Identification, evolution, expression, and docking studies of fatty acid desaturase genes in wheat (*Triticum aestivum* L.)

Zahra Hajiahmadi 1, Amin Abedi 1, Hui Wei 2, Weibo Sun 2, Honghua Ruan 2, Qiang Zhuge 2 and Ali Movahedi 2*

**Abstract**

**Backgrounds:** Fatty acid desaturases (FADs) introduce a double bond into the fatty acids acyl chain resulting in unsaturated fatty acids that have essential roles in plant development and response to biotic and abiotic stresses. Wheat germ oil, one of the important by-products of wheat, can be a good alternative for edible oils with clinical advantages due to the high amount of unsaturated fatty acids. Therefore, we performed a genome-wide analysis of the wheat FAD gene family (*TaFADs*).

**Results:** 68 FAD genes were identified from the wheat genome. Based on the phylogenetic analysis, wheat FADs clustered into five subfamilies, including *FAB2*, *FAD2/FAD6*, *FAD4*, *DES/SLD*, and *FAD3/FAD7/FAD8*. The *TaFADs* were distributed on chromosomes 2A-7B with 0 to 10 introns. The Ka/Ks ratio was less than one for most of the duplicated pair genes revealed that the function of the genes had been maintained during the evolution. Several cis-acting elements related to hormones and stresses in the *TaFAD* promoters indicated the role of these genes in plant development and responses to environmental stresses. Likewise, 72 SSRs and 91 miRNAs in 36 and 47 *TaFADs* have been identified. According to RNA-seq data analysis, the highest expression in all developmental stages and tissues was related to *TaFAB2.5, TaFAB2.12, TaFAB2.15, TaFAB2.17, TaFAB2.20, TaFAD2.1, TaFAD2.6*, and *TaFAD2.8* genes while the highest expression in response to temperature stress was related to *TaFAD2.6, TaFAD2.8, TaFAB2.15, TaFAB2.17*, and *TaFAB2.20*. Furthermore, docking simulations revealed several residues in the active site of TaFAD2.6 and TaFAD2.8 in close contact with the docked oleic acid that could be useful in future site-directed mutagenesis studies to increase the catalytic efficiency of them and subsequently improve agronomic quality and tolerance of wheat against environmental stresses.

**Conclusions:** This study provides comprehensive information that can lead to the detection of candidate genes for wheat genetic modification.

**Keywords:** Expression pattern, Gene family, Molecular docking, Protein structure

*Correspondence: ali_movahedi@njfu.edu.cn
Zahra Hajiahmadi and Amin Abedi are equally as the first author.

© The Author(s). 2020 Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.
Background
Polysaturated fatty acids (PUFAs) are essential components of the plasma membrane. Various PUFAs have crucial roles in plant physiological and cellular processes such as cold acclimation, defense mechanisms against biotic and abiotic stresses, and chloroplast development [1]. PUFAs biosynthesis occurs through different and complex pathways of desaturation and elongation steps [2]. Fatty acid desaturase (FAD) enzymes introduce double bond into fatty acids hydrocarbon chain. Two groups of FAD have been identified in plants, including acyl-acyl carrier protein (acyl-ACP) desaturases and membrane-bound FADs or acyl-lipid desaturases [3]. While identified FADs in plants, animals, algae, and fungi are membrane-bound desaturase, the plant acyl-ACP desaturase (FAB2/SAD) is the only soluble FAD [4, 5]. The acyl-ACP desaturases introduce the first double bond into the acyl chain of saturated fatty acid in plastids. Besides, Membrane-bound FADs exist in chloroplast and endoplasmic reticulum (ER). Desaturation processes occur through two different pathways in the plastid and ER [6]. In the chloroplast and ER, double bond formation requires NADPH/ferredoxin and NADH/cytochrome b₅ systems as the electron donors, respectively [7].

On the other hand, the quality of edible oils depends on the unsaturated fatty acids content [8]. FADs are essential to determine the quality of edible oils [9]. They have been attracted more attention due to their ability to adjust the level of unsaturated fatty acids to increase the quality of these oils and plant resistance against various stresses including drought, salt, heat, cold, and pathogen [10–13]. For instance, the cell membrane is the primary site for cold-induced injury, and the melting temperature of the unsaturated fatty acids is less than saturated fatty acids. Therefore, adjustment of membrane lipid fluidity through manipulation of FADs and changing the levels of unsaturated fatty acids might seem helpful for cold acclimation [14]. To date, several studies have been conducted to assess the expression of genes encoding fatty acid desaturase in response to biotic and abiotic stresses [12, 15–17]. Investigation of the expression of SACPD-A and SACPD-B genes (encoding soluble Δ9 stearoyl-ACP desaturases) and the amount of stearic acid (C18:0) and oleic acid (C18:1) in soybean revealed that the number of transcripts of both genes and oleic acid had been dramatically increased in low temperature. Reversely, we observed an increased amount of C18:0 and decreased the expression of the genes above at high temperatures [18]. Wang et al. (2012) ascertained the expression of oleate desaturase (GbFAD2 and GbFAD6) and GbSAD genes under various temperatures in *Ginkgo biloba* L. leaves. Based on their results, the expression of GbFAD2 and GbSAD genes has been increased in 4 and 15°C, while it has been prevented in 35 and 45°C.

In contrast, the expression of GbFAD6 was constant at different temperatures [19]. The expression of FAD2–1 and FAD2–2 genes of olive has been increased in response to wounding [20]. Likewise, FAD2 and FAD6 genes are necessary for salt tolerance during early seedling in Arabidopsis [21, 22]. Zhang et al. (2005) developed transgenic tobacco plants with the overexpressing FAD3 or FAD8 genes. According to their findings, the over-expression of FAD8 or FAD3 genes caused enhanced tolerance to drought [23]. The importance of FADs in plant pathways has been confirmed previously. A homologous region based on a conserved sequence of a gene family can be applied to identify new genes. The FAD gene family is vital for the production of PUFAs in plants; thus, a comprehensive understanding of FAD genes using bioinformatics studies can help disclose their functions in the studied plants.

Wheat (*Triticum aestivum* L.) is one of the most important cereal crops. Because of the high amount of unsaturated fatty acids, wheat germ oil, one of the essential by-products of wheat, can be a good alternative for edible oils with clinical benefits. Based on studies, wheat germ oil contain different fatty acids, including linoleic acid (C 18:2), palmitic acid (C 16:0), oleic acid (C 18:1), linolenic acid (C18:3), and stearic acid (C 18:0) [24]. Wheat is a good source of edible oil, and the characterization and analysis of the FAD family in wheat plants have not yet been performed. On the other hand, comprehensive analyses on gene families help to address a better understanding of their evolutions and functions in plants [25]. Therefore, in this study, identification, evolutionary relationship, duplication and selection pressure, exon-intron structure, promoter analysis, transcript-targeted miRNA and simple sequence repeat markers prediction, RNA-seq data analysis, three-dimensional structure, and docking studies of the *TaFADs* have been investigated in wheat using bioinformatics tools. Figure 1 provides a flow-chart of the data analysis process.

Results
Identification of *T. aestivum* FAD genes
In the current study, 82 TaFAD (coding by 68 genes) have been identified. All identified FAD genes were named based on a phylogenetic dendrogram, and their numbering was according to the location of genes on chromosomes. The physicochemical study of the identified genes was carried out using the ProtParam tool. Based on the results, these genes are different in the number of amino acids, molecular weight (MW), and isoelectric point (pI). The protein sequence encoded by these 68 FAD genes ranged in length from 281 amino acids...
acids of TaFAD4.2 to 518 amino acids of TaFAD6.2. The theoretical molecular weights of these proteins ranged from 29.74 to 59.63 kDa, with the isoelectric points varied from 5.42 to 9.72 (Additional file 1: Table S1). The results of cellular localization of proteins revealed that they are active in chloroplast, mitochondria, endoplasmic reticulum, and plasma membrane. The spatial diversity of these genes is likely related to the diverse functional roles of these genes in different cell processes.

Phylogenetic analysis of the wheat FAD gene family
We carried out a phylogenetic tree using the Neighbor-joining method to investigate relationships of wheat FAD proteins. According to Fig. 2, wheat FAD proteins were divided into five groups. The first group is FAB2, which is called stearoyl-ACP desaturase (SADs) in plants. It can introduce a double bond into the stearoyl-ACP at the delta-9 position. Based on the phylogenetic tree, the FAB2 subfamily has been noticeably separated from other subfamilies. All members of wheat and rice FAB2 were clustered together, whereas FAB2 of Arabidopsis and soybean (as dicot plants) were grouped. FAD4 group introduces a double bond into a saturated acyl chain at the delta-3 position. Likewise, all FAD4 proteins of monocot plants were clustered together and separated from dicot plants clade. In FAD2/FAD6 (Delta-12 desaturase, omega 6) subfamily, FAD2 and FAD6 were divided into two separate clades due to their subcellular localization (Additional file 1: Table S1). FAD2 is endoplasmic reticulum-type omega 6, while FAD6 is chloroplast-type omega 6. In FAD3/FAD7/FAD8 (Delta-15 desaturase, omega 3) subfamily, the FAD7 clade is closer to the FAD8 clade in wheat, which may be due to the high similarity of their sequences. The subcellular localization of FAD7 and FAD8 is the chloroplast, while the FAD3 is endoplasmic reticulum-type omega 3.

Gene location on a chromosome, gene duplication, and selection pressure of FAD genes
To assess the distribution of FAD family genes, a chromosome map has been constructed, and an uneven distribution of 68 FAD genes on wheat chromosomes was determined (Fig. 3). Chromosomes (Chr) 2A, 2D, 2B, 5A, and 5B had the highest number of genes while there was no gene of this family on chr1A-1D and 7D. Only one gene has been found on chr7A and chr7B related to TaFAB2.33 and TaFAB2.34, respectively. Different members of the FAB2 subfamily are located on chr2A-7B except for chrB and chr4D. In SLD and DES subfamilies, the genes were determined on chr5A-5D and chr2A-2D, respectively. FAD3, FAD6, FAD7, and FAD8 subfamilies have been found on chr4A-4D, chr4A-5D, chr2A-2D, and chr2B-2D, respectively.
Two types of tandem and segmental duplication were observed in the wheat FAD gene family. However, according to the higher frequency of segmental duplication, its role in the expansion of the FAD gene family is much greater than tandem duplication. Gene duplication is one of the important mechanisms to increase genetic diversity and generate new genes in plants. In the current study, it was found that three genes on chr6A (TaFAD2.1–3), four genes on chr6B (TaFAD2.4–7), and four genes on chr6D (TaFAD2.8–11) are the result of tandem duplication. Likewise, three genes on chr2A (TaFAB2.1–3), chr2D (TaFAB2.9–11), chr5B (TaFAB 2.26–28), and chr5A (TaFAB2.23–25) are the result of tandem duplication. We investigated Ka, Ks, and Ka/Ks parameters for 153 paired genes (Additional file 1: Table S2) to reveal a functional selection pressure between duplicated genes. In general, Ka/Ks > 1, Ka/Ks = 1, and Ka/Ks < 1 indicate positive, neutral, and negative selections, respectively [26]. The Ka/Ks ratio was less than one for most of the paired genes. This negative selection is to maintain the function of the FAD gene family in wheat plants, and they were under a slow evolutionary process, and almost their role in evolution has been maintained. However, the Ka/Ks ratio for seven paired genes (TaFAD2.4/TaFAD2.7, TaFAD2.6/TaFAD2.5, TaFAD2.2 /TaFAD2.6, TaFAD2.1/TaFAD2.2, TaFAD2.9/TaFAD2.10, TaFAD2.15/TaFAD2.20, and TaFAD2.1/ TaFAD2.11) was greater than one, indicated the aforementioned paired genes were under the positive selection during evolution and they have different functions due to the mutations that have been occurred during their evolution. The divergence time of duplications was estimated at 1.98–47.75 Myra.

**Exon-intron structures and conserved motifs**

Based on the analysis of the exon-intron structure of the FAD gene family, they had 0 to 10 introns with high structural diversity (Fig. 4). 21.95% of wheat FAD genes are intronless. The longest intron is related to the TaFAD2.7 gene. Three intron splicing phases have been observed for the wheat FAD gene family, including phase zero, splicing after the third nucleotide of the codon; phase one, splicing after the first nucleotide of the codon; and phase two, splicing after the second nucleotide [27]. Most of the genes in this family have phases zero and two, whereas TaFAD6.1–2 and TaFAB2.4 genes have all three intron splicing phases. In TaFAB2.13, TaFAB2.18, TaFAB2.22a, TaFAB2.15, and
TaDES.1–3 genes only phase zero have been observed. Likewise, TaFAB2.1–3, TaFAB2.23–25, TaFAB2.27–28, TaFAB2.6–7, TaFAB2.9–11, TaFAB2.33–34, and TaFAB2.31 genes demonstrated only phase two. The genes of the same subfamily were more similar in exon-intron structure compared with the genes of the other subfamilies. The FAD3/7/8 and DES subfamilies have eight and two exons, respectively. However, the FAB2 subfamily has 2–3 exons except for TaFAB2.4b with 10 exons. The SLD, FAD2, and FAD4 subfamilies have one exon except for TaFAD2.2, TaFAD2.5b, TaFAD2.6c, and TaFAD2.7 with two exons. The FAD6 subfamily contains 10 exons, while TaFAD6.1b has seven exons. A comparison of gene structure and phylogenetic tree demonstrated that the gene structure and intron splicing phases were similar in each gene cluster except TaFAB2.4b (Fig. 4).

MEME tool has been applied to detect conserved motifs in protein sequences of the FAD family (Additional file 1: Table S3 and S4). Based on the results, 30 and 10 conserved motifs in the FAD and FAB2 subfamilies have been identified, respectively (Figs. 5 and 6). Evaluation of motifs with Pfam showed that motifs 1, 2, 8, 9, and 10 in the FAD subfamily are related to the FAD gene family. The motifs 1–5 in the FAB2 subfamily are also related to the fatty acid desaturases. Conserved motif 13 (Cytochrome b5-like Heme/Steroid binding domain) was presented only in the TaSLD1–3 (Additional file 1: Table S4). Motif 18 (AAAARADSGEA) with no function was found in all the members of the FAD subfamily. The lowest number of motifs was related to TaFAD4.1–3 with two motifs (Motifs 18 and 10). Motifs 1, 2, 4, and 9 were common in all the members of the FAB2 subfamily except TaFAB2.4. TaFAB2.4 has only motifs 8 and 10 with no function. All members of the FAD subfamily contain three conserved histidine boxes (Additional file 1: Table S5). The amino acid composition of His-boxes is highly conserved in the same subfamilies. The third His-box with consensus sequence HX3HH exists in all TaFAD except in the TaSLD1–3 that the first amino acid is glutamine instead of histidine, which is vital for

---

**Fig. 3** Chromosomal locations of wheat fatty acid desaturase genes. Chromosomes are represented by Colored boxes. Dark teal curves connecting the genes indicate duplications. The location of genes on chromosomes and the duplication relationship between them were presented using TBtools.
its enzymatic activity. All members of the \textit{FAB2} subfamily contain two conserved histidine boxes. The first His-box with consensus sequence EENRHG exists in all TaFAB2 except TaFAB2.8 with the consensus sequence of EENRML, whereas the sequence of the second His-box (DEKRHE) is identical in all \textit{FAB2} subfamily members.

Detection of cis-acting elements in TaFAD promoters

To better understanding the TaFAD genes expression regulation mechanisms, the search of cis-acting elements 1500 bp upstream of starting codon (ATG) of TaFAD genes was performed using the PlantCare database. In total, 82 motifs were identified in this family that the abundance of these elements was highly varied for each.
gene (Additional file 1: Table S6). The highest frequency of motifs was related to STRE (95.58%), MYC (94.11%), MYB (89.70%), TGACG-motif (88.23%), CGTCA-motif (88.23%), as-1 (88.23%), ABRE (85.29%), and G-box (83.82%). Likewise, the lowest frequency of motifs was related to BoxIII (only in TaFAB2.21), AT1-motif (only in TaFAD2.11), xhs-MA1a (only in TaFAB2.24), L-box (only in TaFAD3.2), LAMP-element (only in TaFAD8.1), AAAC-motif (only in TaFAB2.10), F-box (only in TaFAB2.3), Plant_AP_2_like (only in TaFAB2.15), re2f-1 (only in TaFAB2.23), and OCT (only in TaFAB2.28).

Simple sequence repeats (SSRs) in TaFAD genes and TaFAD-targeted miRNAs prediction

72 SSRs have been identified in 36 of the 68 TaFAD genes including 33, 25, 11, 2, and 1 SSRs in FAD3/ FAD7/FAD8, FAB2, FAD2/FAD6, DES/SLD, and FAD4 subfamilies, respectively (Additional file 1: Table S7). Most genes had a single SSR except FAB2.22 (5SSRs), FAD8.1 (5 SSRs), FAD8.3 (5 SSRs), FAD2.7 (5 SSRs), FAD3.1 (4 SSRs), FAD3.3 (4 SSRs), FAD3.2 (3 SSRs), FAD3.4 (3 SSRs), FAD3.6 (2 SSRs), FAD8.2 (3 SSRs), FAB2.15 (3 SSRs), FAB2.20 (2 SSRs), FAB2.34 (3 SSRs), FAB2.31 (2 SSRs), and FAD2.8 (2 SSRs). The highest frequency was related to tri-nucleotide repeats (36 SSRs) followed by tetra-nucleotide repeats (6 SSRs), penta-nucleotide repeats (13 SSRs), hexa-nucleotide repeats (4 SSRs), and di-nucleotide repeats (3 SSRs). 91 miRNAs for 47 TaFADs possible targets have been identified (Additional file 1: Table S8). The relationship between miRNAs and their targets was not one by one and many miRNAs had a common target. For instance, TaSLD3 transcript was co-targeted by 5 miRNAs named tae-miR9659-3p, tae-miR1137a, tae-miR395a, tae-miR395b, and tae-miR9677b. On the contrary, one miRNA had multiple transcript targets such as tae-miR5384-3p can suppress the expression of TaFAB2.1, TaFAB2.9, TaFAB2.11, TaFAB2.21, TaFAB2.32, TaFAB2.33, and TaFAB2.34.

Expression analysis of TaFAD genes

To understand the functional role of TaFAD genes, their expression patterns of different developmental stages have been studied in the wheat root, shoot, leaf, grain,
The members of the TaFAD family showed different expression patterns. All members of DES/SLD subfamily revealed low transcript level at all developmental stages and tissues except TaSLD3 (reduced expression in leaves and shoots of seedling and growth cycles, high expression at the vegetative stage in roots and spikes), TaSLD1 (moderate expression at seedling stage in roots), and TaSLD2 (reduced expression at the vegetative stage in spikes). In the FAD2/FAD6 subfamily, all TaFAD2 genes had low expression except TaFAD2.1, TaFAD2.6, and TaFAD2.8 with a high level of transcripts in all stages and tissues. However, TaFAD6.1–2 genes had a high level of expression in shoots, leaves at all developmental stages, and spike at the reproductive phase. Based on RNA-seq data analysis of the TaFAD

![Image](image-url)
subfamily, all members demonstrated a low level of transcripts in all stages and tissues. In the FAD3/FAD7/FAD8 subfamily, the low level of transcript expression of TaFAD7.1–3 genes has been observed in vegetative, reproductive, and seedling tissues, while high abundance expression of them has been detected in roots and vegetative phase of spikes. TaFAD8.1–3 genes had a high level of expression in leaves and shoots of all wheat developmental stages. However, TaFAD8.1, TaFAD8.3, and TaFAD8.2 genes had low, moderate, and high expressions at the reproductive stage in spikes, respectively.

The expression patterns of TaFAD3.1–3 genes were similar and low in all stages and tissues except TaFAD3.1 with moderate expression in roots at the seedling stage. However, TaFAD3.4–6 genes showed a high level of expression in all tissues at vegetative and spikes at reproductive stages. In FAB2 subfamily, TaFAB2.1–3, TaFAB2.6–7, TaFAB2.9–11, TaFAB2.14, TaFAB2.16, TaFAB2.18–19, and TaFAB2.21–34 revealed a low level of transcripts in all stages and tissues whereas TaFAB2.4–5, TaFAB2.12, TaFAB2.15, TaFAB2.17, and TaFAB2.20 had a high abundance of transcripts in all developmental stages except TaFAB2.4 with low expression in roots and spike at growth phase, roots at seedling, roots, and grains at reproductive stage. Likewise, TaFAB2.8 had high expression in all conditions except seedling phases with moderate expressions. Finally, TaFAB2.13 showed a low level of transcripts in all conditions except roots at seedling and growth cycles with high expression and leaves/shoots at the growth cycle with moderate expression. Besides, we investigated the expression patterns of TaFAD genes to predict their roles responding to biotic and abiotic stresses (Figs. 8 and 9) (Additional file 1: Table S9). In response to powdery mildew pathogen, we observed the up-regulated expression of TaFAB2.4 after 24 h of treatment. In this observation, TaFAD8.1–3, TaFAB2.1, TaFAB2.15, and TaFAB2.17 revealed high expressions during the stress.

On the other hand, the expression of the TaFAD3.4, TaFAD3.5, TaFAD3.6, TaFAD3.8, TaSLD1–3, and TAFA B2.20 (moderate expression) remained unchanged and it has been down-regulated in TaFAB2.5 and TaFAB2.12 at 72 h compared with 24 h. The expression of TaFAD8.1–3, TaFAD4.2, TaFAD4.3, and TaFAD3.5–6 has been increased in response to the stripe rust pathogen CYR31. Although, the transcript level of TaFAD2.1, TaFAD2.6, TaFAD2.8, TaFAB2.1, TaFAB2.12, TaFAB2.17 (high expression) and TaSLD1–3, TaFAD6.1–2, TaFAB2.20 (moderate expression) remained unchanged. After 6 h of drought stress, the expression of all TaFAD genes has been down-regulated, obviously except TaFAB2.5, TaFAB2.8, and TaFAB2.12 that showed up-regulation. The TaFAD2.1, TaFAD2.6, TaFAD2.8, TaFA B2.5, TaFAB2.8, TaFAB2.12, TaFAB2.15, TaFAB2.17, and TaFAB2.20 could be up-regulated by heat stress after 6 h. Under drought and heat treatment, the expression of TaFAD7.1–3 and TaFAD8.1–3 has been decreased whereas the expression of TaFAD2.1, TaFAD2.6, TaFAD2.8, TaFAB2.5, TaFAB2.8, TaFAB2.12 TaFAB2.15, TaFAB2.17, and TaFAB2.20 have been induced more significantly at 6 h. The expression of 11 TaFAD genes has been up-regulated under cold stress including TaFAD2.6, TaFAD2.8, TaFAD6.1–2, TaFAD8.1–3, TaFAB2.15, TaFAB2.17, and TaFAB2.20 while the transcript level of the TaFAB2.4, TaFAB2.5, TaFAB2.8, TaFAD2.12, TaFAD3.4–6, TaFAD4.2–3, and TaSLD1–3 was moderate.

TaFAD2.6 and TaFAD2.8 structural modeling and docking studies

The FAD2 members have been shown to play important roles in wheat response to environmental stresses. Microsomal omega-six fatty acid desaturases (FAD2) introduce a double bond to the oleic acid (C18:1) carbon chain resulting in linoleic acid (C18:2) [28]. Linoleic acid is a precursor of other polyunsaturated fatty acids, and it is essential for the growth of all eukaryotes [29]. On the other hand, among the FAD2 members, TaFAD2.6 and TaFAD2.8 proteins had the highest expression in all studied stresses. Therefore, in the present study, their molecular structure and interaction of ligand-enzyme have been evaluated. Three-dimensional structures of TaFAD2.6 and TaFAD2.8 proteins have been predicted and refined by I-TASSER and ModRefinder servers, respectively. According to the results of the Ramachandran analysis of primary and refined models, the residue...
count increased in favored regions from 68.4 to 86.0% and from 69.1 to 87.0% in TaFAD2.6, and TaFAD2.8, respectively, confirming that the refined models are of good qualities. The modeled structure for TaFAD2.6 has 18 \(\alpha\)-helices and 19 loops (Fig. 10b) while TaFAD2.8 structure has 19 \(\alpha\)-helices and 20 loops (Fig. 11b). Most of the predictor programs predicted that the TaFAD2.6 and TaFAD2.8 proteins have six transmembrane (TM) \(\alpha\)-helices (Fig. 10a and Fig. 11a) with a consensus domains of TM1 (Val86-Val109), TM2 (Ala114-Ile132), TM3 (Ser145-Trp158), TM4 (Arg208-Val213), TM5 (Gln254-Val273), TM6 (Trp280-Gln303) for TaFAD2.6 and TM1 (Asp65-Val86), TM2 (ALA91-Ile109), TM3 (Ser122-Trp135), TM4 (Trp197-Asn204), TM5 (Asp236-Ser251), TM6 (Phe255-Thr282) for TaFAD2.8. Three conserved histidine boxes have been found in the TaFAD2.6 and TaFAD2.8. Based on the reports in other plant species, the histidine boxes interact with two irons at the catalytic sites of the FAD2 enzymes [30, 31]. The predicted structures of TaFAD2.6 contains three conserved histidine boxes with eight histidine residues of His134, His138, His170, His173, His174, His345, His348, and His349 residues involved substrate-binding are part of HECGH, HRRHH, and HVAHH histidine boxes. In the TaFAD2.8, His111, His147, His150, His151, His322, His325, and His326 residues involved in the active site are also part of three conserved histidine boxes (Additional file 1: Table S5). These results suggest that the conserved histidine boxes play critical roles in the binding of the ligand to the active site. The results mentioned above may be useful for future site-directed mutagenesis studies to increase the catalytic efficiency of the FAD2 enzymes and subsequently enhance wheat tolerance to different stresses.

**Discussion**

In the current study, 68 members of the \(FAD\) gene family have been founded in wheat that are significantly larger than the number of \(FAD\) genes in \(A. thaliana\) (25) [16], \(O. sativa\) (19) [33], \(Glycine max\) (29) [34], and \(Medicago truncatula\) (20) [35]. Therefore, the expansion of the \(FAD\) gene family is species-specific in different plants, and this expansion is the result of gene duplication events [36]. One possible reason for this expansion may be related to the larger genome of wheat compared to other species [37]. The uneven distribution of the \(FAD\) gene family has been observed in wheat, which reported in soybean [34], alfalfa [35], cotton [16], and rapeseed [12] as well. The \(FAD\) family genes have been...
widely distributed in the wheat genome, which may indicate that they originated from different ancestors [34]. Studying the structure of genes in life sciences is necessary, and it can be a great guide to identifying the evolution of genes. The stability of genes structure is a prerequisite to maintain their functional role, while variation in gene structure is essential for the functional evolution of gene families [38]. In the present survey, segmental duplication (83.66%) of the *FAD* gene family in wheat was far more than tandem duplication (16.33%). In a study on duplication of 50 large gene families in Arabidopsis, it was concluded that there was a negative relationship between tandem and segmental duplications and when each of these duplication events was greater, the other type played the least role in the expansion of gene family members [39]. It should be...
noted that if the duplicated genes are maintained in evolution, they generally change in their regulatory or coding region. Changing in coding regions, especially if they cause changes in function, are caused by amino acid substitution or changes in exon-intron structure [40]. The study of the selection pressure of the FAD gene family in wheat showed that a strong purifying selection has been occurred indicating the importance of the functional role of these genes. Generally, the divergence time of tandem duplication is more recent phenomena than segmental duplication and stress-responsive genes tend to located on tandem clusters [41, 42].

There are two major groups of desaturases, including soluble desaturase and membrane-bound desaturase [34]. Based on the phylogenetic study, five desaturase subfamilies have been identified, including delta-15 desaturase (FAD3/FAD7/FAD8), stearoyl-ACP desaturase (FAB2), delta-12 desaturase (FAD2/FAD6), delta-3
desaturase (FAD4), and front-end desaturase (DES/SLD). To date, stearoyl-ACP desaturase is the only identified soluble desaturase. The phylogenetic tree and the type of clustering of FAD proteins are consistent with the other plants, including sunflower [10], cucumber [43], and soybean [34]. Each cluster had similar amino acid compositions, and it can be concluded that the phylogenetic distribution of wheat FAD proteins is related to their motif contents. All members of the FAB2 cluster contained motifs 1, 2, 4, and 9 except TaFAB2.4 with motifs 8 and 10. The FAD2/FAD6 subfamily had common motifs 1, 2, 16, and 18. The members of the FAD6 group contained specific motifs 20, 22, and 28, while the members of the FAD2 cluster had special motifs 11, 12, 14, and 17 which may be the reason for the separation of these groups in the dendrogram. All members of the FAD subfamily contained motif 1 except DES, SLD, and TaFAD4. The main difference between TaSLD1–3 and TaDES1–3 was the presence of specific motifs 8, 19, 29, and 30 in TaSLD1–3. On the other hand, the FAD4 group had only two motifs (10 and 18); thus, it was completely separated from the other groups of FAD subfamily. The members of the FAD3/FAD7/FAD8 group had six common motifs as well. TaFAD7 and TaFAD8 had quite similar motifs. Therefore, their clusters were closer to each other than to FAD3. The protein sequence of FAD7 and FAD8 is very similar. The transcript levels of FAD7 and FAD8 increase at high temperatures and low temperatures, respectively [16, 44]. Two and three His-boxes have been observed in the FAB2 and FAD subfamilies, respectively. The identified His-boxes in the FAB2 subfamily are related to their active sites [7]. All members of the FAD subfamily contain three His-boxes except TaFAD2.4, which had two His-Boxes and also found in a separate clade from the other TaFAD2 in the phylogenetic tree. The third His-box was located in the carboxy-terminus of TaFAD proteins, and its consensus sequence was H/QX3HH, which is similar to FAD protein sequences in various plants [7]. The number of residues between the first and second His-boxes varied between different members of the TaFAD subfamily (Additional file 1: Table S5). For instance, the amino acid length between His-box 2 and His-box 3 was 111 or 171 in TaFAD2.1–11, 156 in TaFAD6.1–2, 162 in TaFAD3.1–6, 163 in TaFAD7.1–3 and TaFAD8.1–3, 173 in TaSLD1–3, and 127 in TaDES1–3. The exception was the TaFAD4 cluster, in which this length was only 25 amino acids. Synthesis of fatty acids occurs in plastids but their desaturation is in plastids and endoplasmic reticulum [45]. In most studies, subcellular localization of the FAB2 and FAD subfamilies is chloroplast and chloroplast/ER, respectively. However, according to our analysis, the subcellular localization of the FAB2 subfamily was both chloroplast and mitochondria (Additional file 1: Table S1) which is in line with Diaz et al. (2018) findings. Likewise, it was observed that all members of TaFAD4 and DES/SLD clusters were localized in the plasma membrane. The FAD gene family in wheat revealed all three types of splicing phase (zero, one, and two). The splicing phase of wheat FAD gene family was almost similar to those of other plants, indicating a high degree of the FAD genes conservation during evolution [21, 34, 35, 43]. Eighteen members of the wheat FAD gene family lacked introns. In eukaryotic genomes, there are three ways to form intronless genes including horizontal gene transfer from prokaryotes, duplication of intronless genes, and retrotransposons [46].

Investigating the promoter regions helps to understand the interaction and function of genes. Transcription factors play an essential role in the regulation of signaling pathways in response to abiotic stresses. The proteins of these factors bind to the promoter of the target genes, regulating their induction or repression activity [47]. The TGACG and CGTCA motifs are on the genes that respond to methyl jasmonate [48]. ABRE and MBS are also the regulatory elements in response to abscisic acid (ABA) and drought, respectively. Jasmonate involves in seed germination, senescence, and response to biotic and abiotic stresses [49]. ABRE motif with the sequence of TACGGTC activates in response to ABA, resulting in plant tolerance against drought and salinity. The high abundance of cis-acting elements related to response to jasmonate, ABA, drought, cold, pathogen, auxin, gibberellic, and ethylene suggest that the TaFAD genes are involved in response to a wide range of stress. However, the existence of a particular cis-acting element in the promoter of a gene is not a definite reason for the expression of that gene in response to the same stress or hormone. This could be due to the complexity of the mechanism of gene expression regulation and the limitation of bioinformatics tools in predicting the cis-acting elements of the promoters [50]. Therefore, using experimental methods such as qRT-PCR is essential for identifying functional regulatory elements in TaFAD promoters. In general, it is known that desaturases can respond to different stresses. For instance, when plants exposed to low temperatures, changes in the expression of their FAD genes occur that result in changes in membrane lipid fluidity and eventually increase cold tolerance [51]. It has been reported that the changes in the desaturation of fatty acids result from the regulation of the FAD genes in response to different stresses [23].

SSRs are tandem repeats of 1–6 nucleotides that have been reported to play important roles in regulating gene expression [52]. The location of tri-nucleotide SSRs in untranslated regions can induce gene silencing. Their distribution in the coding regions also affects transcription, translation, and gene function [53]. In the current
study, the highest abundance has been observed for the tri-nucleotide repeats (50%) in TaFADs that is in line with previous studies in the wheat genome [54, 55]. The dominant SSRs are different in various plant species. For instance, in monocots (except maize), most of the dominant SSRs are CCG/CGG/GGC/GCC, while most of the dominant SSRs in dicots (except Arabidopsis) are AAT/ATT/ATA/TAT/TAA/TTA [56]. Based on the results of Qin et al. (2015), the type of dominant SSRs is taxon-dependent, and the abundance of AT in the dicots genome is higher than that of the monocots. In the future, SSR polymorphisms in TaFADs may be investigated in different cultivars and may be suitable for marker-assisted selection development in wheat breeding programs to select genotypes with a better quality of germ oil. MicroRNAs (miRNA) are kind of non-coding small RNAs usually 19–24 bp in length. They play an important role in controlling post-transcriptional changes. miRNAs are widespread in plants, animals, and viruses. Likewise, they play important roles in plant development and responses to biotic and abiotic stresses [57]. Predicting the target of miRNAs using bioinformatics tools has made it convenient and fast. In FAB2, FAD2, and FAD4 subfamilies, 24, 8, and 1 transcript were targeted by wheat miRNAs, respectively. However, all members of FAD7, FAD8, and SLD transcripts were targeted by miRNAs. In wheat, studies have been carried out to identify miRNAs and their crucial role in stress tolerance [58–61]. miR159 plays a role in leaf development. The members of the MYB family influence response to stress through hormones and signaling networks [62]. miR160 can play a role in response to cold stress by targeting MYB3 transcription factors [63]. They also involve in root development [64]. In the current study, the putative target of miR160 was TaFAD8.2 with an auxin-responsive element in the promoter. The accumulation of auxin-responsive elements through miR160 downregulation enhances the response to auxin and resulting in enhancement of root and leaf development [65]. Induction of miR408 expression occurs under root and leaf dehydration stress [66]. miR395 and miR530 are more expressed in the leaves of plants and control the expression of the genes related to plant development [59]. Therefore, TaFAB2.3 and TaFAD8.3 are probably involved in leaf development. miR1120 is important in regulation of the meiosis and early anther development in wheat [67], thus, the TaFAB2.15 may be involved in the reproductive development of wheat plants. miR164 is also essential for reproductive development and responding to various abiotic stresses [68, 69] that targets TaDAD2.5 in this study. miR408 is involved in response to copper deprivation [70] and drought stress [71], which has also been demonstrated in wheat [72]. Likewise, miR1139 is involved in wheat’s response to phosphate deficiency [73]. miR9659 and miR9657b are seed-specific [74], whereas miR9666b and miR9670 affect grain development in wheat [75]. Therefore, TaFAD2.1, TaFAD2.4, TaFAD2.6–8, TaFAD2.10–11 are probably involved in reproductive development. According to other studies, miR5048 and miR5049 are involved in response to drought [75, 76]. miR5384 can regulate the two pathways of cell death and growth [77]. Finally, miR9772 and miR9678 are essential for the response to heat stress [78] and germination of wheat seeds [79].

Gene expression analysis provides important information regarding the function of the identified genes. The expression profiles of TaFADs have been investigated under various biotic and abiotic stresses. The highest expression in all developmental stages and tissues was related to TaFAB2.5, TaFAB2.12, TaFAB2.15, TaFAB2.17, TaFAB2.20, TaFAD2.1, TaFAD2.6, and TaFAD2.8. The expression pattern of TaFAD2.8 and TaFAB2.5 was similar to high frequency in all tissues except shoots/roots at the growth cycle with moderate abundance. Sever changes in the expression of the TaFAB.13 gene in all of the developmental stages have been observed. The expression of this gene was low in all cycles and tissues, while the moderate and high levels of transcripts have been observed in shoots/root at the growth cycle and roots at vegetative and seedling stages, respectively. Other FAB2 subfamily members showed high level of expression in all tissues except TaFAB2.1–3, TaFAB2.6–7, TaFAB2.9–11, TaFAB2.14, TaFAB2.16, TaFAB2.18–19, and TaFAB2.21–34. The expression pattern of TaFAB2.4 of the FAB2 subfamily was similar to TaFAD6.1–2 and TaFAD8.1–3 genes. They revealed drastic changes in gene expression as high expression at all phases of shoots/leaves and reproductive phases of spikes, but low expression in roots and grains tissues. The expression patterns of the TaFAD7.1–3 gene were almost similar with high transcripts level in roots (at all developmental cycles), grains, and spikes (at vegetative phase). These findings contradict the results of Nishii et al. (1997). They reported no expression of AtFAD7 genes in Arabidopsis roots [80]. However, our results are in agreement with Chi et al. (2011) findings. Based on their study, GmFAD7 had expression in all tissues of soybean [34]. TaFAD8.1–3 genes also demonstrated the same expression patterns, and they had the highest expression in shoots and leaves. These results are inconsistent with those of Soria-García et al. (2019). They detected AtFAD8 in Arabidopsis leaves, not roots [81]. In the case of TaFAD3.1–3 genes, the expression was relatively low in all tissues, whereas the expression level of TaFAD3.4–6 was highly abundant in all tissues at the growth stage and in spike at the reproductive phase. The expression of TaFAD2.1, TaFAD2.6, and TaFAD2.8 genes was high in all tissues while the expression of
TaFAD2.2–5, TaFAD2.7, and TaFAD2.9–11 was found in low abundance in all tissues and phases. The expression patterns of duplicated paired genes was mostly similar except TaFAD2.5/TaFAD2.6, TaFAD2.8/TaFAD2.9, TaFAD2.1/TaFAD2.2, TaFAD2.1/TaFAD2.3, TaFAD2.6/TaFAD2.7, TaFAD2.4/TaFAD2.6, TaFAD2.8/TaFAD2.10, and TaFAD2.8/TaFAD2.11. Different expression patterns of TaFAD genes under biotic and abiotic stresses indicating their different functions. The expression patterns of the TaFAB2 subfamily were different under biotic stress (Fig. 8). For instance, the TaFAB2.15 and TaFAB2.17 constantly expressed at high levels, whereas other FAB2 genes showed a minimum expression except for TaFAB2.12 and TaFAB2.20 with moderate to high expressions. Therefore, they have different roles in plant response to pathogen attacks. The expression of TaFAD8.1–3 has been induced in response to pathogen infection as well. Omega 3 fatty acid desaturases convert dienoic fatty acids to trienoic fatty acids, including linolenic acid, which is subsequently converted to jasmonic acid (JA) [82]. Accumulation of JA is essential in plant response to pathogen infection [82]. The expression of TaFAD2.1, TaFAD2.6, and TaFAD2.8 was also increased during biotic stress. The FAD2 genes appear to be involved in response to pathogen attack through increasing biosynthesis of linoleic acid and palmito-linoleic acid in olive [83]. Temperature stress can affect plant yield [16]. Based on the other studies, several members of the FAD gene family play important roles in response to temperature stress in various plant species [16, 43, 84–86]. According to the Fig. 9a, the expression levels of TaFAD2.1–2, TaFAD2.6, TaFAD2.8, TaFAB2.5, TaFAB2.8, TaFAB2.12, TaFAB2.15, TaFAB2.17, and TaFAB2.20 was increased after 6 h of heat treatment relative to their expression after 1 h of heat stress while the expression of other TaFADs was almost constant. Likewise, the expression of TaFAD2.6, TaFAD2.8, TaFAB2.15, TaFAB2.17, and TaFAB2.20 was high during 2 weeks of cold treatment. The results suggest that TaFAD2.6, TaFAD2.8, TaFAB2.15, TaFAB2.17, and TaFAB2.20 might play an important role in the response of wheat plants to temperature stress. Furthermore, the expression of TaFAD6.1–2 and TaFAD8.1–3 was high in response to cold stress. These results are in line with the results reported by Wang et al. (2006). They ascertained that the over-expression of FAD8 in rice resulted in a reduction in the cold stress damage [87]. Along with the results reported by other researchers, the expression of most TaFAD genes was more related to low temperature than to heat stress response [16, 86, 88, 89]. The expression level of the TaFAD2.8 under low temperature is much higher than that of the other TaFAD genes, suggesting that this gene may be related to enhancing cold tolerance in wheat. The different expression patterns of the duplicated genes illustrated the theory of divergence that these duplicated genes could be the result of two mechanisms; 1- subfunctionalization, and 2- neofunctionalization. In the subfunctionalization mechanism, some of the functional aspects of new genes are different from the function of the parental genes [90]. However, in the neofunctionalization process, the new gene has a different role due to changes in its amino acid composition compared with the parental gene [91]. Taken together, the results suggest that members of the wheat FAD gene family may maintain their main functions or obtain different roles during evolution. The use of accurate expression assessment methods such as northern blot and qRT-PCR are necessary to determine the exact expression patterns of the wheat FAD gene family at various tissues and developmental stages under biotic and abiotic stresses.

**Conclusions**

In recent years, using bioinformatics tools and computational biology methods is essential to clarify complicated biological problems in genomics, proteomics as well as in metabolomics. Therefore, researchers can identify various stress-related elements responding to environmental changes and subsequently develop improved crop plants in terms of quality and productivity, showing enhanced tolerance against biotic and abiotic stresses. On the other hand, fatty acid desaturases have essential roles in plant development and response to various stresses. Therefore, in the present study, 68 TaFAD genes were identified using bioinformatics approaches that they revealed 0 to 10 introns with high structural diversity. Some of TaFAD genes were intronless, and probably their genesis was based on the horizontal gene transfer from prokaryotes or retrotransposons. Based on the phylogenetic study, TaFAD genes were divided into five groups, including FAB2, FAD2/FAD6, FAD4, DES/SLD, and FAD3/FAD7/FAD8. The analysis of the mechanism of gene family expansion revealed that tandem and segmental duplications have occurred. Based on the results of the Ka/Ks ratio, the function of most of the duplicated TaFAD genes has been maintained during the evolution due to negative selection. Promoter analysis showed hormones and stresses-responsive elements in the TaFADs promoters suggesting the role of them in plant development and responses to biotic and abiotic stresses. Likewise, the TaFAD gene expression patterns indicated their functional roles in different tissues and developmental stages under environmental stresses in wheat. The expression patterns of genes resulting from tandem duplication suggested new functional roles for some of the duplicated genes. In addition, several residues in the active site of TaFAD2.6 and TaFAD2.8 in close contact with the docked oleic acid were found.
based on docking simulations which could be useful in future site-directed mutagenesis studies to improve the catalytic efficiency of the FAD2 enzymes and thereby to enhance the resistance of wheat to different stresses. This study was the first step in understanding the potential role of the TaFAD genes and could be useful in future studies on the specific role of each of the TaFAD genes at different developmental stages and responses to biotic and abiotic stresses.

**Methods**

**In silico identification of FAD genes**

FAB enzymes have the functional FA_desaturase 2 domain, while the FAD enzymes have the FA_desaturase or TMEM189 domains. Therefore, to determine the FAD gene family in wheat, the HMM profile of FA_desaturase (PF00487), FA_desaturase 2 (PF03405) or TMEM189 (PF10520) domains were obtained from Pfam database [92], and hmmSearch tools in HMMER server [93] has been used to find wheat FAD proteins in Ensembl Plants database [94]. Default parameters including significance E-values 0.01 for sequence, 0.03 for hit matches, and reporting E-values 1 for both sequence and hit. Molecular weight, length, and theoretical isoelectric points of wheat FAD were calculated using the ProtParam tool of the ExPASY Bioinformatics Resource Portal [95]. To recognize the cellular localization of proteins, CELLO, and DeepLoc have been applied [96, 97].

**Evolutionary relationships of wheat FAD gene family**

We used ClustalX 2.0.8 software to investigate the evolutionary relationships of the FAD gene family using full-length protein sequence alignment of wheat, Arabidopsis, rice, and soybean (Additional file 2). A phylogenetic tree of FAD proteins was constructed using MEGA 7 [98] based on Neighbor-joining (NJ) method with 1000 bootstraps [99].

**Duplication of FAD members and selection pressure**

The location image of FAD genes on the wheat chromosomes has been generated using TBtools [100]. Genes on the same chromosome with a maximum spacing of 10 genes have been considered as tandem duplication [101]. Likewise, two criteria have been considered to identify segmental duplication, including the identity of the aligned region more than 90% and the coverage of alignment more than 90% compared with longer genes [102]. Also, we computed the type of selection pressure on the tandem and segmental duplication genes, substitution rates of synonymous (Ks) and non-synonymous (Ka) by DnaSP ver. 5 software [103]. Then, the type of selection on genes has been determined using the Ka/Ks rate. Divergence time (T) of the duplicated genes has been estimated using the $T = \frac{Ks}{2\lambda}$ (MYA) formula that $\lambda$ is $6.5 \times 10^{-9}$ in wheat [104].

**Investigation of exon-intron structure and conserved motifs**

Exon-intron distributions and the type of splicing phase for wheat FAD genes have been appraised using a gene structure display server (GSDS 2.0) [105]. To find specific motifs of the FAD gene family, Multiple Em for Motif Elicitation (MEME 5.0.5) was used [106]. Various parameters, including a minimum length of motifs (6 amino acid) and a maximum length of motifs (200 amino acid), have been considered. Identification of 10 and 30 motifs have been taken into account for FAB2 and FAD subfamilies, respectively. Pfam and SMART databases have been used to evaluate the function of the motifs above [107, 108].

**Promoter analysis and prediction of simple sequence repeats (SSR) markers and miRNAs targets**

1500 bp upstream of starting codon (ATG) of FAD genes were obtained from Ensemble Plants database [109], and identification of cis-regulatory elements was carried out using PlantCare [110]. SSR markers have been identified in TaFAD gene sequences using the BatchPrimer3v1.0 server [111]. To find TaFAD-targeted miRNAs, CDS sequences of them were analyzed in the psRNATarget database by considering default parameters.

**Digital gene expression analysis**

To assess the expression profile of the wheat FAD genes, wheat gene expression data in the expVIP database [112] was applied, including RNA-seq data related to reproductive and vegetative shoots, leaf, root, spike, and grain tissues. Likewise, the expression profile of TaFADs under different stresses, including drought, cold, heat, and pathogen infection, has been evaluated. The obtained log10 (TPM + 1) value related to FAD genes expression was used, and heat map via an Amazing Heatmap in TBtools software [100].

**TaFAD2.6 and TaFAD2.8 structural modeling and validation**

In situ full-length atomic structure of TaFAD2.6 and TaFAD2.8 proteins were constructed by iterative template-based fragment assembly simulations to predict protein structures in the I-TASSER server [113]. The best models from I-TASSER were further refined by ModRefinder software [114]. The predicted structures were then validated via Ramachandran plot by measuring the backbone dihedral phi (ϕ) and psi (Ψ) angles using the PROCHECK module of the PDBSum server [115], and
RAMPAGE server has been applied for further confirmation [116]. The transmembrane α-helices of the Tafa D2.6 and TaFAD2.8 has been predicted using a CCTOP program (http://cctop.enzim.ttk.mta.hu/).

Molecular docking
The structure of oleic acid ligand has been obtained from the PubChem database [117] and converted into PDB format using Discovery Studio software. An enhanced version of the COACH server (COACH-D) has been used to predict protein-ligand binding site [118]. The server, as mentioned above, uses five methods to predict protein-ligand binding sites; four methods are template-based, including TM-SITE [119], S-SITE [119], COFACTOR [120], and FINDSITE [121] whereas the latter method (ConCavity) is structure-based [122]. Then, the results of each method have been combined using the COACH algorithm [119]. To analyze ligand-enzyme interaction, AutoDock v4.2.6 has been applied [123]. To prepare grid maps, the Auto Grid program developed with AutoDock has been used. The grid box size for x, y, and z was set at 60, 60, and 60 Å, respectively. The grid center for x, y, and z was set at 63.065, 63.843, and 63.387 Å, respectively with a grid spacing of 0.375 Å. To find the best conformers, Lamarckian Genetic Algorithm (LGA) has been selected. For ligand a limit of 100 conformers was considered during the docking process. Most of docking processes were carried out with AutoDock4’s default parameters [123]. Population size was set at 150, the maximum number of tests at 2,500,000, the maximum number of generations at 27,000, the maximum number of automatically surviving top individuals 1, gene mutation rate 0.02 and crossover rate 0.8. The interaction of the enzymes and substrates has been demonstrated in 2D and 3D using Discovery Studio Visualizer and Chimera software (Available on https://github.com/Amin62123/Chimera/tree/TaFAD) [124], respectively.

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12864-020-07199-1.

Additional file 1: Table S1 Features of wheat ATG proteins. Table S2 Ka/Ks analysis of the TaFAD duplicated paired genes. Table S3 Putative conserved motifs of the TaFAD2 subfamily. Table S4 Putative conserved motifs of the TaFAD subfamily. Table S5 Conserved histidine-rich boxes of FAD in wheat. Table S6 List of cis-acting regulatory elements in the TaFAD promoter. Table S7 Simple sequence repeats detected in TaFAD genes. Table S8 Putative TaFAD-targeted miRNA. Table S9 Gene expression data of TaFADs.

Additional file 2. The full-length protein sequence of wheat (Ta), Arabidopsis (At), rice (Os), and soybean (Gm).

Abbreviations
ABA: Abscisic acid; Chr: Chromosome; FAD: Fatty acid desaturase; JA: Jasmonic acid; LGA: Lamarckian Genetic Algorithm; MW: Molecular weight; pI: Isoelectric point; SSRs: Simple sequence repeats; TaFAD: Wheat fatty acid desaturase

Acknowledgements
We thank all researchers that helped us to improve the quality of this study.

Authors’ contributions
This research was designed and wrote by Z.H, and A.A. In addition, H.W, W.S, H.R, and Q.Z. reviewed and confirmed. This study was supervised and funded by A.M. All authors have read and approved the manuscript.

Funding
This work was supported by the foundation of Nanjing Forestry University (163108059). This funding body had no role in the experimental design, sample collection, data analysis and interpretation, and manuscript writing.

Availability of data and materials
All data generated or analyzed during this study are included in this published article and its supplementary information files.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no conflict of interest.

Author details
1Department of Agricultural Biotechnology, Faculty of Agricultural Sciences, University of Guilan, Rasht 4199613776, Iran. 2-Co-Innovation Center for Sustainable Forestry in Southern China, Key Laboratory of Forest Genetics & Biotechnology, Ministry of Education, Nanjing Forestry University, Nanjing 210037, China.

Received: 27 April 2020 Accepted: 27 October 2020

Reference
1. Wallis JG, Watts JL, Browse J. Polyunsaturated fatty acid synthesis: what will they think of next? Trends Biochem Sci. 2002;27(9):467–73.
2. Li M-J, Wang X-J, Su L, Bi Y-P, Wan S-B. Characterization of five putative acyl carrier protein (ACP) isoforms from developing seeds of Arabis hypogaea L. Plant Mol Biol Report. 2010;28(3):365–72.
3. Sperling P, Temes P, Zank T, Heine Z. The evolution of desaturases. Prostaglandins Leukot Essent Fat Acids. 2003;68(2):73–95.
4. Luo T, Deng WY, Zeng J, Zhang FL. Cloning and characterization of a stearyl-acyl carrier protein desaturase gene from Cinnamomum longepaniculatum. Plant Mol Biol Report. 2009;27(1):13.
5. Singh SC, Sinha RP, Hader D-P. Role of lipids and fatty acids in stress tolerance in cyanobacteria. Acta Protozool. 2002;41(4):297–308.
6. Ohirrogje J, Browse J. Lipid biosynthesis. Plant Cell. 1995;7(9):957.
7. Díaz MI, Cuppari S, Soreli D, Carrera A. In silico analysis of fatty acid Desaturase genes and proteins in grasses. Appl Biochem Biotechnol. 2018; 184(2):484–99.
8. Sharma A, Chauhan RS. In silico identification and comparative genomics of candidate genes involved in biosynthesis and accumulation of seed oil in plants. Comp Func Genom. 2012;2012:914843.
9. Mikulíneni V, Rocheford T. Sequence variation and genomic organization of fatty acid desaturase-2 (fad2) and fatty acid desaturase-6 (fad6) cDNAs in maize. Theor Appl Genet. 2003;106(7):1326–32.
10. Celik Altunoglu Y, Uzel NM, Baloglu MC, Ulu F, Can TH, Cetinkaya R. 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. Biotechnol Biotechnol Equip. 2018;32(4):846–57.
11. Matteucci M, Angelis S, Enrico S, Lanamna R, Perrotta G, Altamura M. Cold affects the transcription of fatty acid desaturases and oil quality in the fruit of Olea europea L. genotypes with different cold hardiness. J Exp Bot. 2011; 62(10):3403–20.
12. Xu L, Zeng W, Li J, Liu H, Yan G, Si P, et al. Characteristics of membrane-bound fatty acid desaturase (FAD) genes in Brassica napus L. and their
expressions under different cadmium and salinity stresses. Environ Exp Bot. 2019;162:144–56.

13. Xue Y, Zhang X, Wang R, Chen B, Jiang J, Win AN, et al. Cloning and expression of Perilla frutescens FAD2 gene and polymorphism analysis among cultivars. Acta Physiol Plant. 2017;39(3):394.

14. Dar AA, Choudhury AR, Kanchana PK, Arumugam N. The FAD2 gene in plants: occurrence, regulation, and role. Front Plant Sci. 2017;8:1789.

15. Feng J, Dong Y, Liu W, He Q, Daud MK, Chen J, Zhu S. Characterization of 19 genes encoding membrane-bound fatty acid desaturases and their expression profiles in Gossypium hirsutum and their expressions during abiotic stress. Sci Rep. 2017;7:45711.

16. Liu W, Li W, He Q, Daud MK, Chen J, Zhu S. Characterization of 19 genes encoding membrane-bound fatty acid desaturases and their expression profiles in Gossypium rosinndii at low temperature. PLoS One. 2015;10(4):e0123281.

17. Zhao X, Wei J, Lei H, Zhang Y, Zhao Y, Xu X, et al. Identification of fatty acid Desaturases in maize and their differential responses to low and high temperature. Genes. 2019;10(6):445.

18. Byfield G, Xue H, Upchurch R. Two genes from soybean encoding soluble Δ9 stearoyl-ACP desaturases. Crop Sci. 2006;46(2):890–6.

19. Wang H, Cao F, Zhang W, Wang G, Yu W. Cloning and expression of Stearoyl-ACP Desaturase and two Oleate Desaturases genes from Ginkgo biloba L. Plant Mol Biol Report. 2013;33(3):633–48.

20. Hernández ML, Mancha M, Martínez-Rivas JM. Molecular cloning and characterization of genes encoding two microsomal oleate desaturases (FAD2) from olive. Phytochemistry. 2005;66(12):1417–26.

21. Zhang J, Liu H, Sun J, Li B, Zhu Q, Chen S, et al. Arabidopsis fatty acid desaturase FAD2 is required for salt tolerance during seed germination and early seedling growth. PLoS One. 2012;7(11):e013055.

22. Zhang J-T, Zhu JQ, Zhu Q, Liu H, Gao X-S, Zhang X-H. Fatty acid desaturase-6 (Fad6) is required for salt tolerance in Arabidopsis thaliana. Biochem Biophys Res Commun. 2009;390(3):649–74.

23. Zhang M, Kang R, Yim M, Gueta-Dahan Y, Salts Y, et al. 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. J Plant Sci. 2005;44(3):361–71.

24. Ghafoor K, Ozcan MM, AL-Juhaib AJ. Characterization of the fatty acid desaturase gene in wheat. Mol Gen Genomics. 2003;270(4):315–24.

25. Dong C-J, Cao N, Zhang Z-G, Shang Q-M. Characterization of the fatty acid desaturase genes in cucumber: structure, phylogeny, and expression patterns in cucumber. BMC Plant Biol. 2015;15(1):209.

26. PMG N, Kang I-S, Moon B-Y, Lee C-H. Effects of low temperature stress on rice (Oryza sativa L.) plastid ω-3 desaturase gene, OsFAD8 and its functional analysis using T-DNA mutants. Plant Cell Tissue Organ Cult (PCTOC). 2009;98(1):87–96.

27. Dyer JM, Mullen RT. Immunocytological localization of two plant fatty acid desaturases in the endoplasmic reticulum. FEBS Lett. 2001;494(1–2):44–7.

28. Liu W, Xian Z, Kang X, Tang N, Li Z. Genome-wide identification, phylogeny and expression analysis of GRAS gene family in tomato. BMC Plant Biol. 2015;15(1):209.

29. Dong C-J, Shang Q-M. Genome-wide characterization of phenylalanine ammonia-lyase gene family in watermelon (Citrus lanatus). Planta. 2013;238(1):35–49.

30. Qin Z, Wang Y, Zhang J, Shi W, Zhang J. Genome wide identification and expression analysis of the MADS-box gene family in Medicago truncatula. Biochem Biophys Res Commun. 2018;499(2):361–7.

31. Johnson DA, Thomas MA. The monosaccharide transporter gene family in Arabidopsis and rice: a history of duplications, adaptive evolution, and functional divergence. Mol Biol Evol. 2007;24(11):2412–23.

32. Zhang Z, Wei X, Liu W, Min X, Jin X, Niyama Zaba B, et al. Genome-wide identification and expression analysis of the fatty acid desaturase genes in Medicago truncatula. Biochem Biophys Res Commun. 2018;499(2):361–7.

33. Wang H, Cao F, Zhang W, Wang G, Yu W. Cloning and expression of Stearoyl-ACP Desaturase and two Oleate Desaturases genes from Ginkgo biloba L. Plant Mol Biol Report. 2013;33(3):633–48.

34. Lu Y, Guo C, Shan H, Kong H. Divergence of duplicate genes in exon–intron structure. Proc Natl Acad Sci. 2012;109(4):1187–92.

35. Wang Z, Qiao Y, Zhang J, Shi W, Zhang J. Genome-wide identification of small RNA and degradome sequencing. Gene. 2017;619:61.
83. Hernández ML, Padilla MN, Sicardo MD, Mancha M, Martínez-Rivas JM. Effect of different environmental stresses on the expression of oleate desaturase genes and fatty acid composition in olive fruit. Phytochemistry. 2011;72(2–3):178–87.

84. Chen M, Markham JE, Caihoon EB. Sphingolipid Δ8 unsaturation is important for glucosylceramide biosynthesis and low-temperature performance in Arabidopsis. Plant J. 2012;69(5):769–81.

85. Iba K. Acclimative response to temperature stress in higher plants: approaches of gene engineering for temperature tolerance. Annu Rev Plant Biol. 2002;53(1):225–45.

86. Román Â, Andreu V, Hernández ML, Lagunas B, Picorel R, Martínez-Rivas JM, et al. Contribution of the different omega-3 fatty acid desaturase genes to the cold response in soybean. J Exp Bot. 2012;63(13):4973–82.

87. Wang J, Ming P, Pittman J, Han Y, Yu J, Guo B, et al. Characterization of a rice (Oryza sativa L.) gene encoding a temperature-dependent chloroplast ω-3 fatty acid desaturase. Biochem Biophys Res Commun. 2006;340(4):1209–16.

88. Wang HS, Yu C, Tang XF, Wang LY, Dong XC, Meng QW. Antisense-mediated depletion of tomato endoplasmic reticulum omega-3 fatty acid desaturase enhances thermal tolerance. J Integr Plant Biol. 2010;52(6):568–77.

89. Yurchenko OP, Park S, Iltut DC, Irnon JJ, Milholton JC, Lichty Z, et al. Genome-wide analysis of the omega-3 fatty acid desaturase gene family in Gossypium. BMC Plant Biol. 2014;14(1):312.

90. He X, Zhang J. Rapid subfunctionalization accompanied by prolonged and substantial neofunctionalization in duplicate gene evolution. Genetics. 2005;169(2):1157–64.

91. Teshima KM, Innan H. Neofunctionalization of duplicated genes under the pressure of gene conversion. Genetics. 2008;178(3):1385–98.

92. Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, et al. The Pfam protein families database. Nucleic Acids Res. 2004;32(suppl_1):D138–41.

93. Finn RD, Clements J, Eddy SR, Bateman A, Durbin R, Hollich V, et al. HMmer web server: interactive sequence similarity searching. Nucleic Acids Res. 2006;34(2):78–9.

94. Bolser D, Staines DM, Pritchard E, Kersey P. Ensembl plants: integrating tools for visualizing, mining, and analyzing plant genomics data. Plant Bioinform. 2016;11:115–40.

95. Antimo P, Jonnalagadda M, Arnold K, Baratin D, Ciadri G, De Castro E, et al. The Pfam protein families database. Nucleic Acids Res. 2012;40(4):4940.

96. Almagro Armenteros J, Sanderby CK, Sanderby SK, Nielsen H, OJB W. DeepLoc: prediction of protein subcellular localization using deep learning. Bioinformatics. 2017;33(2):3387–95.

97. Yu CS, Chen YC, Lu CH, Hwang JKJP. Function, bioinformatics. Prediction of protein subcellular localization. Proteins. 2006;64(3):643–5.

98. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016;33:1870–80.

99. Felsenstein JJE. Confidence limits on phylogenies: an approach using the bootstrap. Evolution. 1985;39(4):783–91.

100. Chen C, Chen H, He Y, Xia R. TBtools, a toolkit for biologists integrating various biological data handling tools with a user-friendly interface. BioReV. 2018;1:289660.

101. Wei K, Pan S, Li Y. Functional characterization of maize C 2 H 2 zinc-finger gene family. Plant Mol Biol Report. 2016;34(4):761–76.

102. Wu C, Ding X, Ding Z, Tie W, Yan Y, Wang Y, et al. The class III peroxidase (POD) gene family in cassava: identification, phylogeny, duplication, and expression. J Int J Mol Sci. 2019;20(11):2730.

103. Hu B, Jin J, Guo A-Y, Zhang H, Luo J, Gao G. GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics. 2014;30(23):4141–7.

104. Balloy T, Contreras-Moreira B, De Silva N, Maslen G, Akaneni W, Allen J, et al. Ensembl genomes 2020—enabling non-vertebrate genomic research. Nucleic Acids Res. 2019;47(D1):D689–D695.
110. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, et al. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res. 2002;30(1):325–7.

111. You FM, Huo N, Gu YQ, Luo M-c, Ma Y, Hane D, et al. BatchPrimer3: a high throughput web application for PCR and sequencing primer design. BMC bioinformatics. 2008;9(1):253.

112. Borrell P, Ramirez-Gonzalez R, Ualay C. expVIP: a customizable RNA-seq data analysis and visualization platform. Plant Physiol. 2016;170(4):2172–86.

113. Zhang Y, RAK, I. TASSER: A unified platform for automated protein structure and function prediction. Nature Protocols. 2010;5(4):725.

114. Xu D, Zhang Y. Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. Biophys J. 2011;101(10):2525–34.

115. Laskowski RA, Chistyakov V, Thornton JM. PDBsum more: new summaries and analyses of the known 3D structures of proteins and nucleic acids. Nucleic Acids Res. 2005;33(suppl_1):D266–D8.

116. Lovell SC, Davis IW, Arendall WB III, De Bakker PI, Word JM, Prisant MG, et al. Structure validation by Cα geometry: ϕ, ψ and Cβ deviation. Proteins. 2003;50(3):437–50.

117. Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, et al. PubChem 2019 update: improved access to chemical data. Nucleic Acids Res. 2019;47(D1):D1102–D8.

118. Wu Q, Peng Z, Zhang Y, Yang J. COACH-D: improved protein–ligand binding site prediction with refined ligand-binding poses through molecular docking. Nucleic Acids Res. 2018;46(W1):W438–W42.

119. Yang J, Roy A, Zhang Y. Protein–ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. Bioinformatics. 2013;29(20):2388–95.

120. Roy A, Yang J, Zhang Y. COFACTOR: an accurate comparative algorithm for structure-based protein function annotation. Nucleic Acids Res. 2012;40(W1):W471–W7.

121. Brylinski M, Skolnick J. A threading-based method (FINDSITE) for ligand-binding site prediction and functional annotation. Proc Natl Acad Sci. 2008;105(1):129–34.

122. Capra JA, Laskowski RA, Thornton JM, Singh M, Funkhouser TA. Predicting protein ligand binding sites by combining evolutionary sequence conservation and 3D structure. PLoS Comput Biol. 2009;5:12.

123. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. J Comput Chem. 2009;30(16):2785–91.

124. Petterson EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF chimera—a visualization system for exploratory research and analysis. J Comput Chem. 2004;25(13):1605–12.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:
• fast, convenient online submission
• thorough peer review by experienced researchers in your field
• rapid publication on acceptance
• support for research data, including large and complex data types
• gold Open Access which fosters wider collaboration and increased citations
• maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.
Learn more biomedcentral.com/submissions