Comparative identification and evolutionary relationship of fatty acid desaturase (FAD) genes in some oil crops: the sunflower model for evaluation of gene expression pattern under drought stress

Yasemin Celik Altunoglu, Necdet Mehmet Unel, Mehmet Cengiz Baloglu, Ferhat Ulu, Tevfik Hasan Can and Rahmi Cetinkaya

Department of Genetics and Bioengineering, Faculty of Engineering and Architecture, Kastamonu University, Kastamonu, Turkey

ABSTRACT
Sunflower (Helianthus annuus L.) and some oil crops have considerable economic value, making them important for commercial use. The fatty acid content of their seeds is crucial for proper nutrition. Additionally, the desaturation reaction via fatty acid desaturases (FADs) leads to formation of double bonds in fatty acids and is an essential step in fatty acid biosynthesis. This study aimed to identify and compare some fatty acid desaturase gene members in sunflower and other oil crops. Totally, 33 FAD genes from sunflower, 19 from sesame, 20 from canola, 25 from cacao and 46 from flax were analysed. The FAD members had roles in oxidation–reduction processes, fatty acid biosynthetic processes or lipid metabolic processes. Comparative phylogenetic and motif analysis revealed a conserved structure among FAD members belonging to various oil crops. The average Ka/Ks rates calculated on the basis of orthologous gene analysis were 0.04, 0.05 and 0.05 between sunflower and, respectively, canola, cacao and sesame. The most recent divergence time between sunflower and cacao was estimated at an average of 100 million years ago (MYA). This was followed by flax, sesame and canola with an average of 101 MYA, 114 MYA and 149 MYA, respectively. Alpha-helices were dominant in the predicted 3D structures of FAD proteins. The FAB2 expression levels from a drought tolerant sunflower variety were not affected by drought. The characterisation of desaturase family members in economically valuable oil crops could be useful for functional cloning studies to enhance the unsaturated fatty acid contents of the plants.

Introduction
Sunflower (Helianthus annuus L.) is an important crop and has considerable economical value worldwide because of the seed oil content. The cultivation of sunflower on 25 million ha leads to production of approximately 35 million tons worldwide, particularly in Eastern Europe, former Soviet Union and Argentina [1,2]. In addition, sunflower can grow in some parts of the Mediterranean region and North Africa owing to its tolerance to water deficit or salt stresses [3,4]. Sunflower seeds have high concentrations of polyunsaturated fatty acids, especially oleic (18:1) and linoleic (18:2) acids [5–7]. Oil with high total unsaturated fatty acid level is healthier for human consumption and helps in regulating blood cholesterol ratios by reducing low density lipoprotein levels [8,9]. This leads to significant reduction in risk ratios for coronary heart diseases [10,11].

Fatty acid contents of seeds contribute to nutrition and have commercial value [12]. The desaturation reactions by fatty acid desaturases (FADs) are necessary for fatty acid biosynthesis because they lead to the formation of double bonds [13]. FAB2, FAD4, FAD5, FAD6, FAD7 and FAD8 are groups of enzymes localised in the plastids and are responsible for the desaturation reactions in lipids. FAD2 and FAD3 are present in endoplasmic reticulum (ER) and direct the desaturation of extra-chloroplastic lipids [14]. Moreover, stearoyl-ACP desaturase (FAB2 or SAD), a member of soluble FADs, has a key role in fatty acid synthesis among distinct subfamilies of these plant desaturases [15,16]. FAB2 adds a double bond to saturated stearic acids (18:0) to transform them to unsaturated oleic acids (18:1). These oleic acids in turn act as precursors for fatty acid biosynthesis in plastids [17,18]. Therefore, the FAB2 group of enzymes attract the attention of researchers because they help in determining the proportions of saturated to unsaturated fatty acids [19]. Once desaturation is ensued, these fatty acids are further subjected to desaturation reactions. This would help in constructing polyunsaturated fatty acids.
e.g. linoleic acid formation (18:2) by FAD2 in ER and by FAD6 in plastids. FAD3, FAD7 and FAD8 help in the formation of linolenic acid (18:3). In addition, FAD4 and FAD5 convert palmitic acid (16:0) to palmitoleic acid (16:1) [14,20–24].

Abiotic stress factors cause variable hard effects on plant growth and development. In particular, the water limitation factor might be responsible for the salinity of many arable fields by 2050 [25]. Drought is responsible for the increase in soluble sugar and proline levels in sunflower leaves [26]. Besides, water deficiency directly affected the seed germination as well as the development of seedlings in sunflower [27].

This study aimed to identify and compare fatty acid desaturase gene members in sunflower and other oil crops including sesame (Sesamum indicum), cacao (Theobroma cacao), flax (Linum usitatissimum) and canola (Brassica rapa) by utilising their whole genome sequence data. In addition, some bioinformatics tools performed molecular characterisation analyses of FAD proteins. This included gene duplication and divergence times of FAD genes, phylogenetic relationships of FAD members in oil crops (sunflower, sesame, cacao, flax and canola), FAD gene ontology, amino acid motif analysis and prediction of three-dimensional structures. Besides, the gene expression profiles of some FAB2 members were also evaluated under drought stress conditions in leaf and root tissues of drought tolerant as well as drought sensitive sunflower varieties.

Materials and methods

Sequence data retrieval for desaturase genes of sunflower and other oil crops

All protein, genomic and coding (CDS) sequences were downloaded from Sunflower Genome Database (https://www.sunflowergenome.org/). These sequences were searched for desaturase genes by using gene annotation numbers in the same database. This was followed by extraction of desaturase sequences. Moreover, extracted protein sequences were also confirmed for their annotations in Pfam database (http://pfam.sanger.ac.uk/) by checking Hidden Markov Model (HMM) profiles. Sesame sequences were retrieved from Sesamum indicum Genome Database. The same approaches were utilised for sunflower to explore desaturase genes. Phytozome database were used for sequence retrieval of desaturase genes from cacao, flax and canola plant genomes. Briefly, confirmed desaturase protein sequences from sunflower were exploited as a reference for identification of desaturase genes from cacao, flax and canola plants in Phytozome v.11 database [28]. In addition, defined Pfam accession numbers for desaturase domains were also utilised to search in Phytozome v.11 database for these plants. Thus the identified desaturase sequences were retrieved and utilised in the present study.

Phylogenetic analysis and determination of conserved motifs

Desaturase genes from sunflower, sesame, cacao, flax and canola were aligned by Clustal W using the MEGA7 software [29]. Neighbour joining method with bootstrap analysis for 1000 replicates was utilised for phylogenetic tree construction with the above aligned sequences. In addition, Jones Taylor Thornton substitution model was selected for phylogenetic tree construction. Motif domain search tool (MEME) (http://meme.nbcr.net/meme3/meme.html) [30] was used to detect amino acid sequence motifs of desaturases. Default parameters were applied for MEME searches. The determined MEME motifs were then screened in InterPro database with InterProScan. This helped in the functional analysis of proteins by classifying them into families thereby predicting domains and important sites [31].

Calculation of synonymous and non-synonymous substitution ratios

The Clustal W tool was used for aligning orthologous gene pairs between sunflower and sesame, cacao, flax and canola. Synonymous (Ks) and non-synonymous (Ka) change ratios were determined using the CODEML program (http://www.bork.embl.de/pal2nal/) that enables aligning amino acid sequences of FAD genes and their complementary DNAs [32]. Duplication and divergence times (million years ago, MYA) of every FAD gene in this study were calculated using the following formula: $T = Ks/2\lambda (t = 6.5 \times 10^{-9})$ [33,34].

Predicted biological role, molecular function and cellular localisation of FAD proteins

The Blast2GO (http://www.blast2go.com) program was used for functional analysis of FAD genes [35]. Three steps (i- BLASTp against the non-redundant protein database of NCBI, ii- mapping and retrieval of GO terms related with the BLAST results, and iii- annotation of GO terms associated with each query to relate the sequences to known protein function) were followed using amino acid sequences of FAD genes. Predicted biological processes, cellular components and molecular functions, which were of GO classification categories, were clarified.
**Predicted 3D structures of FAD proteins**

BLASTP search was used for the identification of similar sequences. The three-dimensional structure of the best known sample for FAD protein sequences of sunflower, sesame, cacao, flax and canola was from Protein Data Bank (PDB) [36]. This information was used in Phyre2 database (Protein Homology/Analogy Recognition Engine; http://www.sbg.bio.ic.ac.uk/phyre2) to predict the 3D structures of FAD proteins with > 90% confidence level [37].

**Plant materials, growth parameters and stress treatment**

Drought tolerant sunflower (Saray variety) seeds were obtained from the Directorate of Trakya Agricultural Research Institute (Edirme, Turkey), and drought sensitive sunflower plants (Turay variety) were obtained from the Directorate of Ege Agricultural Research Institute (Izmir, Turkey). The seeds were washed thrice without capsules, transferred to pots and were grown in a plant growing cabinet. The growth conditions provided were 400 μmol m⁻² s⁻¹ light intensity, 24 °C± 2 °C, 16 h light and 8 h dark photoperiods. Hoagland solution [38] was used as a culture medium for growing plants for 14 days. Drought stress treatment was achieved with 10% Polyethylene glycol 6000 (PEG-6000) by addition into Hoagland medium at the end of the 14th day on the basis of leaf development. Leaf as well as root samples were collected at 0, 3, 12 and 24 h after stress treatment. Zero hour samples were utilised as the control. Control and stress treated plants were grown in the same conditions in the growing cabinet and three biological samplings were made for every treatment. The obtained tissue samples were frozen in liquid nitrogen, and then subjected to RNA isolation to study the expression patterns of some FAD genes in sunflower.

**Total RNA isolation and quantitative real-time PCR (qRT-PCR) analysis**

Trizol reagent (Life Technologies Corporation, Grand Island, NY, USA) was used for total RNA isolation from leaf and root tissues according to the manufacturer’s instructions with small changes. Isolated RNA was controlled for quality as well as quantity by agarose gel electrophoresis and Multi-scan GO nano-spectrophotometer (ThermoFisher Scientific, USA). Further, DNase I (Thermo, Lithuania) was used for the removal of genomic DNA fragments in RNA samples. The Primer 5 software program (http://www.primer-e.com/index.htm) was employed for designing primer for sunflower FAD genes. 18SrRNA gene was selected as a reference gene in the present study [39]. List of primers used in quantitative real-time (qRT)-PCR expression analysis of HaFAB2 and 18SrRNA genes are shown in Table 1. Three biological replicates were used per hour and triple quantitative PCR was applied for duplications. Qiagen SYBR Green master mix (Qiagen, Hilden, Germany) was used for qRT-PCR analysis in RotorGene qPCR detection system (Qiagen, Hilden, Germany). qRT-PCR conditions were 45 cycles of 95 °C for 5 s and 60 °C for 10 s. After the polymerisation reactions, melting curve analysis was performed. The ΔCT and ΔΔCT were calculated according to the formulas: ΔCT = CT target – CT reference and ΔΔCT = ΔCT treated sample – ΔCT untreated sample (0 h). Differences of expression patterns were assessed as 2⁻ΔΔCT [40].

**Results and discussion**

**Identification of FAD genes in the sunflower and other oil crops**

The defined sunflower FAD genes on the basis of annotation numbers and HMM profiles from Sunflower Genome Database were used for determination of sesame, cacao, flax and canola FAD genes. The study identified FAD gene numbers in all organisms after multiple searches and is shown in Table 2. FAD genes were grouped and numbered their chromosomal locations on sunflower, sesame and canola genomes according to Pfam searches. However, there was no chromosomal location for FAD genes on cacao and flax genomes. Further, FAD4

| Name       | Forward primers (5'-3') | Reverse primers (5'-3') |
|------------|------------------------|------------------------|
| HaFAB2.1   | CACTCACACCCCAACACAAA   | TCCAAACCGGTACGAGATA    |
| HaFAB2.2   | ATGCAAGTCCGATATGGAAG   | TACTCCGACCTCTCTGCC     |
| HaFAB2.4   | CTCCATGGCCGACACAAAGA   | TCCCTGAAACCTGTGAGTGC   |
| 18SrRNA    | TTTTAGGCTACCTGACTACGATA | CTCTCCGGCATCGACCTAAATCTCC |

Table 1. List of primers used in quantitative real-time-PCR expression analysis of HaFAB2 and 18SrRNA genes.

| Oil crops | FAB2 | FAD2 | FAD3 | FAD5 | FAD6 | FAD7 | FAD8 | Total FAD gene numbers |
|-----------|------|------|------|------|------|------|------|------------------------|
| Sunflower | 4    | 11   | 16   | 2    | –    | –    | 33   |                        |
| Sesame    | 7    | 2    | 1    | 1    | 4    | –    | 19   |                        |
| Canola    | 7    | 2    | 3    | 5    | 2    | 1    | 20   |                        |
| Cacao     | 8**  | 4    | 1    | 3    | 1*** | 8*** | 25   |                        |
| Flax      | 4*   | 17*  | 3*   | 11   | 8    | –    | 46   |                        |
| Total     | 30   | 36   | 7    | 29   | 22   | 3    | 163  |                        |

Note: Boldface numbers and asterisks (*) indicate the reference of previous studies [45–47].
proteins were not observed in any of studied organisms. Sunflower had FAB2, FAD2, FAD5 and FAD6 proteins. FAD5 and FAD3 proteins were absent in canola and cacao, respectively. Same protein types were detected in flax and sesame. The highest number of FAD protein type was of FAD2 proteins with 17 members in flax. This was followed by FAD5 and FAD2 proteins in sunflower and FAD5 proteins in flax.

According to earlier studies, a total of 4 FAB2, 15 FAD2 and 6 FAD3 proteins were identified for flax [41–45]. In our study, we identified FAD5, FAD6 and FAD8 members of desaturase family in flax. In another study, 8 FAB2 genes and some FAD3, FAD7/FAD8 members were determined in cacao [46,47]. Four FAD2, a FAD5, three FAD6 and seven FAD7/FAD8 proteins were added to the above desaturase family in cacao as observed in our study. Two cDNA clones encoding FAB2 genes were determined in sesame before [48]. In a comprehensive study, three FAB2, two FAD2, fourteen FAD5 genes, and one gene for each FAD3, FAD4, FAD6 and FAD7 were characterised in the cucumber genome [23]. In addition, 9 SAD (FAB2) genes in diploid Gossypium raimondii (D5) and G. arboreum (A2) each, 18 SAD (FAB2) genes for tetraploid G. hirsutum acc. TM-1 (AD1) and 19 SAD (FAB2) genes for G. barbadense cv. Xinhai21 (AD2) were determined [49]. Sunflower, sesame and canola FAD genes were identified for the first time in the present study. FAD genes were located especially on Chromosome 3, 4 and 6 of the sesame genome. Moreover, most of the FAD genes were on Chromosome 3 and 5 of the canola genome. However, sunflower FAD genes were distributed on different sunflower chromosomes. The predicted amino acid lengths of the FAD genes of these plants ranged between 72 aa (HaFAB2.2) and 1171 aa (SinFAB2.4). The predicted molecular weights of the FAD proteins were between 8.15 and 134 kDa. All FAB2 proteins from sunflower, sesame, cacao, flax and canola were acidic, whereas, FAD3, FAD5, FAD6, FAD7 and FAD8 proteins from the above plants were basic. The results are consistent with the result of an earlier study wherein cucumber FAD characterisation of all FAB2 proteins was of acidic character, while the characterisation of other FAD protein groups indicated basic character [23]. This could be explained by the fact that the FAB2 group of proteins is the only group of soluble proteins among other FAD protein groups. All detailed information related to FAD proteins for all oil crops is shown in Supplementary Table S1.

Analyses of predicted biological role, molecular function and cellular localisation of FAD proteins

The FAD proteins have roles in metabolic, cellular and single organism processes according to gene ontology analysis of sunflower FAD genes. In detail, these activities are oxidation–reduction processes, fatty acid biosynthetic processes or lipid metabolism processes. These biological roles appeared to be consistent with their key roles in fatty acid synthesis [17,18]. The dominant molecular function of sunflower FAD proteins was catalytic activity. In addition, they were especially located on membranes or part of a membrane (Figure 1). However, the detailed analysis revealed that FAB2 group members in sunflower were localised in the chloroplast, which was consistent with FAB2 proteins in cucumber, cotton and Arabidopsis [23,49,50]. Furthermore, all FAD5 protein members of sunflower were integral components of the membrane (Supplementary Table S2).

In addition, all FAD proteins from other oil crops were subjected to Blast2GO analysis together (Supplementary Figure 1). Single organism and metabolic processes were mostly found as their biological role. According to detailed analysis, all FAD members were observed to have roles in oxidation and reduction processes. Besides, fatty acid biosynthetic and metabolic processes and lipid metabolic processes were also reported to have biological roles in the studied plants (Supplementary Table S3). These results are in agreement with recent studies indicating the usage of phospholipids as substrate of NADH, NADH-cytochrome b5 reductase and cytochrome b5 as the electron donor by FAD2 and FAD3 enzymes, respectively. Moreover, FAD6 and FAD7/8 members utilise glycolipids as acyl carriers whereas, NAD(P)H, ferredoxin-NAD(P) reductase and ferredoxin are used as an electron donor system [51,52]. Catalytic activity is the prime molecular function among FAD proteins, which might reflect their mode of action. According to previous studies, soluble as well as membrane-bound Δ9 desaturase has been reported to associate with iron for catalytic activity [53,54]. Moreover, FADs were found predominantly on membranes or membrane parts including chloroplast and endoplasmic reticulum membranes. However, the FAB2 members of the studied oil crops were located especially in chloroplasts, which was consistent with their already reported desaturation role in the chloroplast. Further, the FAD2 and FAD3 members of the studied plants were especially integral components of the membrane or in the ER. FAD6, FAD7 and FAD8 proteins were also integral membrane components and a few of them were located in the chloroplast envelope (Supplementary Table S3). These findings were relevant with the roles of these membrane-bound enzymes in the desaturation of fatty acids in different parts of the cell.

Phylogenetic relationships and conserved motifs of FAD proteins

The phylogenetic relationships among the FAD proteins were analysed by utilising the Neighbour Joining
Method with bootstrap analysis of 1000 iterations. According to the phylogenetic tree, seven different clusters (Cluster I–VII) were observed (Figure 2). The FAD proteins in these oil plants were divided into different clusters according to their desaturase group. Cluster I, II and III were composed of FAD2 proteins as observed from all the oil plants included in this study. Cluster VI was comprised of FAD3, FAD6, FAD7 and FAD8 proteins. Further, all FAB2 and FAD5 proteins were present in Cluster V and Cluster VII, respectively. The FAD6 protein group members were mainly found in Cluster VI. The FAD7 and FAD8 members were in the same cluster due to high sequence homology. In addition, FAD3, FAD7 and FAD8 enzymes are known as $\omega$–3 or $\Delta$–15 desaturases which add double bonds to generate 18:3 polyunsaturated fatty acids [14]. According to these results, the same FAD group of enzymes were grouped in the same cluster due to high sequence homology. In addition, FAD3, FAD7 and FAD8 enzymes are known as $\omega$–3 or $\Delta$–15 desaturases which add double bonds to generate 18:3 polyunsaturated fatty acids [14]. According to these results, the same FAD group of enzymes were grouped in the same cluster due to high sequence homology. In addition, FAD3, FAD7 and FAD8 enzymes are known as $\omega$–3 or $\Delta$–15 desaturases which add double bonds to generate 18:3 polyunsaturated fatty acids [14]. According to these results, the same FAD group of enzymes were grouped in the same cluster due to high sequence homology. In addition, FAD3, FAD7 and FAD8 enzymes are known as $\omega$–3 or $\Delta$–15 desaturases which add double bonds to generate 18:3 polyunsaturated fatty acids [14]. According to these results, the same FAD group of enzymes were grouped in the same cluster due to high sequence homology. In addition, FAD3, FAD7 and FAD8 enzymes are known as $\omega$–3 or $\Delta$–15 desaturases which add double bonds to generate 18:3 polyunsaturated fatty acids [14]. According to these results, the same FAD group of enzymes were grouped in the same cluster due to high sequence homology. In addition, FAD3, FAD7 and FAD8 enzymes are known as $\omega$–3 or $\Delta$–15 desaturases which add double bonds to generate 18:3 polyunsaturated fatty acids [14]. According to these results, the same FAD group of enzymes were grouped in the same cluster due to high sequence homology. In addition, FAD3, FAD7 and FAD8 enzymes are known as $\omega$–3 or $\Delta$–15 desaturases which add double bonds to generate 18:3 polyunsaturated fatty acids [14]. According to these results, the same FAD group of enzymes were grouped in the same cluster due to high sequence homology. In addition, FAD3, FAD7 and FAD8 enzymes are known as $\omega$–3 or $\Delta$–15 desaturases which add double bonds to generate 18:3 polyunsaturated fatty acids [14]. According to these results, the same FAD group of enzymes were grouped in the same cluster due to high sequence homology. In addition, FAD3, FAD7 and FAD8 enzymes are known as $\omega$–3 or $\Delta$–15 desaturases which add double bonds to generate 18:3 polyunsaturated fatty acids [14]. According to these results, the same FAD group of enzymes were grouped in the same cluster due to high sequence homology. In addition, FAD3, FAD7 and FAD8 enzymes are known as $\omega$–3 or $\Delta$–15 desaturases which add double bonds to generate 18:3 polyunsaturated fatty acids [14]. According to these results, the same FAD group of enzymes were grouped in the same cluster due to high sequence homology. In addition, FAD3, FAD7 and FAD8 enzymes are known as $\omega$–3 or $\Delta$–15 desaturases which add double bonds to generate 18:3 polyunsaturated fatty acids [14]. According to these results, the same FAD group of enzymes were grouped in the same cluster due to high sequence homology. In addition, FAD3, FAD7 and FAD8 enzymes are known as $\omega$–3 or $\Delta$–15 desaturases which add double bonds to generate 18:3 polyunsaturated fatty acids [14]. According to these results, the same FAD group of enzymes were grouped in the same cluster due to high sequence homology. In addition, FAD3, FAD7 and FAD8 enzymes are known as $\omega$–3 or $\Delta$–15 desaturases which add double bonds to generate 18:3 polyunsaturated fatty acids [14]. According to these results, the same FAD group of enzymes were grouped in the same cluster due to high sequence homology. In addition, FAD3, FAD7 and FAD8 enzymes are known as $\omega$–3 or $\Delta$–15 desaturases which add double bonds to generate 18:3 polyunsaturated fatty acids [14]. According to these results, the same FAD group of enzymes were grouped in the same cluster due to high sequence homology. In addition, FAD3, FAD7 and FAD8 enzymes are known as $\omega$–3 or $\Delta$–15 desaturases which add double bonds to generate 18:3 polyunsaturated fatty acids [14]. According to these results, the same FAD group of enzymes were grouped in the same cluster due to high sequence homology. In addition, FAD3, FAD7 and FAD8 enzymes are known as $\omega$–3 or $\Delta$–15 desaturases which add double bonds to generate 18:3 polyunsaturated fatty acids [14]. According to these results, the same FAD group of enzymes were grouped in the same cluster due to high sequence homology. In addition, FAD3, FAD7 and FAD8 enzymes are known as $\omega$–3 or $\Delta$–15 desaturases which add double bonds to generate 18:3 polyunsaturated fatty acids [14].
desaturases and soluble desaturases in higher plants [52]. In addition, the similar motif content of the FAD proteins from different oil plants might be explained by the evolutionarily conserved role of FADs.

**Analysis of synonymous and non-synonymous substitution ratios**

Since gene duplications have been reported to be the leading factor behind the distribution of the gene families, the predicted duplication times of the FAD genes were estimated [55]. The nonsynonymous (Ka) versus synonymous (Ks) substitution (Ka/Ks) rates were computed between orthologous genes of sunflower FAD genes with those of flax, cacao, sesame and canola, in order to better understand the divergence of FAD genes. The average Ka/Ks rates were 0.04, 0.05 and 0.05 between sunflower and canola, cacao and sesame, respectively. The observed average Ka/Ks values were <1, suggesting that the FAD genes in these oil plants

---

**Figure 2.** Phylogenetic relationships of FAD proteins from all oil crops.  
Note: Alignment of sequences was achieved by ClustalW. Phylogenetic tree was constructed with bootstrap analysis for 1000 repetitions using the MEGA v7 software.
were under a purifying selection. Further, the most recent divergence time was observed between sunflower and cacao with the average of 100 MYA and this was followed by flax, sesame and canola with the average of 101 MYA, 114 MYA and 149 MYA, respectively. The highest number (355 pairs) of orthologous genes were observed between sunflower and flax (Figure 3, Supplementary Table S4).

In addition, the FAD gene family members in all oil crops were compared for evaluation of orthologous relationships (Figure 4, Supplementary Table S4. According to the analysis of the FAB2 genes, the number of orthologous FAB2 genes (57 pairs) between sesame and cacao was the highest. Canola and flax FAB2 genes were the
genes to diverge at the earliest point (220–240 MYA) among FAB2 genes from all organisms. When considering the FAD2 genes, 146 pairs of orthologous genes were determined between sunflower and flax. The earliest divergence time was observed between sesame and flax (140–160 MYA). For FAD3 genes, the highest number of orthologous gene pairs (57 pairs) were observed between sesame and cacao. In addition, the most recent divergence time was that between sesame and cacao (80–90 MYA). Considering the FAD5 genes, the earliest point of divergence was that between sunflower and flax (50–70 MYA). However, among all the FAD genes in oil crops, the FAD5 genes from sunflower and flax indicated the highest gene pair rate. A total of 150 FAD5 orthologous genes were observed between sunflower and flax plants. For FAD6 genes, the earliest and the last divergence time rates were calculated between cacao and canola, canola and flax (260–280 MYA) and sesame and flax (25–30 MYA). The highest number of orthologous FAD6 genes was revealed between canola and flax (18 pair of genes). For FAD7 genes, orthologous genes were identified only between cacao and canola (2 pairs of genes). The divergence time of these FAD7 genes ranged between 15 and 25 MYA. According to the analysis of FAD8 genes, the earliest divergence time for this gene family was observed between canola and flax (290–310 MYA). With this exception, the FAD8 gene family members showed the most recent divergence time rate (average between 10–20 MYA) between sesame and cacao, sesame and canola, cacao and canola. To our knowledge, this study is the first to show orthologous relationships of FAD gene families and their divergence time rates in oil crops. Further, flax was observed to be the key plant for the evolution of FAD genes. This means that the FAD genes in flax indicated a higher orthologous relationship in comparison to other oil crops, including sunflower, sesame, canola and cacao. The highest number of gene orthologous for FAD2 and FAD5 were observed between sunflower and flax, with 146 gene pairs and 150 gene pairs, respectively. On the other hand, the lowest number of orthologous genes for FAD6, FAD7 and FAD8 were identified between sunflower and cacao (2 pairs), cacao and canola (2 pairs) and sesame and canola (2 pairs), respectively. Moreover, for FAB2 and FAD3 genes, the highest number of orthologous genes were defined between sesame and cacao plants (57 gene pairs). This might be explained by the genetic closeness of these plants on the basis of orthologous desaturase genes. In addition, the FAB2 gene family members were identified in all oil crop plants. Therefore, all combinations of sunflower, sesame, canola, cacao and flax for FAB2 genes were compared with each other. The calculation of their divergence time rates and identification of orthologous gene pairs were performed using all oil crops.

**Analysis of predicted 3D structures of FAD proteins**

A total of 27 FAD proteins from sunflower, 10 proteins from flax, seven proteins from cacao and four proteins from sesame were selected on the basis of their high homology as observed by BLASTP search against PDB. In Phyre2 database, HMM scanning was applied by detection rate for identification of a homology model [56]. These FAD proteins were distributed with high homology (Supplementary Table S5). The percentage of modelled residues ranged between 11% and 99% at the >90% confidence interval when intensive mode was selected. The FAB2, FAD2 and FAD5 groups displayed high homology according to the predicted 3D structure analysis of FAD proteins. Further, only alpha-helix structure was observed in the above selected proteins, but no β-sheets (Figure 5). This structure was consistent with the study that analysed the crystal structure of Δ9 stearoyl-acyl carrier protein desaturase (FAB2) from castor seed by protein crystallographic methods [54]. It reported that 363 amino acid monomers contained a single domain of 11 a-helices. Moreover, nine of these helices constituted an antiparallel helix bundle [54].

**Drought stress effects on the expression profiles of sunflower FAB2 genes in sunflower varieties**

We selected sunflower as the model organism for gene expression studies because of its use in identifying desaturase genes from other oil crops and its economic importance in our country. The expression patterns of FAB2 genes were also extensively studied in other plants. However, few studies on the effect of drought stress on the expression levels of FAB2 genes have been conducted. For better evaluation of FAD gene expression studies, two sunflower varieties which are drought tolerant and sensitive were utilised. qRT-PCR was performed to evaluate the effects of drought on the expression levels of sunflower HaFAB2.1, HaFAB2.2 and HaFAB2.4 genes in the root and leaf tissues of Turay and Saray sunflower varieties. Further, the expression levels of selected genes involved in fatty acid production were downregulated by drought stress in the leaf and root tissues of Turay, the drought sensitive variety. The HaFAB2.1 gene was only upregulated in the leaf tissues of the Turay variety. In the Saray variety, which is drought tolerant, the expression of the studied FAB2 genes was not affected by drought stress either in leaf or in root tissues. The expression of the HaFAB2.2 genes was downregulated after 3 h of stress treatment in the leaf tissues, which
might be explained as the first response to drought stress. However, the expression of the above gene was highest after 24 h of stress application in comparison to other genes in leaf tissues. In addition, the expression levels of the HaFAB2.1 and HaFAB2.4 genes were not influenced by drought. Their expression levels increased from the first hours of stress when compared to the control in leaf tissues. In the root tissues of the Saray variety, the expression levels of the HaFAB2.1 and HaFAB2.2 genes were increased at 3 and 12 h of stress treatments, respectively. The expression of HaFAB2.4 was also upregulated at 24 h of stress application in root tissues (Figure 6).

Previous studies on the expression patterns of FAB2 genes in oil crops revealed that some FAB2 genes were mainly upregulated in embryo and seed tissues. For example, the OeSAD2 gene in olive appeared to be the main gene contributing to the oleic acid content of the olive fruit as per lipid and gene expression levels [57]. In maize, the expression level of the ZmSAD1 gene was the highest in the embryo, followed by that in the endosperm, leaf, seed, shoot, tassel, ear, ovule, anther and shoot apical meristem tissues [58]. The expression pattern of the TcSAD1 gene in cacao seeds correlated with the changes in fatty acid composition during seed maturation, while this gene was universally expressed across all the studied tissues [46]. According to Northern blot analysis, two clones of sesame FAB2 gene transcripts (CDES01 and CDES04) indicated an accumulation in seeds with a peak at 21 days after anthesis [48]. In our sunflower varieties, expression of the HaFAB2.1, HaFAB2.2 and HaFAB2.4 genes was detected in the leaf and root tissues. These results appear to be consistent with the suggestion that three cucumber CsFAB2 genes were dominantly expressed in the cucumber seedling leaves [23]. Additionally, the gene expression levels of the FAB2 genes increased after the first shock of drought stress in the drought tolerant Saray variety. In contrast to this, the expression levels of the genes related to fatty acid production were observed to be reduced by drought stress in the leaf and root tissues of the drought sensitive Turay variety. These results supported the results from earlier studies. For example, there was an approximately 10-fold increase in two cucumber CsFAB2 isoforms under cold stress conditions with distinct inducible time courses, whereas 3-fold downregulation was observed for another CsFAB2 gene expression under cold temperatures [23]. It could be concluded that some FAB2 gene members were not affected by stress and might be essential for unsaturated fatty acid production in these plants.

The present study is one of the limited studies evaluating the drought stress responses of the FAB2 group of
fatty acid desaturase genes and gives an idea about the FAB2 gene expression patterns under drought conditions. The results demonstrated that water deficiency affects the unsaturated fatty acid biosynthesis. In addition, our study characterised some desaturase genes and their proteins, which are important in the synthesis of unsaturated fatty acids. The obtained results could be used in future studies to increase the desaturase gene expression under drought conditions in oil crops.

Conclusions

This study reported detailed and comparative analyses of FADS in sunflower and some oil crops. FADS have key roles in unsaturated fatty acid biosynthesis. Comparative phylogenetic and motif analysis of FAD proteins belonging to different oil crops reflected a conserved structure between FAD protein groups. In addition, the expression patterns of FAB2 gene members in different sunflower varieties might be informative for the activity of these desaturases under drought stress conditions. This detailed characterisation analysis of desaturase family members in these economically valuable oil crops could be useful for functional cloning and validation studies to increase the proportion of unsaturated fatty acid content in the plants.

Disclosure statement

Yasemin CELIK ALTUNOGLU, Necdet Mehmet UNEL, Mehmet Cengiz BALOGLU, Ferhat ULU, Tevfik Hasan CAN, Rahmi CETINKAYA declare that they have no conflict of interest.

Funding

This study was supported by The Scientific and Technological Research Council of Turkey [grant number TUBITAK 2209-A Program].

ORCID

Mehmet Cengiz Baloglu http://orcid.org/0000-0003-2976-7224

References

[1] Liu X, Baird W. Differential expression of genes regulated in response to drought or salinity stress in sunflower. Crop Sci. 2003;43(2):678–687.
[2] Seiler GJ, Rieseberg LH. Systematics, origin, and germplasm resources of the wild and domesticated sunflower. In: Schneiter AA, editor. Sunflower technology and production. Agronomy Monograph. Madison (WI): American Society of Agronomy, Crop Science Society of America, Soil Science Society of America; 1997. p. 21–65.
[3] Miller JF. Inheritance of salt tolerance in sunflower. Helia. 1995;236:9–16.
[4] Baloglu MC, Kavas M, Aydin G, et al. Antioxidative and physiological responses of two sunflower (Helianthus annuus) cultivars under PEG-mediated drought stress. Turk J Bot. 2012;36(6):707–714.
[5] León AJ, Andrade FH, Lee M. Genetic analysis of seed-oil concentration across generations and environments in sunflower. Crop Sci. 2003;43(1):135–140.
[6] López Pereira M, Trápani N, Sadrás VO. Genetic improvement of sunflower in Argentina between 1930 and 1995. Field Crops Res. 2000;67(3):215–221.
[7] Monotti M. Growing non-food sunflower in dryland conditions. Ital J Agron. 2003;8(1):3–8.
[8] Delplanque B. Intérêt nutritionnel des huiles de tournesol: tournesol linoléique et tournesol à haute teneur en oléique The nutritional value of sunflower oils: linoleic sunflower seeds and seeds with high oleic content. Oléagineux, Corps Gras, Lipides. 2000;7(6):467–472.
[9] Ramsden CE, Narayana M, Ramakrishnan C, et al. Use of dietary linoleic acid for secondary prevention of coronary heart disease and death: evaluation of recovered data from the Sydney Diet Heart Study and updated meta-analysis. BMJ. 2013 [cited 2017 Nov 30];346:e8707. [18 p.]
DOI:10.1136/bmj.e8707.
[10] Grundy SM. Comparison of monounsaturated fatty acids and carbohydrates for lowering plasma cholesterol. N Engl J Med. 1986;314(12):745–748.
[11] Premnath A, Narayana M, Ramakrishnan C, et al. Mapping quantitative trait loci controlling oil content, oleic acid and linoleic acid content in sunflower (Helianthus annuus L.). Mol Breed. 2016 [cited 2017 Nov 30];36(7):106.[7 p.]
DOI:10.1007/s11032-016-0527-2.
[12] Yen CL, Stone SJ, Kolliwad S, et al. Thematic review series: glycerolipids. DGAT enzymes and triacylglycerol biosynthesis. J Lipid Res. 2008;49(11):2283–2301.
[13] Nakamura MT, Nara TY. Structure, function, and dietary regulation of delta6, delta5, and delta9 desaturases. Annu Rev Nutr. 2004;24:345–376.
[14] Walls JC, Browse J. Mutants of Arabidopsis reveal many roles for membrane lipids. Prog Lipid Res. 2002;41(3):254–278.
[15] Chi X, Yang Q, Lu Y, et al. Genome-wide analysis of fatty acid desaturases in soybean (Glycine max). Plant Mol Biol. 2011;29(5):679–783.
[16] Chi X, Yang Q, Pan L, et al. Isolation and characterization of fatty acid desaturase genes from peanut (Arachis hypogea L.). Plant Cell Rep. 2011;30(8):1393–1404.
[17] Ohrogwe J, Browse J. Lipid biosynthesis. Plant Cell. 1995;7:957–970.
[18] Ohrogwe J, Jaworski JG. Regulation of fatty acid synthesis. Annu Rev Plant Physiol Plant Mol Biol. 1997;48(1):109–136.
[19] Zhang YM, Wang CC, Hu HH, et al. Cloning and expression of three fatty acid desaturase genes from cold-sensitive lima bean (Phaseolus lunatus L.). Biotechnol Lett. 2011;33(2):395–401.
[20] Chapman KD, Ohrogwe JB. Compartmentation of triacylglycerol accumulation in plants. J Biol Chem. 2012;287(4):2288–2294.
[21] Harwood JL, Guschina IA. Regulation of lipid synthesis in oil crops. FEBS Lett. 2013;587(13):2079–2081.
[22] Zhao N, Zhang Y, Li Q, et al. Identification and expression of a stearoyl-ACP desaturase gene responsible for oleic acid accumulation in Xanthoceras sorbifolia seeds. Plant Physiol Biochem: PPB. 2015;87:9–16.
[23] Dong CJ, Cao N, Zhang ZG, et al. Characterization of the fatty acid desaturase genes in cucumber: structure, phylogeny, and expression patterns. PLoS One. 2016 [cited 2017 Nov 30];11(3):e0149917. [22 p.]
DOI:10.1371/journal.pone.0149917.
[24] Ruddell P 2nd, Whetten R, Cardinal A, et al. Effect of a novel mutation in a Delta9-stearoyl-ACP-desaturase on soybean seed oil composition. TAG Theor Appl Genet Theoretische und angewandte Genetik. 2013;126(1):241–249.
[25] Wang W, Vinocur B, Altman A. Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. Planta. 2003;218(1):1–14.
[26] Nazarli H, Faraji F, Zardashti M. Effect of drought stress and polymer on osmotic adjustment and photosynthetic pigments of sunflower. Cercetari Agronomice in Moldova. 2011;44(1):35–41.
[27] Albuquerque MdF, de Carvalho Nd. Effect of the type of environmental stress on the emergence of sunflower (Helianthus annuus L.), soybean (Glycine max (L.) Merrill) and maize (Zea mays L) seeds with different levels of vigor. Seed Sci Technol (Switzerland). 2003;31(2):465–479.
[28] Goodstein DM, Shu S, Howson R, et al. Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res. 2012;40(Database issue):D1178–1186.
[29] Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016;33(7):1870–1874.
[30] Bailey TL, Boden M, Buske FA, et al. MEME Suite: tools for motif discovery and searching. Nucleic Acids Res. 2009;37 (suppl 2):W202–W208.
[31] Quevillon E, Silventoinen V, Pillai S, et al. InterProScan: protein domains identifier. Nucleic Acids Res. 2005;33 (Web Server):W116–W120.
[32] Suyama M, Torres D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. 2006;34(Web Server issue):W609–W612.
[33] Lynch M, Conery JS. The evolutionary fate and consequences of duplicate genes. Science. 2000;290 (5494):1151–1155.
[34] Yang Z, Gu S, Wang X, et al. Molecular Evolution of the CPP–like Gene Family in Plants: Insights from Comparative Genomics of Arabidopsis and Rice. J Mol Evol. 2008;67 (3):266–277.
[35] Conesa A, Götz S. Blast2GO: A comprehensive suite for functional analysis in plant genomics. Int J Plant Genom. 2008;2008:619832. [12 p.]
DOI:10.1155/2008/619832.
[36] Berman HM, Westbrook J, Feng Z, et al. The protein data bank. Nucleic Acids Res. 2000;28(1):235–242.
[37] Kelley LA, Sternberg MJ. Protein structure prediction on the Web: a case study using the Phyre server. Nat Protoc. 2009;4(3):363–371.
[38] Hoagland DR, Arnon DI. The water-culture method for growing plants without soil. Circular California Agric Exp Station. 1950:347:32.
[39] Chaki M, Valderrama R, Fernández-Ocaña AM, et al. Mechanical wounding induces a nitrosative stress by down-regulation of GSNO reductase and an increase in S-
nitrosothiols in sunflower (Helianthus annuus) seedlings. J Exp Bot. 2010;62(6):1803–1813.

[40] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCT method. Methods. 2001;25(4):402–408.

[41] Banik M, Duguid S, Cloutier S. Transcript profiling and gene characterization of three fatty acid desaturase genes in high, moderate, and low linolenic acid genotypes of flax (Linum usitatissimum L.). Genome. 2011;54(6):471–483.

[42] Vrinten P, Hu Z, Munchinsky MA, et al. Two FAD3 desaturase genes control the level of linolenic acid in flax seed. Plant Physiol. 2005;139(1):79–87.

[43] Jain RK, Thompson RG, Taylor DC, et al. Isolation and Characterization of two promoters from linseed for genetic engineering NRCC No. 42620. Crop Sci. 1999;39(6):1696–1701.

[44] Fofana B, Cloutier S, Duguid S, et al. Gene expression of stearoyl-ACP desaturase and delta12 fatty acid desaturase 2 is modulated during seed development of flax (Linum usitatissimum). Lipids. 2006;41(7):705–712.

[45] You FM, Li P, Kumar S, et al. Genome-wide identification and characterization of the gene families controlling fatty acid biosynthesis in flax (Linum usitatissimum L.). J Proteomics Bioinform. 2014;7:310–326.

[46] Zhang Y, Maximova SN, Guiltinan MJ. Characterization of a stearoyl-acyl carrier protein desaturase gene family from chocolate tree, Theobroma cacao L. Front Plant Sci. 2015;6:239. [12 p.] DOI:10.3389/fpls.2015.00239.

[47] Yurchenko OP, Park S, Ilut DC, et al. Genome-wide analysis of the omega-3 fatty acid desaturase gene family in Gossypium. BMC Plant Biol. 2014 [cited 2017 Nov 30];14:312. [15 p.] DOI:10.1186/s12870-014-0312-5.

[48] Yukawa Y, Takaia F, Shoji K, et al. Structure and expression of two seed-specific cDNA clones encoding stearoyl-acyl carrier protein desaturase from sesame, Sesamum indicum L. Plant Cell Physiol. 1996;37(2):201–205.

[49] Shang X, Cheng C, Ding J, et al. Identification of candidate genes from the SAD gene family in cotton for determination of cottonseed oil composition. Mol Genet Genom: MGG. 2016;292(1):173–186.

[50] Kachroo A, Shanklin J, Whittle E, et al. The Arabidopsis stearoyl-acyl carrier protein-desaturase family and the contribution of leaf isoforms to oleic acid synthesis. Plant Mol Biol. 2007;63(2):257–271.

[51] Parvini F, Sicardo MD, Hosseini-Mazinani M, et al. Transcriptional analysis of stearoyl-acyl carrier protein desaturase genes from olive (Olea europaea) in relation to the oleic acid content of the virgin olive oil. J Agric Food Chem. 2016;64(41):7770–7781.

[52] Shanklin J, Cahoon EB. Desaturation and related modifications of fatty acids 1. Annu Rev Plant Physiol Plant Mol Biol. 1998;49:611–641.

[53] Strittmatter P, Spatz L, Corcoran D, et al. Purification and properties of rat liver microsomal stearyl coenzyme A desaturase. PNAS. 1974;71(11):4565–4569.

[54] Lindqvist Y, Huang W, Schneider G, et al. Crystal structure of A9 stearoyl-acyl carrier protein desaturase from castor seed and its relationship to other di-iron proteins. EMBO J. 1996;15(16):4081–4092.

[55] Cao J, Cao F, Feng S. Evolution of the RALF gene family in plants: gene duplication and selection patterns. Evol Bioinform Online. 2012;8:271–292.

[56] Söding J. Protein homology detection by HMM–HMM comparison. Bioinformatics. 2005;21(7):951–960.

[57] Parvini F, Sicardo MD, Hosseini-Mazinani M, et al. Transcriptional analysis of stearoyl-acyl carrier protein desaturase genes from olive (Olea europaea) in relation to the oleic acid content of the virgin olive oil. J Agric Food Chem. 2016;64(41):7770–7781.

[58] Han Y, Xu G, Du H, et al. Natural variations in stearoyl-acp desaturase genes affect the conversion of stearic to oleic acid in maize kernel. Theor Appl Genet. 2017;130(1):151–161.