location coordinates (5′-3′) was obtained from the NCBI database. Physical parameters of the omega-3 fatty acid desaturase proteins, including molecular mass (kDa), and isoelectric point (pI) were calculated using compute pI/Mw tool in ExPASy (http://www.expasy.org/ tools/), with parameters (resolution) set to 'average' (Gasteiger et al., 2005). Several cysteine residues, aliphatic index, instability index, and grand average of hydropathicity were predicted using the ProtParam tool (Gasteiger et al., 2005). The aliphatic index of a protein is defined as the relative volume occupied by the aliphatic side chains of alanine, valine, isoleucine, and leucine residues. Proteins with a calculated instability index smaller than 40 were predicted as stable, while those with a value above 40 were considered as unstable. ProtParam calculates the GAH value as the sum of the amino acids' hydropathy value divided by the sequence's number of residues (Kyte and Doolittle, 1982).

Table 1. List of Omega-3 fatty acid desaturase genes identified in *Gossypium raimondii, Glycine soja, Brassica napus, Brassica juncea, Zea mays, Gossypium arboretum,* and their predicted physicochemical properties.

Accession no.	Species	Length (aa)	MW (Da)	PI	Instability index	Alphatic index
gi|728424424	*Gossypium*	435	50497.15	7.81	43.05(unstable)	87.33
gi|728424414	*Gossypium*	450	51171.86	8.91	36.18(unstable)	85.58
gi|728424404	*Gossypium*	446	50716.97	8.52	35.56(stable)	83.05
gi|728424394	*Gossypium*	376	43588.11	8.68	38.55(stable)	86.06
gi|728424384	*Gossypium*	388	45029.85	9.05	39.84(stable)	79.36
gi|408794	*Glycine soja*	380	44184.89	8.49	38.18(stable)	93.37
gi|734428355	*Glycine soja*	453	51215.88	8.17	33.80(stable)	90.97
gi|7734408411	*Glycine soja*	452	51376.92	7.40	38.56(stable)	87.94
gi|7734373505	*Glycine soja*	453	51247.81	8.45	32.78(stable)	88.01
gi|734344277	*Glycine soja*	453	51301.66	7.39	38.04(stable)	85.85
gi|734407330	*Glycine soja*	381	44147.63	8.72	33.04 (stable)	86.72
gi|734419357	*Glycine soja*	312	36753.14	7.48	36.50 (stable)	89.62
gi|734329362	*Glycine soja*	370	42763.07	8.87	37.41 (stable)	86.41
gi|734427032	*Glycine soja*	195	22236.48	8.93	46.83(unstable)	82.97
gi|728838741	*Gossypium*	311	35841.13	8.33	35.53 (stable)	92.70
gi|728837144	*Gossypium*	330	38132.82	8.79	33.02 (stable)	91.82
gi|195627062	*Zea mays*	443	49385.72	9.16	46.37 (unstable)	85.51
gi|195635609	*Zea mays*	384	44072.42	8.69	38.79(stable)	85.52
gi|195612756	*Zea mays*	408	45905.81	8.87	44.43 (unstable)	90.78
gi|408492	*Brassica napus*	377	43258.58	7.85	29.55 (stable)	90.74
gi|408490	*Brassica napus*	329	38089.77	6.59	25.72 (stable)	92.13
gi|49355342	*Brassica napus*	378	43377.84	8.52	27.62 (stable)	92.04
gi|47028567	*Brassica napus*	383	43850.36	7.86	33.08 (stable)	91.59
gi|46849975	*Brassica napus*	439	50273.62	8.21	34.81 (stable)	86.38
gi|7378667	*Brassica juncea*	429	49174.49	8.42	31.54 (stable)	83.85

Solubility upon overexpression, disulfide bond formation and potential glycosylation sites

Solubility upon overexpression in *Escherichia coli* and disulfide bond formation were predicted using SOLpro (Magnan et al., 2009) and DIpro (Scratch protein predictor), respectively (Cheng et al., 2006). Potential N-glycosylation sites were determined using the NetNGlyc 1.0 (Gupta and Brunak, 2001).

Phylogenetic analysis of members of omega-3 fatty acid desaturase family

Multiple sequence alignments of the full-length protein sequences from cotton, soybean, rapeseed, and corn were performed using MEGA 7.0 (Kumar et al., 2016) with default parameters. A phylogenetic tree based on the alignment was constructed using MEGA 7.0 and the Neighbor-Joining (NJ) method (Saitou and Nei, 1987). The Maximum Parsimony (MP) method (Tamura et al., 2011) was also used to create a phylogenetic tree and to validate the results from the N-J method. Bootstrap analysis was performed using 1000 replicates in pairwise gap deletion mode, which allows divergent domains to contribute to the topology of the NJ tree (Hu et al., 2010).

Conserved motif analysis

Conserved protein motifs were analyzed Online MEME (Multiple Expectation Maximization for Motif Elicitation) (Bailey and Elkan, 1995). The parameters were as follows: The number of repetitions for any, with maximum number of motifs = 20, and the optimum motif width was constrained to between 6 and 200 residues. In addition, structural motif annotation was performed using Pfam (http://pfam.sang er.ac.uk /search) and SMART (http://smart.embl-heidelb erg.de/) tools. Putative conserved domains within the omega-3 fatty acid desaturase were identified using the NCBI conserved domain database (CDD) (Marchler-Bauer et al., 2005; Wang et al., 2023).

Results

Physical properties and chromosomal distribution of omega-3 fatty acid desaturase gene family in five species genomes

Twenty-five genes were identified as members of the omega-3 fatty acid desaturase gene family, seven genes in cotton, nine genes in soybean, three genes in corn and six genes in rapeseed. These findings determined the physical location of individual omega-3 fatty acid desaturase genes on chromosomes. The chromosome location of the omega-3 fatty acid desaturase genes indicated that this gene in *Gossypium raimondii* were distributed on five out of 26 chromosomes (chr 1, 2, 6, 11 and 12). The chromosomal locations of some genes were unknown. Among the 25 genes, gi|734428355, gi|7734373505 and gi|734344277 encodes the longest protein (453 amino acids (aa)), while gi|734427032 encoded the shortest (195 aa) belongs to *Glycine soja*. The average length of the proteins encoded by omega-3 fatty acid desaturase proteins was 388.58 amino acid. The molecular weights of these proteins ranged from 22.2 kDa to 51.3 kDa, with an average of 44.7 kDa. Furthermore, the theoretical pI values of all proteins except gi|408490 were above 7, indicating that they were alkaline. The 25 omega-3 fatty acid desaturase genes contained different numbers of exons, ranging from 5 to 9. Furthermore, eight omega-3 fatty acid desaturase genes were found to possess eight exons. Detailed parameters are listed in Table 1. According to Protparam as shown in Table 1, the most unstable Omega-3 fatty acid desaturase gene is probably obtained from gi|734427032 and the most stable form is gi|408490. Stable enzymes can be used for a longer time to biocatalize a reaction. The most alphatic index omega-3 fatty acid desaturase gene is obtained from gi|408794. It may be regarded as a positive factor for the increase of thermostability of globular proteins. The chloroplast transit peptide (CTP) of these proteins ranged from 0.002 to 0.904, with an average of 0.342 (Table 2). CTPs can be used to transport non-chloroplastic proteins into the chloroplasts.

Solubility upon overexpression, disulfide bond formation and potential glycosylation sites

The omega-3 fatty acid desaturase protein from gi|734419357 and gi|728424394 were predicted to have the highest probabilities of solubility and insolubility, respectively (Table 3). Of 25 FAD, only eight sequences (three *G. raimondii,* two G. *soja,* one in each of *G. arboretum, Z. mays* and *B. juncea*) were predicted to be soluble while the other 17 FAD were predicted to be insoluble upon overexpression in E. coli. Table 3 shows the prediction by DIpro of

disulfide bond formation for different FAD. Most of FAD were predicted to have the lowest number (only one) of disulfide bonds (Table 3). Enzymes from G. *soja* (gi|734344277 and gi|7734408411) contained the highest amount of cysteine residues, 7 and 6, respectively. This means that according to our prediction, recombinant production of these FAD might be performed using expression systems that provide a suitable redox medium for disulfide bond formation. FAD from *Z. mays* (gi|195635609), has no potential Nglycosylation site. The highest numbers of N-glycosylation site were found for FAD from gi|46849975 and gi|7378667 (9 potential sites), and gi|728424384 (8 potential sites).

Table 2. Predicts the subcellular location of Omega-3 fatty acid desaturase proteins sequences.

Accession no.	Len	cTP	mTP	SP	other	Loc	RC
gi|728424424	435	0.193	0.370	0.043	0.233	M	5
gi|728424414	450	0.844	0.052	0.016	0.056	C	2
gi|728424404	446	0.831	0.037	0.011	0.159	C	2
gi|728424394	376	0.154	0.134	0.020	0.878	-	2
gi|728424384	388	0.021	0.091	0.155	0.976	-	1
gi|408794	380	0.051	0.194	0.017	0.846	-	2
gi|734428355	453	0.734	0.085	0.030	0.097	C	2
gi|7734408411	452	0.788	0.181	0.014	0.091	C	2
gi|7734373505	453	0.705	0.110	0.027	0.087	C	3
gi|734344277	453	0.923	0.188	0.008	0.047	C	1
gi|734407330	381	0.037	0.163	0.032	0.936	-	2
gi|734419357	312	0.004	0.066	0.938	0.057	S	1
gi|734329362	370	0.096	0.112	0.022	0.894	-	2
gi|734427032	195	0.098	0.338	0.228	0.278	M	5
gi|728838741	311	0.002	0.133	0.968	0.091	S	1
gi|728837144	330	0.002	0.133	0.968	0.091	S	1
gi|195627062	443	0.566	0.765	0.002	0.007	M	5
gi|195635609	384	0.131	0.137	0.029	0.871	-	2
gi|195612756	408	0.533	0.426	0.237	0.030	C	5
gi|408492	377	0.078	0.171	0.018	0.856	-	2
gi|408490	329	0.012	0.180	0.935	0.042	S	2
gi|49355342	378	0.078	0.171	0.018	0.856	-	2
gi|47028567	383	0.049	0.161	0.024	0.936	-	2
gi|46849975	439	0.904	0.162	0.001	0.021	C	2
gi|7378667	429	0.718	0.260	0.005	0.097	C	3

The location assignment is based on the predicted presence of any of the N-terminal presequences: Sequence length (**Len**), chloroplast transit peptide (**cTP**), mitochondrial targeting peptide (**mTP**) or secretory pathway signal peptide (**SP**), Prediction of localization (**Loc**), Reliability class (**RC**).

Table 3. Prediction of solubility upon overexpression, disulfide bond formation and N-glycosylation of linoleic acid isomerases (LAI).

Accession no.	Length (aa)	Solubility on Overexpression [a]	Cys	Disulfide bonds [b]	N glycosylation [c]
gi|728424424	435	Insoluble (0.712733)	4	1 (42-103)	6 (29, 71, 91, 171, 222, 257)
gi|728424414	450	Soluble (0.507310)	5	2 (9-42,115-171)	3 (30, 64, 183)
gi|728424404	446	Soluble (0.500000)	3	1 (113-169)	6 (37, 47, 57, 81, 254, 258)
gi|728424394	376	Insoluble (0.720550)	4	1 (67-94)	5 (106, 137, 150, 190, 194)
gi|728424384	388	Soluble (0.709559)	2	1 (39-155)	8 (43, 117, 122, 152, 168, 205, 209, 355)
gi|408794	380	Insoluble (0.597998)	4	1 (85,102)	3 (14, 162, 198)
gi|734428355	453	Insoluble (0.600799)	5	2 (9-52,118-174)	4 (39, 53, 106, 186)
gi|773440841	452	Insoluble (0.572344)	6	2 (9-27,117-173)	5 (39, 46, 56, 147, 236)
gi|773437350	453	Insoluble (0.616692)	5	2 (9-52,117-173)	6 (39, 53, 83, 105, 185, 219)
gi|734344277	453	Soluble (0.569171)	7	2 (9-27,118-174)	5 (39, 46, 84, 88, 237)
gi|734407330	381	Insoluble (0.698000)	3	1 (44-100)	7 (12, 22, 74, 160, 196, 200, 381)
gi|734419357	312	Soluble (0.713550)	2	1 (34-150)	2 (94, 130)
gi|734329362	370	Insoluble (0.663430)	3	1 (33-89)	5 (2, 149, 185, 189, 370)
gi|734427032	195	Insoluble (0.535586)	4	1 (186-193)	5 (14, 24, 72, 142, 158)
gi|728838741	311	Soluble (0.597674)	3	1 (21-48)	6 (60, 88, 91, 104, 144, 148)
gi|728837144	330	Insoluble (0.524630)	3	1 (21-48)	6 (60, 88, 91, 104, 144, 148)
gi|195627062	443	Insoluble (0.650700)	4	2 (9-10,107-163)	2 (175, 259)
gi|195635609	384	Soluble (0.572108)	5	2 (38-94,210-212)	0
gi|195612756	408	Insoluble (0.762427)	5	2 (35-73,129-245)	4 (141, 157, 214, 357)
gi|408492	377	Insoluble (0.617692)	3	1 (38-94)	3 (12, 190, 377)
gi|408490	329	Insoluble (0.612866)	3	1 (164-174)	4 (22, 23, 60, 156)
gi|49355342	378	Insoluble (0.645880)	3	1 (38-94)	3 (12, 191, 378)
gi|47028567	383	Insoluble (0.698698)	3	1 (44-100)	5 (10, 12, 160, 196, 383)
gi|46849975	439	Insoluble (0.614346)	4	1 (102-158)	9 (3, 29, 46, 50, 57, 90, 132, 170, 266)
gi|7378667	429	Soluble (0.514721)	5	2 (95-151,267-277)	9 (3, 28, 42, 52, 83, 125, 163, 247, 259)

a) Probability values are presented within parenthesis.
b) The position of disulfide bonds is written within the parenthesis.
c) The position of N -glycosylation sites is written within the parenthesis.

Phylogenetic analysis of omega-3 fatty acid desaturase family in cotton, soybean, rapeseed and corn

To estimate the similarity and evolutionary ancestry of omega-3 fatty acid desaturase genes in cotton, soybean, rapeseed and corn, we created an unrooted phylogenetic tree of the 25 omega-3 fatty acid desaturase protein sequences. The phylogenetic tree divided the omega-3 fatty acid desaturase proteins into five clades (Figure 1). Clade 2 has the fewest omega-3 fatty acid desaturase gene members (2), in contrast clade 1 contains the most members (12), followed by clades 3 and 4 (3) and clade 5 (5). The four species (except *G. arboreum*) contributed at least one omega-3 fatty acid desaturase gene to clade 1, while the members of clades 2, 3, 4 and clade 5 included one two or three species, for example, clade 5 consisted of *G. raimondii, Z. mays and B. napus*, which may correspond to some special in the evolutionary process.

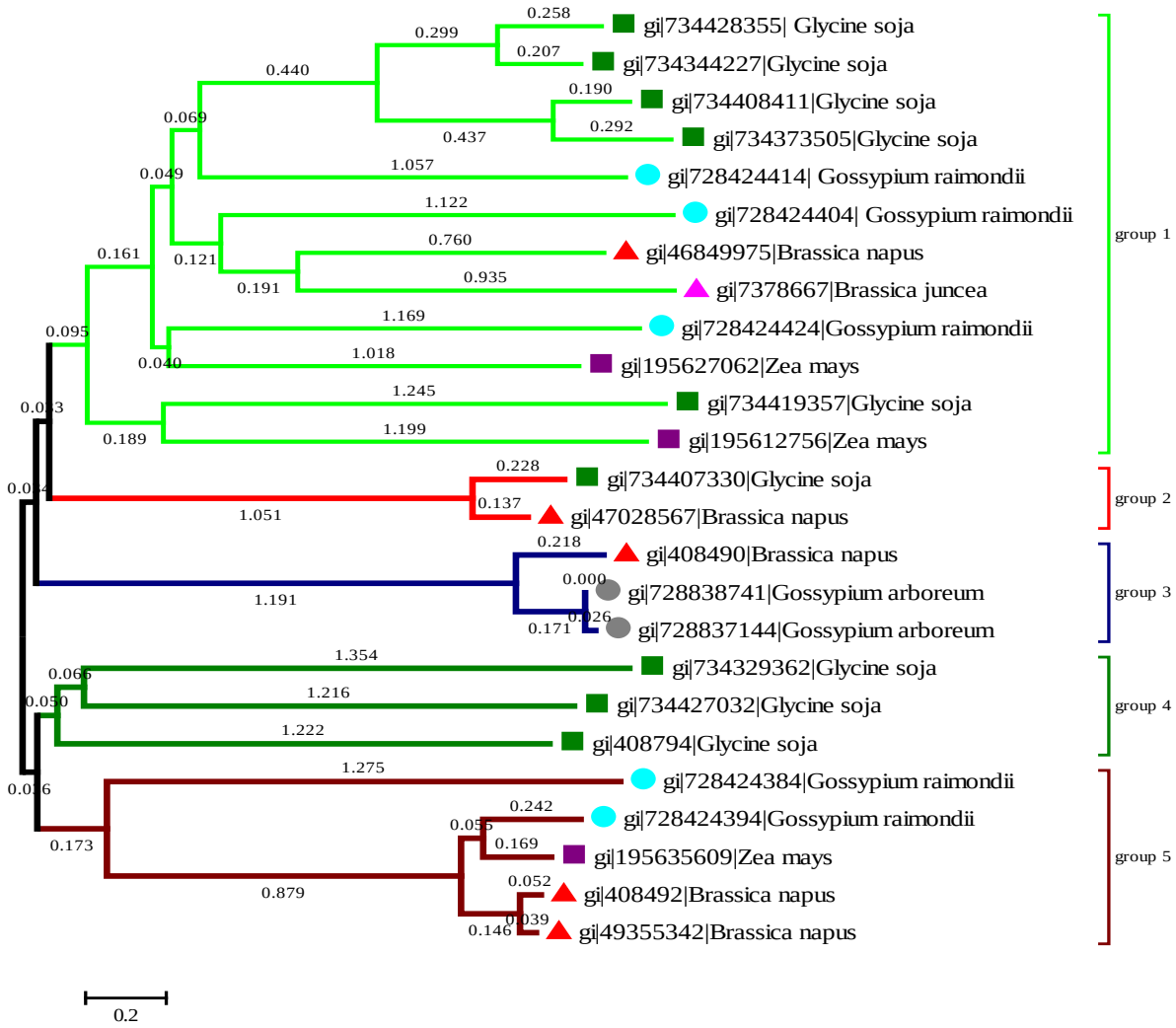

Figure 1. Phylogenetic tree of full-length Omega-3 fatty acid desaturase proteins from *G.soja* ( ), *B. napus* ( ), *B. juncea* ( ), *Z. mays* ( ), *G. arboretum* ( ) and *G. raimondii* ( ).

Based on phylogenetic analysis, five pairs of orthologous genes were identified among the omega-3 fatty acid desaturase genes: gi/195635609 and gi/728424394, gi/47028567 and gi/734407330, gi/195612756 and gi/734419357, gi/195627062 and gi/728424424, and gi/7378667 and gi/46849975. The remaining genes of the omega-3 fatty acid desaturase gene family were represented by paralogous pairs.

Conserved motifs of omega-3 fatty acid desaturase genes

Twenty distinct conserved motifs were identified, and the conserved amino acid sequences and length of each motif are shown in Figure 2. Each putative motif obtained from MEME was annotated by searching Pfam and SMART. While four motifs were found to encode the fatty acid desaturases domain, four different domains were found related to the DUF3474 domain, while the other motifs did not function in the annotation. Fatty acid desaturases are enzymes that catalyze the insertion of a double bond at the delta position of fatty acids. Additionally, the DUF3474 domain remains functionally uncharacterized (Figure 3). This domain is presented in both bacteria and eukaryotes. This domain is typically 126 - 140 amino

acids in length. Motifs 3 and 4 were the most common motifs, found in all twenty-five omega-3 fatty acid desaturase genes. Motif 15 was unique to the proteins in clade 4 and unique motifs 18 were found in clade 3; These motifs might be important to the functions of unique omega-3 fatty acid desaturase protein.

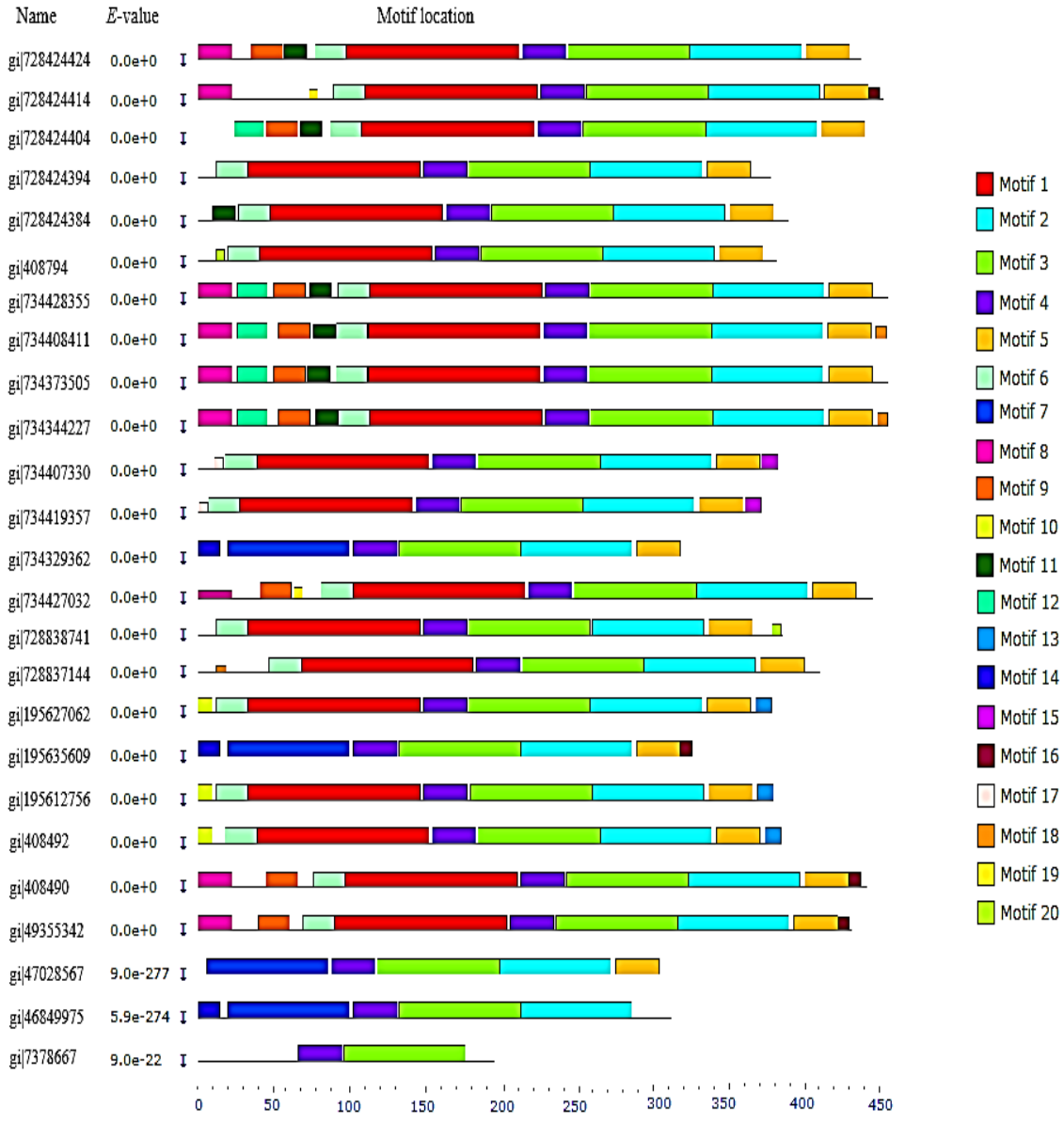

Figure 2. Distribution of conserved motifs in the omega-3 fatty acid desaturase family members. All motifs were identified by MEME using the complete amino acid sequences. Names of all members among the defined genes are shown on the left side of the figure, and motif lengths are indicated at the bottom of the figure. The positions of Zn-finger domains predicted by the SMART tool. Database are indicated by vertical tick marks below each protein model. The different-colored boxes represent different motifs and their position in each omega-3 fatty acid desaturase sequence.

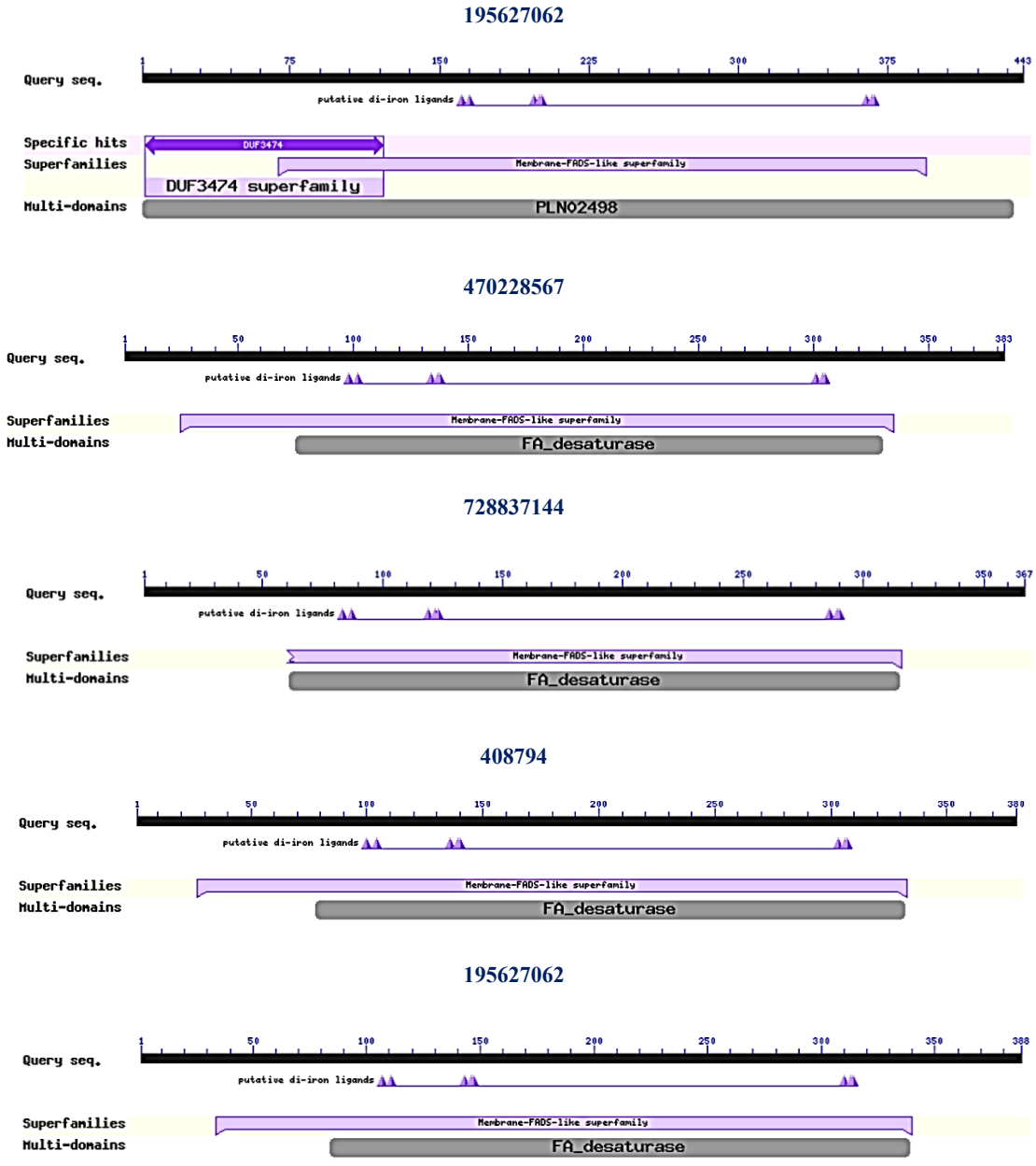

Figure 3. Conserved domain search on Omega-3 fatty acid desaturase gene sequences.

1

Discussion

Omega-3 fatty acid desaturase gene family is involved in the regulation of a variety of processes. In the present study, the complex features and functions of this group of proteins have been studied in the plant cotton, in soybean, in rapeseed and in corn. FAD genes encode enzymes that catalyze the desaturation of fatty acids, which affect the nutritional value and oxidative stability of seed storage oils (Clemente and Cahoon, 2009). Although the distribution of these omega-3 fatty acid desaturase genes is diverse, their genetic features and physicochemical properties tend to be identified. There are anatomical and physiological differences between the four species, which might be reflected in the diversity of omega-3 fatty acid

desaturase genes structure and conserved motifs. We identified that the 25 omega-3 fatty acid desaturase genes contain different numbers of amino acids and exons, indicating that there is some diversity in these four species. Protein length is shown to correlate with the probability that the protein is encoded by an essential gene. Increased protein length appears to be a notable mechanism by which the increasing complexity of protein–protein interaction networks is accommodated within the natural evolution of species (Tan et al., 2005). gi|728424384 and gi|408794 had the lowest and highest aliphatic indices among the protein sequences, respectively. This may be considered as positive factor for the increase in thermo stability of proteins (Ikai, 1980). MEME server identified the different conserved protein motifs that are present in each of the FAD proteins. Conserved motifs could be important for the diverse functions of omega-3 fatty acid desaturas proteins from cotton, soybean, rapeseed, and corn (Lynch and Conery, 2000). Some closely related members shared similar structures, implying functional similarities for these FAD proteins. The specific sequence motifs present in each clade may impart specific functions to the FAD proteins. The similarities in gene structure and motif composition of most FAD proteins consistented with phylogenetic analysis of the omega-3 fatty acid desaturase gene family. The differences in these characteristics among the different clades suggested that the FAD members were functionally diversified. Certainly, the characterization of FADs would provide candidate genes to improve new oilseed crops.

To explore how the omega-3 fatty acid desaturase gene family evolved, we performed a genome-wide comparison of plant FAD members from four species of various oilseed crops. Considerable phylogenetic analysis of FAD proteins has been conducted in cotton, soybean, rapeseed and corn, respectively. To obtain an overall picture of the 25 FAD proteins and their relationships with each other, a phylogenetic tree of FAD proteins was constructed, which divided the 25 FAD members into five clades. The plant FAD members from soybean appear to be more closely related to each other than to FAD genes of the other oilseed crops. Yurchenko et al in a genome wide analysis of the omega 3 tree fatty acid desaturase gene family in Gossypium indicated that species separated into three well-defined monophyletic groups in the phylogenetic analysis (Yurchenko et al., 2014). The result of connecting the thiol groups of two cysteine amino acids is formation of disulfides bond on the polypeptide chains. These bonds are responsible for stabilizing the globular structure and correct conformation of the protein (Bulleid and Ellgaard, 2011). In addition, extracellular proteins often have several disulfide bonds, whereas intracellular proteins usually lack them. Disulfide bonds' prediction can be beneficial in selecting recombinant expression system (Bulleid and Ellgaard, 2011). Glycosylation is one of the most current and structurally diverse forms of post-translational alteration of proteins. The bacterial host lacks post-translational modification systems such as glycosylation, but species from plants, fungi and animals have the ability to glycosylate proteins. Glycosylation may increase the stability of the protein, and in some cases, decrease the biological activity because of masking of the active site (Easton and Leader, 2011).

Conclusion

In the current study, 25 members of the FAD3 were analyzed, a comprehensive analysis including their chromosomal location, physical parameters, aliphatic index, instability index phylogeny, and conversed motifs. Phylogenetic analysis revealed that the genes could be grouped into five major clades. In each clade, the characteristics of exon/intron structure and motif compositions were relatively conserved. Five pairs of orthologous genes among the omega-3 fatty acid desaturase genes and a total of twenty distinct conserved motifs were identified. Four of these motifs were associated with encoding the fatty acid desaturase domain, while an additional four were related to the DUF3474 domain. The remaining motifs did not exhibit any functional annotation. Certainly, the identification of these FADs can help improve new varieties of oilseeds by introducing valuable candidate genes.

Supplementary Materials

There is no supplementary material available for this article.

Author contributions

Conceptualization, A.A.G.; software and formal analysis, A.A.G.; investigation, A.A.G and S.J.H.; data curation, A.A.G.; writing—original draft preparation, A.A.G. and S.J.H.; writing—review and editing, A.G.; All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Acknowledgments

The authors would like to express their gratitude to the Dr. Ebrahim Sharafi and Dr. Jafar Ahmadi for their valuable technical assistance.

Conflict of interest statement

The authors declare no conflict of interest.

References

Alloatti, A., and Uttaro, A.D. (2011). Highly specific methyl-end fatty-acid desaturases of trypanosomatids. *Mol Biochem Parasitol* 175(2): 126-132.

Bailey, T.L., and Elkan, C. (Year). "The value of prior knowledge in discovering motifs with MEME", in: *Proceedings of the Third International Conference on Intelligent Systems for Molecular Biology July, 1995 (Ismb)*), 21-29.

Bocianowski, J., Mikołajczyk, K., and Bartkowiak-Broda, I. (2012). Determination of fatty acid composition in seed oil of rapeseed (*Brassica napus* L.) by mutated alleles of the *FAD3* desaturase genes. *J Appl Genet* 53: 27-30.

Bulleid, N.J., and Ellgaard, L. (2011). Multiple ways to make disulfides. *Trends Biochem Sci* 36(9): 485-492.

Cheng, J., Saigo, H., and Baldi, P. (2006). Large‐scale prediction of disulphide bridges using kernel methods, two‐dimensional recursive neural networks, and weighted graph matching. *Proteins: Struct* 62(3): 617-629.

Chi, X., Yang, Q., Lu, Y., Wang, J., Zhang, Q., Pan, L., Chen, M., He, Y., and Yu, S. (2011). Genome-wide analysis of fatty acid desaturases in soybean (*Glycine max*). *Plant Mol Biol Rep* 29: 769-783.

Clemente, T.E., and Cahoon, E.B. (2009). Soybean oil: genetic approaches for modification of functionality and total content. *Plant Physiol* 151(3): 1030-1040.

Domínguez, T., Hernández, M.L., Pennycooke, J.C., Jiménez, P., Martínez-Rivas, J.M., Sanz, C., Stockinger, E.J., Sánchez-Serrano, J.J., and Sanmartín, M. (2010). Increasing *ω-3* desaturase expression in tomato results in altered aroma profile and enhanced resistance to cold stress. *Plant Physiol* 153(2): 655-665.

Easton, R., and Leader, T. (2011). Glycosylation of proteins–structure, function and analysis. *Life Sci Tech Bull* 60: 1-5.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S.e., Wilkins, M.R., Appel, R.D., and Bairoch, A. (2005). *Protein identification and analysis tools on the ExPASy server.* Humana press. Springer.

Gupta, R., and Brunak, S. (2001). "Prediction of glycosylation across the human proteome and the correlation to protein function," in *PubMed. Pacific Symp. Biocomput.*), 310-322.

Guschina, I.A., and Harwood, J.L. (2006). Mechanisms of temperature adaptation in poikilotherms. *FEBS Lett* 580(23): 5477-5483.

Hernández, M.L., Sicardo, M.D., and Martínez-Rivas, J.M. (2016). Differential contribution of endoplasmic reticulum and chloroplast *ω-3* fatty acid desaturase genes to the linolenic acid content of olive (*Olea europaea*) fruit. *Plant Cell Physiol* 57(1): 138-151.

Hu, R., Qi, G., Kong, Y., Kong, D., Gao, Q., and Zhou, G. (2010). Comprehensive analysis of *NAC* domain transcription factor gene family in *Populus trichocarpa*. *BMC Plant Biol* 10(1): 145.

Ikai, A. (1980). Thermostability and aliphatic index of globular proteins. *J Biol Chem* 88(6): 1895-1898.

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33(7): 1870-1874.

Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157(1): 105-132.

Liu, W., Li, W., He, Q., Daud, M.K., Chen, J., and Zhu, S. (2015). Characterization of 19 genes encoding membrane-bound fatty acid desaturases and their expression profiles in *Gossypium raimondii* under low temperature. *PloS One* 10(4): e0123281.

Lynch, M., and Conery, J.S. (2000). The evolutionary fate and consequences of duplicate genes. *Sci* 290(5494): 1151-1155.

Magnan, C.N., Randall, A., and Baldi, P. (2009). SOLpro: accurate sequence-based prediction of protein solubility. *Bioinformatics* 25(17): 2200-2207.

Marchler-Bauer, A., Anderson, J.B., Cherukuri, P.F., DeWeese-Scott, C., Geer, L.Y., Gwadz, M., He, S., Hurwitz, D.I., Jackson, J.D., and Ke, Z. (2005). CDD: a Conserved Domain Database for protein classification. *Nucleic Acids Res* 33(suppl_1): D192-D196.

Martin, B.A., Schoper, J.B., and Rinne, R.W. (1986). Changes in soybean (*Glycine max* [L.] Merr.) glycerolipids in response to water stress. *Plant Physiol* 81(3): 798-801.

Meesapyodsuk, D., and Qiu, X. (2012). The front-end desaturase: structure, function, evolution and biotechnological use. *Lipids* 47: 227-237.

Ohlrogge, J., and Browse, J. (1995). Lipid biosynthesis. *Plant Cell* 7(7): 957.

Román, Á., Andreu, V., Hernández, M.L., Lagunas, B., Picorel, R., Martínez-Rivas, J.M., and Alfonso, M. (2012). Contribution of the different omega-3 fatty acid desaturase genes to the cold response in soybean. *J Exp Bot* 63(13): 4973-4982.

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4(4): 406-425.

Shanklin, J., and Cahoon, E.B. (1998). Desaturation and related modifications of fatty acids. *Annu Rev Plant Biol* 49(1): 611-641.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28(10): 2731-2739.

Tan, T., Frenkel, D., Gupta, V., and Deem, M.W. (2005). Length, protein–protein interactions, and complexity. *Physica A* 350(1): 52-62.

Wang, J., Chitsaz, F., Derbyshire, M.K., Gonzales, N.R., Gwadz, M., Lu, S., Marchler, G.H., Song, J.S., Thanki, N., and Yamashita, R.A. (2023). The conserved domain database in 2023. *Nucleic Acids Res* 51(D1): D384-D388.

Williams, J.P., Khan, M.U., Mitchell, K., and Johnson, G. (1988). The effect of temperature on the level and biosynthesis of unsaturated fatty acids in diacylglycerols of *Brassica napus* leaves. *Plant Physiol* 87(4): 904-910.

Yang, Q., Fan, C., Guo, Z., Qin, J., Wu, J., Li, Q., Fu, T., and Zhou, Y. (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.

Yurchenko, O.P., Park, S., Ilut, D.C., Inmon, J.J., Millhollon, J.C., Liechty, Z., Page, J.T., Jenks, M.A., Chapman, K.D., and Udall, J.A. (2014). Genome-wide analysis of the omega-3 fatty acid desaturase gene family in *Gossypium*. *BMC Plant Biol* 14(1): 1-15.

Zhang, M., Barg, R., Yin, M., Gueta‐Dahan, Y., Leikin‐Frenkel, A., Salts, Y., Shabtai, S., and Ben‐Hayyim, G. (2005). Modulated fatty acid desaturation via overexpression of two distinct *ω‐3* desaturases differentially alters tolerance to various abiotic stresses in transgenic tobacco cells and plants. *Plant J* 44(3): 361-371.

JPMB
Journal of Plant Molecular Breeding
Published by Genetics and Agricultural Biotechnology Institute of Tabarestan
2023 | Volume 11 | Issue 1

مقاله پژوهشی

مقایسه توالی‌های پروتئینی پیش‌بینی شده از خانواده ژنی اسیدهای چرب غیراشباع امگا۳ در برخی از گیاهان دانه روغنی

عمار افخمی قادی*[۱]، سید جابر حسینی[۲]، علی قنبری[۲]

[۱] گروه مهندسی ژنتیک و اصلاح نباتات، دانشکده کشاورزی، دانشگاه امام خمینی (ره)، قزوین، ایران
[۲] گروه زراعت، دانشگاه تربیت مدرس، تهران، ایران
[۳] گروه ژنتیک و مهندسی تولیدات گیاهی، دانشگاه زنجان، زنجان، ایران

چکیده: هدف از این مطالعه، تجزیه مقایسه‌ای بین توالی‌های پروتئینی پیش‌بینی‌شده در خانواده ژنی اسید چرب غیراشباع امگا۳ (FAD) در گیاهان مختلف دانه روغنی، همچون پنبه، سویا، کلزا و ذرت بود. بیست و پنج ژن اسید چرب غیراشباع امگا۳ با حذف توالی‌های مشابه، متمایز شدند. با تجزیه و تحلیل فیلوژنتیک، خانواده ژنی اسید چرب غیراشباع امگا۳ شناسایی شد. اطلاعات ژنی اسید چرب غیراشباع امگا۳ با استخراج ویژگی‌های فیزیکی شامل تعداد اسیدهای آمینه، مکان کروموزومی، تعداد اگزون و غیره به دست آمد. نتایج نشان داد میانگین طول توالی‌های پروتئینی کدگذاری شده اسید چرب غیراشباع امگا۳، ۳۸۸/۵۸ اسید آمینه بود. وزن مولکولی پروتئین‌های بررسی‌شده از ۲۲/۲ تا ۵۱/۳ کیلو دالتون متغیر بود. درخت فیلوژنتیکی، پروتئین‌های اسید چرب امگا۳ را در پنج گروه قرار داد. گروه دوم، کمترین تعداد اعضای ژن اسید چرب غیراشباع امگا۳ را شامل شده، در حالی که گروه اول بیشترین تعداد را در بر می‌گیرد. بنابراین پنج جفت ژن ارتولوگ در میان ژن‌های اسید چرب غیراشباع امگا۳ و در مجموع بیست موتیف حفاظت‌شده مجزا شناسایی شد. چهار مورد از این موتیف‌ها با رمزگذاری دمین اسید چرب غیراشباع مرتبط بوده، در حالی که چهار مورد دیگر مربوط به دمین DUF3474 بودند. بدون شک، مشخص کردن خصوصیات این اسیدهای چرب غیراشباع، می‌تواند ژن‌های گزینشی ارزشمندی را برای بهبود واریته‌های جدید دانه روغنی آشکار سازد.

کلمات کلیدی: اسید چرب امگا۳، خانواده ژنی، گیاهان دانه روغنی، موتیف‌های حفاظت‌شده.

ویراستار علمی
دکتر پرویز حیدری،
دانشگاه صنعتی شاهرود، ایران

تاریخ
دریافت: ۱۰ مرداد ۱۴۰۲
پذیرش: ۲۷ مهر ۱۴۰۲
چاپ: ۸ بهمن ۱۴۰۲

نویسنده مسئول
دکتر عمار افخمی قادی
a.afkhami@sanru.ac.ir

ارجاع به این مقاله
Afkhami Ghadi, A., Hosseini, S. J., and Ghanbari, A. (2023). Comparison of protein predicted sequences of the omega-3 fatty acid desaturase gene family in some of the oil seed crops. *J. Plant Mol. Breed 11* (1): 28-40. doi: 10.22058/JPMB.2023.2008246.1281.