Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Identification, Evolution, Expression, and Docking Studies of Fatty Acid Desaturase Genes in Wheat (*Triticum aestivum* L.)

Zahra Hajiahmadi
University of Guilan
Amin Abedi
University of Guilan
Hui Wei
Nanjing Forestry University
Weibo Sun
Nanjing Forestry University
Honghua Ruan
Nanjing Forestry University
Qiang Zhuge
Nanjing Forestry University
Ali Movahedi (ali_movahedi@njfu.edu.cn)
Nanjing Forestry University https://orcid.org/0000-0001-5062-504X

Research article

Keywords: Expression pattern, Gene Family, Molecular Docking, Protein structure

Posted Date: May 8th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-25679/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Version of Record: A version of this preprint was published at BMC Genomics on November 10th, 2020.
See the published version at https://doi.org/10.1186/s12864-020-07199-1.

Abstract

Backgrounds: Fatty acid desaturases (FADs) introduce a double bond into the fatty acids acyl chain resulting in unsaturated fatty acids which have important 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, genome-wide analysis of wheat *FAD* gene family (*TaFADs*) has been performed.

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 has been maintained during the evolution. Several cis-acting elements related to hormones and stresses in the *TaFAD*s 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 *TaFAD*s 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.

Background

Polyunsaturated 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 band 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 band into the acyl chain of saturated fatty acid in plastids. Membrane-bound FADs exist in chloroplast and endoplasmic reticulum (ER). Desaturation processes occur through two different pathways in the chloroplast and the ER [6]. In the chloroplast and ER, double bond formation requires NADPH/ferredoxin and NADH/cytochrome b_5 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 important 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 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 amount of transcripts of both genes and oleic acid has been dramatically increased in low temperature, while the expression of the aforementioned genes and the amount of C18:0 have been decreased and increased in high temperature, respectively [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 tobacco transgenic plants over-expressing *FAD3* or *FAD8* genes. According to their findings, the over-expression of *FAD8* or *FAD3* genes caused enhanced tolerance to drought [23]. Taken together, 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 important for the production of PUFAs in plants, thus, a comprehensive understanding of *FAD* genes using bioinformatics studies can be helpful to 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.

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 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

To investigate phylogenetic relationships of wheat FAD proteins, a phylogenetic tree has been constructed using the Neighbor-joining method. According to Fig. 1, 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 together. 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 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. 2). 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 ch4B and ch4D. 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, threes genes on chr2A (*TaFAB2.1-3*), chr2D (*TaFAB2.9-11*), chr5B (*TaFAB2.26-28*), and chr5A (*TaFAB2.23-25*) are the result of tandem duplication. To reveal selection pressure between duplicated genes, Ka, Ks, and Ka/Ks parameters have been investigated for 153 paired genes (Additional file 1: Table S2). 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*, *TaFADB2.1/TaFADB2.2, TaFADB2.9/TaFADB2.10*, *TaFADB2.15/TaFADB2.20*, and *TaFADB2.1/TaFADB2.11*) 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. 3). 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 *TaFAB2.4*. (Fig. 3).

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. 4 and 5). 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. Third His-box with consensus sequence HX_2HH exists in all TaFAD except in the TaSLD1-3 that the first amino acid is glutamine instead of histidine which is important for its enzymatic activity. All members of the *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 *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, leave, grain, and spike tissues (Fig. 6). 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* (moderate 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* (moderate 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 the all studied developmental stages, and spike at the reproductive phase. Based on RNA-seq data analysis of the *TaFAD4* 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. Although *TaFAD3.4-6* genes showed a high level of expression in all tissues at vegetative and spike 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 stages except seedling phases with moderate expressions. Finally, *TaFAB2.13* showed a low level of transcripts in all stages except roots at seedling and growth cycles with high expression and leaves/shoots at the growth cycle with moderate expression. To predict the role of *TaFAD* genes responding to biotic and abiotic stresses, their expression patterns have been investigated (Figs. 7 and 8). In response to powdery mildew pathogen, the up-regulated expression has been observed in *TaFAB2.4* after 24 hours of treatment while the expression of *TaFAD8.1-3*, *TaFAB2.1*, *TaFAB2.15*, and *TaFAB2.17* was high during the stress. On the other hand, the expression of the *TaFAD3.4*, *TaFAD3.5*, *TaFAD3.6*, *TaFAD3.8*, *TaSLD1-3*, and *TAFAB2.20* (moderate expression) remained unchanged and it has been down-regulated in *TaFAB2.5* and *TaFAB2.12* at 72 hours compared with 24 hours. 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 hours 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*, *TaFAB2.5*, *TaFAB2.8*, *TaFAB2.12*, *TaFAB2.15*, *TaFAB2.17*, and *TaFAB2.20* could be up-regulated by heat stress after 6 hours. 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 hours. 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*, *TaFAB2.12*, *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 6 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 68.4–86.0% and from 69.1–87.0% in TaFAD2.6, and TaFAD2.8, respectively, confirming that the refined models are of good qualities. The modeled structure for TaFAD2.6 has 18 α-helices and 19 loops (Fig. 9b) while TaFAD2.8 structure has 19 α-helices and 20 loops (Fig. 10b). Most of the predictor programs predicted that the TaFAD2.6 and TaFAD2.8 proteins have six transmembrane (TM) α-helices (Fig. 9a and 10a) 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 at position loop 7, His170, His173, His174 at positions α-helix 8, and His345, His348, His349 at position α-helix 16 of the protein structure (Fig. 9d). Likewise, the TaFAD2.8 structure possesses three conserved histidine boxes with eight histidine residues of His111, His115 at position α-helix 4, His147, His150, His151 at positions α-helix 6, and His322, His325, His326 at position α-helix 17 (Fig. 10d). The histidine residues revealed essential catalytic features in plant FADs [32].

Delta-12 desaturase introduces a double bond at the delta-12 position of oleic acid to produce linoleic acid. To assess the ligand specificity of wheat omega 6 desaturase, docking studies of oleic acid on the refined model structure have been carried out using AutoDock 4.2. Based on the docking simulation with the ligand-enzyme binding energy of -5.94 kcal/mol, Val129, His134, Leu160, Val161, Ser165, Trp166, Ser169, His170, His173, Pro188, Lys189, Gln190, Lys191, Glu192, Ala193, Ile212, Val213, Leu216, Leu296, Val297, Ile299, Thr300, and His345 of TaFDA2.6 formed closed contacts with the docked oleic acid, while Arg29, Val106, Ala110, His111, Trp143, Ser146, His147, Arg149, His150, His151, Asn153, Gln192, Ala202,

Pro211, Leu273, Ile276, Thr277, His281, His322, and His326 were involved in ligand-TaFAD2.8 binding with − 4.75 kcal/mol docking energy. The oleic acid formed a hydrogen band with His170 and His147 in TaFAD2.6 and TaFAD2.8, respectively (Fig. 9c and 10c). Based on the Fig. 9d, in TaFDA2.6, 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, His115, His147, His 150, 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 aforementioned results 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 trancatula* (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 an important 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/QX_2HH 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 important 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, gibberellin, and ethylene suggest that the *TaFAD* genes are involved in the 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 are able to 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/CGC/GCG/GCC/GGC 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 greater 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 key 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 the 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 the response to drought [75, 76]. miR5384 can regulate the two pathways of cell death and growth [77]. Finally, miR9772 and miR9678 are important 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. In fact, 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 Nishiuchi 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. 7). 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 [84]. Based on the other studies, several members of the *FAD* gene family play important roles in response to temperature stress in various plant species [84–88]. According to the Fig. 8a, 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 hours of heat treatment relative to their expression after 1 hour 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 [89]. 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 [84, 88, 90, 91]. 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 [92]. 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 [93]. 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

Taken collectively, 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 5 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 *TaFAD*s 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, a number of residues in the active site of TaFAD2.6 and TaFAD2.8 in close contact with the docked oleic acid were found on the basis of 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

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 [94] and HMMERserach tool [95] has been used to find wheat FAD proteins in Ensembl Plants database [96]. Molecular weight, length, and theoretical isoelectric points of wheat FAD were calculated using the ProtParam tool of the ExPASY database [97]. To recognize cellular localization of proteins, CELLO and DeepLoc have been applied [98, 99].

Evolutionary relationships of wheat FAD gene family

To investigate the evolutionary relationships of the *FAD* gene family, full-length protein sequence alignment of wheat, *Arabidopsis*, rice, and soybean has been conducted using ClustalX 2.0.8 software. A phylogenetic tree of FAD proteins was constructed using MEGA 7 [100] based on Neighbor-joining (NJ) method with 1000 bootstraps [101].

Duplication of FAD members and selection pressure

The location image of *FAD* genes on the wheat chromosomes has been generated using MapGene2Chrom 2 server [102]. Genes on the same chromosome with a maximum spacing of 10 genes have been considered as tandem duplication [103]. 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 [104]. To determine the type of selection pressure on the tandem and segmental duplication genes, substitution rates of synonymous (Ks) and non-

synonymous (Ka) were computed by DnaSP ver. 5 software [105]. 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 = Ks/2λ (MYA) formula that λ is 6.5×10^{-9} in wheat [106].

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) [107]. To find specific motives of the *FAD* gene family, Multiple Em for Motif Elicitation (MEME 5.0.5) was used [108]. Various parameters including a minimum length of motives (6 amino acid) and a maximum length of motives (200 amino acid) have been considered. Identification of 10 and 30 motives have been taken into account for *FAB2* and *FAD* subfamilies, respectively. Pfam and SMART databases have been used to evaluate the function of the aforementioned motives [109, 110].

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 [111] and identification of cis-regulatory elements was carried out using PlantCare [112].

SSR markers have been identified in *TaFAD* gene sequences using the BatchPrimer3v1.0 server [113]. 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 [114] was applied including RNA-seq data related to reproductive and vegetative shoot, 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 data related to *FAD* genes expression was used to perform Hierarchical clustering using Pearson correlation, the complete linkage algorithm, and heat map via an Amazing Heatmap in TBtools software [115].

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 I-TASSER server [116]. The best models from I-TASSER were further refined by ModRefinder software [117]. 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 [118] and RAMPAGE server has been applied for further confirmation [119]. The transmembrane α-helices of the TAFAD2.6 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 [120] 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 [121]. The aforementioned server uses five methods to predict protein-ligand binding sites, four methods are template-based including TM-SITE [122], S-SITE [122], COFACTOR [123], and FINDSITE [124] whereas the last method (ConCavity) is structure-based [125]. Then, the results of each method have been combined using the COACH algorithm [122]. To analyze ligand-enzyme interaction, AutoDock v4.2.6 has been applied [126]. 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 [126]. Population size was set at150, 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 [127], respectively.

Abbreviations

ABA: abscisic acid; Chr: choromosome; FAD: fatty acid desaturase; JA: jasmonic acid; LGA: Lamarckian Genetic Algorithm; MW: molecular weigh; pI: isoelectric point; SSRs: Simple sequence repeats ;TaFAD: wheat fatty acid desaturase

Declarations

Acknowledgments

This work was supported by the foundation of Nanjing Forestry University (163108059).

Authors' contributions

This research was designed and wrote by Z.H. and A.A. In addition, H.W., H.R., and Q.Z. reviewed and confirmed. This study was supervised and funded by A.M.

Funding

Not applicable.

Availability of data and materials

All data generated or analysed 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.

Additional information

Supplementary information is available in this paper.

References

1. Wallis JG, Watts JL, Browse J. Polyunsaturated fatty acid synthesis: what will they think of next? Trends in biochemical sciences. 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 Arachis hypogaea L. Plant molecular biology reporter. 2010;28(3):365-72.
3. Sperling P, Ternes P, Zank T, Heinz E. The evolution of desaturases. Prostaglandins, leukotrienes, and essential fatty acids. 2003;68(2):73-95.
4. Luo T, Deng WY, Zeng J, Zhang FL. Cloning and characterization of a stearoyl–acyl carrier protein desaturase gene from Cinnamomum longepaniculatum. Plant molecular biology reporter. 2009;27(1):13.
5. Singh SC, Sinha RP, Hader D-P. Role of lipids and fatty acids in stress tolerance in cyanobacteria. Acta protozoologica. 2002;41(4):297-308.
6. Ohlrogge J, Browse J. Lipid biosynthesis. The Plant Cell. 1995;7(7):957.
7. Díaz ML, Cuppari S, Soresi D, Carrera A. In Silico Analysis of Fatty Acid Desaturase Genes and Proteins in Grasses. Applied biochemistry and biotechnology. 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. Comparative and functional genomics. 2012;2012.
9. Mikkilineni V, Rocheford T. Sequence variation and genomic organization of fatty acid desaturase-2 (fad2) and fatty acid desaturase-6 (fad6) cDNAs in maize. Theoretical and applied genetics. 2003;106(7):1326-32.
10. Celik Altunoglu Y, Unel 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. Biotechnology & Biotechnological Equipment. 2018;32(4):846-57.

11. Matteucci M, D'angeli S, Errico S, Lamanna R, Perrotta G, Altamura M. Cold affects the transcription of fatty acid desaturases and oil quality in the fruit of Olea europaea L. genotypes with different cold hardiness. Journal of experimental botany. 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. Environmental and Experimental Botany. 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 physiologiae plantarum. 2017;39(3):84.
14. Dar AA, Choudhury AR, Kancharla PK, Arumugam N. The FAD2 gene in plants: occurrence, regulation, and role. Frontiers in plant science. 2017;8:1789.
15. Feng J, Dong Y, Liu W, He Q, Daud M, Chen J, et al. Genome-wide identification of membrane-bound fatty acid desaturase genes in Gossypium hirsutum and their expressions during abiotic stress. Scientific reports. 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 raimondii under low temperature. PloS one. 2015;10(4):e0123281.
17. Zhao X, Wei J, He L, 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 science. 2006;46(2):840-6.
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 molecular biology reporter. 2013;31(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(1):e30355.
22. Zhang J-T, Zhu J-Q, Zhu Q, Liu H, Gao X-S, Zhang H-X. Fatty acid desaturase-6 (Fad6) is required for salt tolerance in Arabidopsis thaliana. Biochemical and biophysical research communications. 2009;390(3):469-74.
23. Zhang M, Barg R, Yin M, Gueta-Dahan Y, Leikin-Frenkel A, 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. The Plant Journal. 2005;44(3):361-71.

24. Ghafoor K, Özcan MM, AL-Juhaımı F, Babıker EE, Sarker ZI, Ahmed IAM, et al. Nutritional composition, extraction, and utilization of wheat germ oil: A review. European Journal of Lipid Science and Technology. 2017;119(7):1600160.
25. Kong W, Gong Z, Zhong H, Zhang Y, Zhao G, Gautam M, et al. Expansion and evolutionary patterns of glycosyltransferase family 8 in Gramineae crop genomes and their expression under salt and cold stresses in Oryza sativa ssp. japonica. Biomolecules. 2019;9(5):188.
26. Zhang W, Wang S, Yu F, Tang J, Yu L, Wang H, et al. Genome-wide identification and expression profiling of sugar transporter protein (STP) family genes in cabbage (Brassica oleracea var. capitata L.) reveals their involvement in clubroot disease responses. Genes. 2019;10(1):71.
27. Sharp PA. Speculations on RNA splicing. Cell. 1981;23(3):643-6.
28. Miao X, Zhang L, Hu X, Nan S, Chen X, Fu H. Cloning and functional analysis of the FAD2 gene family from desert shrub Artemisia sphaerocephala. BMC plant biology. 2019;19(1):481.
29. Chodok P, Eiamsa-ard P, Cove DJ, Quatrano RS, Kaewsuwan S. Identification and functional characterization of two Δ 12-fatty acid desaturases associated with essential linoleic acid biosynthesis in Physcomitrella patens. Journal of industrial microbiology & biotechnology. 2013;40(8):901-13.
30. Rajwade AV, Joshi RS, Kadoo NY, Gupta VS. Sequence characterization and in silico structure prediction of fatty acid desaturases in linseed varieties with differential fatty acid composition. Journal of the Science of Food and Agriculture. 2016;96(15):4896-906.
31. Shanklin J, Whittle E, Fox BG. Eight histidine residues are catalytically essential in a membrane-associated iron enzyme, stearoyl-CoA desaturase, and are conserved in alkane hydroxylase and xylene monooxygenase. Biochemistry. 1994;33(43):12787-94.
32. Broadwater JA, Whittle E, Shanklin J. Desaturation and Hydroxylation residues 148 and 324 of arabidopsis fad2, in addition to substrate chain length, exert a major influence in partitioning of catalytic specificity. Journal of Biological Chemistry. 2002;277(18):15613-20.
33. Xue Y, Chen B, Wang R, Win AN, Li J, Chai Y. Genome-wide survey and characterization of fatty acid desaturase gene family in Brassica napus and its parental species. Applied biochemistry and biotechnology. 2018;184(2):582-98.
34. Chi X, Yang Q, Lu Y, Wang J, Zhang Q, Pan L, et al. Genome-wide analysis of fatty acid desaturases in soybean (Glycine max). Plant molecular biology reporter. 2011;29(4):769-83.
35. Zhang Z, Wei X, Liu W, Min X, Jin X, Ndayambaza B, et al. Genome-wide identification and expression analysis of the fatty acid desaturase genes in Medicago truncatula. Biochemical and biophysical research communications. 2018;499(2):361-7.
36. Johnson DA, Thomas MA. The monosaccharide transporter gene family in Arabidopsis and rice: a history of duplications, adaptive evolution, and functional divergence. Molecular biology and evolution. 2007;24(11):2412-23.
37. Ma J, Yang Y, Luo W, Yang C, Ding P, Liu Y, et al. Genome-wide identification and analysis of the MADS-box gene family in bread wheat (Triticum aestivum L.). PLoS One. 2017;12(7).

38. Cao J, Shi F. Evolution of the RALF gene family in plants: gene duplication and selection patterns. Evolutionary Bioinformatics. 2012;8:EBO. S9652.
39. Cannon SB, Mitra A, Baumgarten A, Young ND, May G. The roles of segmental and tandem gene duplication in the evolution of large gene families in Arabidopsis thaliana. BMC plant biology. 2004;4(1):10.
40. Xu G, Guo C, Shan H, Kong H. Divergence of duplicate genes in exon–intron structure. Proceedings of the National Academy of Sciences. 2012;109(4):1187-92.
41. Hanada K, Zou C, Lehti-Shiu MD, Shinozaki K, Shiu S-H. Importance of lineage-specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli. Plant physiology. 2008;148(2):993-1003.
42. Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, et al. The Sorghum bicolor genome and the diversification of grasses. Nature. 2009;457(7229):551.
43. 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. PLoS One. 2016;11(3):e0149917.
44. Nair PMG, Kang I-S, Moon B-Y, Lee C-H. Effects of low temperature stress on rice (Oryzasativa L.) plastid ω-3 desaturase gene, OsFAD8 and its functional analysis using T-DNA mutants. Plant Cell, Tissue and Organ Culture (PCTOC). 2009;98(1):87-96.
45. Dyer JM, Mullen RT. Immunocytological localization of two plant fatty acid desaturases in the endoplasmic reticulum. FEBS letters. 2001;494(1-2):44-7.
46. Huang W, Xian Z, Kang X, Tang N, Li Z. Genome-wide identification, phylogeny and expression analysis of GRAS gene family in tomato. BMC plant biology. 2015;15(1):209.
47. Lindemose S, O'Shea C, Jensen M, Skriver K. Structure, function and networks of transcription factors involved in abiotic stress responses. International journal of molecular sciences. 2013;14(3):5842-78.
48. Dong C-J, Shang Q-M. Genome-wide characterization of phenylalanine ammonia-lyase gene family in watermelon (Citrullus lanatus). Planta. 2013;238(1):35-49.
49. Cao J, Li M, Chen J, Liu P, Li Z. Effects of MeJA on Arabidopsis metabolome under endogenous JA deficiency. Scientific reports. 2016;6:37674.
50. Liu Z, Zhang M, Kong L, Lv Y, Zou M, Lu G, et al. Genome-wide identification, phylogeny, duplication, and expression analyses of two-component system genes in Chinese cabbage (Brassica rapa ssp. pekinensis). 2014;21(4):379-96.
51. Upchurch RG. Fatty acid unsaturation, mobilization, and regulation in the response of plants to stress. Biotechnology letters. 2008;30(6):967-77.
52. Haasl RJ, Payseur BA. Microsatellites as targets of natural selection. Molecular biology and evolution. 2012;30(2):285-98.
53. Li Y-C, Korol AB, Fahima T, Nevo E. Microsatellites within genes: structure, function, and evolution. Molecular biology and evolution. 2004;21(6):991-1007.

54. Gupta PK, Rustgi S, Sharma S, Singh R, Kumar N, Balyan H. Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. Molecular genetics and genomics. 2003;270(4):315-23.
55. Kumar A, Batra R, Gahlaut V, Gautam T, Kumar S, Sharma M, et al. Genome-wide identification and characterization of gene family for RWP-RK transcription factors in wheat (Triticum aestivum L.). PloS one. 2018;13(12).
56. Qin Z, Wang Y, Wang Q, Li A, Hou F, Zhang L. Evolution analysis of simple sequence repeats in plant genome. PloS one. 2015;10(12).
57. Wang Z, Qiao Y, Zhang J, Shi W, Zhang J. Genome wide identification of microRNAs involved in fatty acid and lipid metabolism of Brassica napus by small RNA and degradome sequencing. Gene. 2017;619:61-70.
58. Agharbaoui Z, Leclercq M, Remita MA, Badawi MA, Lord E, Houde M, et al. An integrative approach to identify hexaploid wheat miRNAome associated with development and tolerance to abiotic stress. BMC genomics. 2015;16(1):339.
59. Fileccia V, Bertolini E, Ruisi P, Giambalvo D, Frenda AS, Cannarozzi G, et al. Identification and characterization of durum wheat microRNAs in leaf and root tissues. Functional & integrative genomics. 2017;17(5):583-98.
60. Liu H, Able AJ, Able JA. Water-deficit stress-responsive microRNAs and their targets in four durum wheat genotypes. Functional & integrative genomics. 2017;17(2-3):237-51.
61. Zhao Y-Y, Guo C-J, Li X-J, Duan W-W, Ma C-Y, Chan H-M, et al. Characterization and expression pattern analysis of microRNAs in wheat under drought stress. Biologia plantarum. 2015;59(1):37-46.
62. Phillips JR, Dalmay T, Bartels D. The role of small RNAs in abiotic stress. FEBS letters. 2007;581(19):3592-7.
63. Liu S, Wang N, Zhang P, Cong B, Lin X, Wang S, et al. Next-generation sequencing-based transcriptome profiling analysis of Pohlia nutans reveals insight into the stress-relevant genes in Antarctic moss. Extremophiles. 2013;17(3):391-403.
64. Wang J-W, Wang L-J, Mao Y-B, Cai W-J, Xue H-W, Chen X-Y. Control of root cap formation by microRNA-targeted auxin response factors in Arabidopsis. The Plant Cell. 2005;17(8):2204-16.
65. Ma X, Xin Z, Wang Z, Yang Q, Guo S, Guo X, et al. Identification and comparative analysis of differentially expressed miRNAs in leaves of two wheat (Triticum aestivum L.) genotypes during dehydration stress. BMC Plant Biology. 2015;15(1):21.
66. Kantar M, Unver T, Budak H. Regulation of barley miRNAs upon dehydration stress correlated with target gene expression. Functional & integrative genomics. 2010;10(4):493-507.
67. Sun L, Sun G, Shi C, Sun D. Transcriptome analysis reveals new microRNAs-mediated pathway involved in anther development in male sterile wheat. BMC genomics. 2018;19(1):333.
68. Luo Y, Guo Z, Li L. Evolutionary conservation of microRNA regulatory programs in plant flower development. Developmental biology. 2013;380(2):133-44.

69. Nakashima K, Takasaki H, Mizoi J, Shinozaki K, Yamaguchi-Shinozaki K. NAC transcription factors in plant abiotic stress responses. Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms. 2012;1819(2):97-103.
70. Ma C, Burd S, Lers A. mi R 408 is involved in abiotic stress responses in A rabidopsis. The Plant Journal. 2015;84(1):169-87.
71. Hajyzadeh M, Turktas M, Khawar KM, Unver T. miR408 overexpression causes increased drought tolerance in chickpea. Gene. 2015;555(2):186-93.
72. Bai Q, Wang X, Chen X, Shi G, Liu Z, Guo C, et al. Wheat miRNA TaemiR408 acts as an essential mediator in plant tolerance to Pi deprivation and salt stress via modulating stress-associated physiological processes. Frontiers in plant science. 2018;9:499.
73. Liu Z, Wang X, Chen X, Shi G, Bai Q, Xiao K. TaMIR1139: a wheat miRNA responsive to Pi-starvation, acts a critical mediator in modulating plant tolerance to Pi deprivation. Plant cell reports. 2018;37(9):1293-309.
74. Han R, Jian C, Lv J, Yan Y, Chi Q, Li Z, et al. Identification and characterization of microRNAs in the flag leaf and developing seed of wheat (Triticum aestivum L.). BMC genomics. 2014;15(1):289.
75. Li T, Ma L, Geng Y, Hao C, Chen X, Zhang X. Small RNA and degradome sequencing reveal complex roles of miRNAs and their targets in developing wheat grains. PloS one. 2015;10(10).
76. Akdogan G, Tufekci ED, Uranbey S, Unver T. miRNA-based drought regulation in wheat. Functional & integrative genomics. 2016;16(3):221-33.
77. Li M, Liang Z, He S, Zeng Y, Jing Y, Fang W, et al. Genome-wide identification of leaf abscission associated microRNAs in sugarcane (Saccharum officinarum L.). BMC genomics. 2017;18(1):754.
78. Ravichandran S, Ragupathy R, Edwards T, Domaratzki M, Cloutier S. MicroRNA-guided regulation of heat stress response in wheat. BMC genomics. 2019;20(1):488.
79. Guo G, Liu X, Sun F, Cao J, Huo N, Wuda B, et al. Wheat miR9678 affects seed germination by generating phased siRNAs and modulating abscisic acid/gibberellin signaling. The Plant Cell. 2018;30(4):796-814.
80. Nishiuchi T, Hamada T, Kodama H, Iba K. Wounding changes the spatial expression pattern of the arabidopsis plastid omega-3 fatty acid desaturase gene (FAD7) through different signal transduction pathways. The Plant Cell. 1997;9(10):1701-12.
81. Soria-García Á, Rubio MC, Lagunas B, López-Gomollón S, Luján MdlÁ, Díaz-Guerra R, et al. Tissue Distribution and Specific Contribution of Arabidopsis FAD7 and FAD8 Plastid Desaturases to the JA- and ABA-Mediated Cold Stress or Defense Responses. Plant and Cell Physiology. 2019;60(5):1025-40.
82. Im YJ, Han O, Chung GC, Cho BH. Antisense expression of an Arabidopsis omega-3 fatty acid desaturase gene reduces salt/drought tolerance in transgenic tobacco plants. Molecules and cells. 2002;13(2):264-71.
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. 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 raimondii under low temperature. PloS one. 2015;10(4).
85. Chen M, Markham JE, Cahoon EB. Sphingolipid Δ8 unsaturation is important for glucosylceramide biosynthesis and low-temperature performance in Arabidopsis. The plant journal. 2012;69(5):769-81.
86. 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. PLoS One. 2016;11(3).
87. Iba K. Acclimative response to temperature stress in higher plants: approaches of gene engineering for temperature tolerance. Annual Review of Plant Biology. 2002;53(1):225-45.
88. Román Á, 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. Journal of experimental botany. 2012;63(13):4973-82.
89. Wang J, Ming F, Pittman J, Han Y, Hu J, Guo B, et al. Characterization of a rice (Oryza sativa L.) gene encoding a temperature-dependent chloroplast ω-3 fatty acid desaturase. Biochemical and biophysical research communications. 2006;340(4):1209-16.
90. 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. Journal of integrative plant biology. 2010;52(6):568-77.
91. Yurchenko OP, Park S, Ilut DC, Inmon JJ, Millhollon JC, Liechty Z, et al. Genome-wide analysis of the omega-3 fatty acid desaturase gene family in Gossypium. BMC plant biology. 2014;14(1):312.
92. He X, Zhang J. Rapid subfunctionalization accompanied by prolonged and substantial neofunctionalization in duplicate gene evolution. Genetics. 2005;169(2):1157-64.
93. Teshima KM, Innan H. Neofunctionalization of duplicated genes under the pressure of gene conversion. Genetics. 2008;178(3):1385-98.
94. Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, et al. The Pfam protein families database. Nucleic acids research. 2004;32(suppl_1):D138-D41.
95. Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity searching. Nucleic acids research. 2011;39(suppl_2):W29-W37.
96. Bolser D, Staines DM, Pritchard E, Kersey P. Ensembl plants: integrating tools for visualizing, mining, and analyzing plant genomics data. Plant bioinformatics: Springer; 2016. p. 115-40.
97. Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, De Castro E, et al. ExPASy: SIB bioinformatics resource portal. Nucleic acids research. 2012:gks400.
98. Almagro Armenteros JJ, Sønderby CK, Sønderby SK, Nielsen H, Winther OJB. DeepLoc: prediction of protein subcellular localization using deep learning. 2017;33(21):3387-95.
99. Yu CS, Chen YC, Lu CH, Hwang JKJPS, Function,, Bioinformatics. Prediction of protein subcellular localization. 2006;64(3):643-51.

100. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Molecular biology and evolution. 2016:msw054.
101. Felsenstein JJE. Confidence limits on phylogenies: an approach using the bootstrap. 1985;39(4):783-91.
102. Jiangtao C, Yingzhen K, Qian W, Yuhe S, Daping G, Jing L, et al. MapGene2Chrom, a tool to draw gene physical map based on Perl and SVG languages. 2015;37(1):91-7.
103. Wei K, Pan S, Li Y. Functional Characterization of Maize C 2 H 2 Zinc-Finger Gene Family. Plant molecular biology reporter. 2016;34(4):761-76.
104. 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. International journal of molecular sciences. 2019;20(11):2730.
105. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009;25(11):1451-2.
106. Malviya N, Jaiswal P, Yadav DJP, plants mbo. Genome-wide characterization of Nuclear Factor Y (NF-Y) gene family of sorghum [Sorghum bicolor (L.) Moench]: a bioinformatics approach. 2016;22(1):33-49.
107. Hu B, Jin J, Guo A-Y, Zhang H, Luo J, Gao G. GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics. 2014:btu817.
108. Bailey TL, Williams N, Misleh C, Li WW. MEME: discovering and analyzing DNA and protein sequence motifs. Nucleic acids research. 2006;34(suppl 2):W369-W73.
109. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. Pfam: the protein families database. Nucleic acids research. 2013:gkt1223.
110. Letunic I, Doerks T, Bork P. SMART 7: recent updates to the protein domain annotation resource. Nucleic acids research. 2012;40(D1):D302-D5.
111. Howe KL, Contreras-Moreira B, De Silva N, Maslen G, Akanni W, Allen J, et al. Ensembl Genomes 2020—enabling non-vertebrate genomic research. Nucleic acids research. 2019;48(D1):D689-D95.
112. 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 research. 2002;30(1):325-7.
113. 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.
114. Borrill P, Ramirez-Gonzalez R, Uauy C. expVIP: a customizable RNA-seq data analysis and visualization platform. Plant physiology. 2016;170(4):2172-86.
115. Chen C, Chen H, He Y, Xia R. TBtools, a toolkit for biologists integrating various biological data handling tools with a user-friendly interface. BioRxiv. 2018:289660.
116. RoyA K. ZhangY. I TASSER: Aunifiedplatform forautomated proteinstructureandfunction prediction. NatureProtocols. 2010;5(4):725.

117. Xu D, Zhang Y. Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. Biophysical journal. 2011;101(10):2525-34.
118. Laskowski RA, Chistyakov VV, Thornton JM. PDBsum more: new summaries and analyses of the known 3D structures of proteins and nucleic acids. Nucleic acids research. 2005;33(suppl_1):D266-D8.
119. Lovell SC, Davis IW, Arendall III WB, De Bakker PI, Word JM, Prisant MG, et al. Structure validation by Cα geometry: ϕ, ψ and Cβ deviation. Proteins: Structure, Function, and Bioinformatics. 2003;50(3):437-50.
120. Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, et al. PubChem 2019 update: improved access to chemical data. Nucleic acids research. 2019;47(D1):D1102-D9.
121. Wu Q, Peng Z, Zhang Y, Yang J. COACH-D: improved protein–ligand binding sites prediction with refined ligand-binding poses through molecular docking. Nucleic acids research. 2018;46(W1):W438-W42.
122. 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):2588-95.
123. Roy A, Yang J, Zhang Y. COFACTOR: an accurate comparative algorithm for structure-based protein function annotation. Nucleic acids research. 2012;40(W1):W471-W7.
124. Brylinski M, Skolnick J. A threading-based method (FINDSITE) for ligand-binding site prediction and functional annotation. Proceedings of the National Academy of sciences. 2008;105(1):129-34.
125. Capra JA, Laskowski RA, Thornton JM, Singh M, Funkhouser TA. Predicting protein ligand binding sites by combining evolutionary sequence conservation and 3D structure. PLoS computational biology. 2009;5(12).
126. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. Journal of computational chemistry. 2009;30(16):2785-91.
127. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera—a visualization system for exploratory research and analysis. Journal of computational chemistry. 2004;25(13):1605-12.

Additional Files

Table S1 Features of wheat FAD proteins. Ch chloroplast, Cy cytoplasmic, ER endoplasmic reticulum, M mitochondria, PM plasma membrane.

Table S2 Ka/Ks analysis of the *TaFAD* duplicated paired genes.

Table S3 Putative conserved motifs of *TaFAB2* subfamily.

Table S4 Putative conserved motifs of *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.

Figures

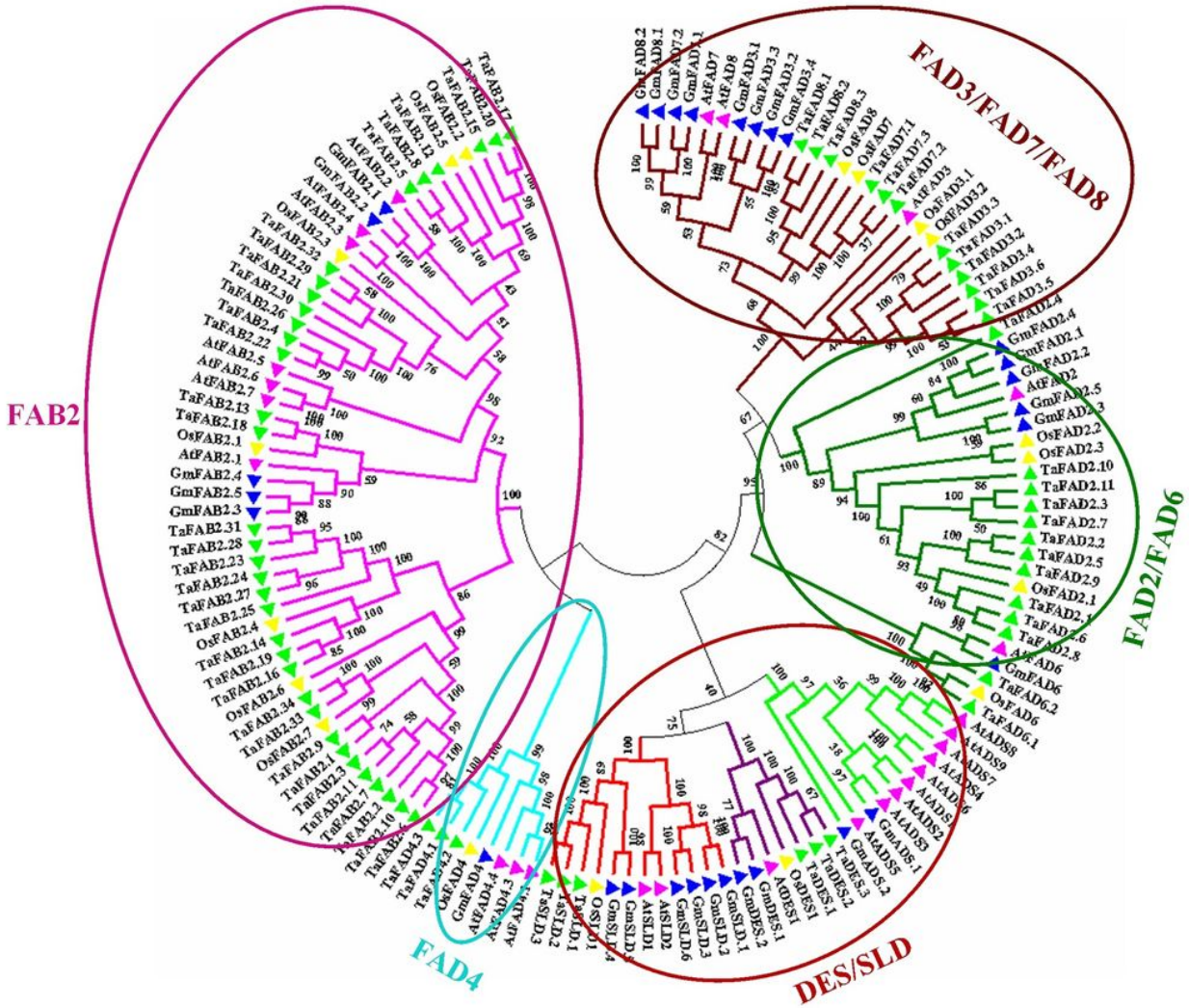

Figure 1

Phylogenetic relationships of FAD genes from wheat, Arabidopsis, rice, and soybean. The colored branch shows a different subfamily.

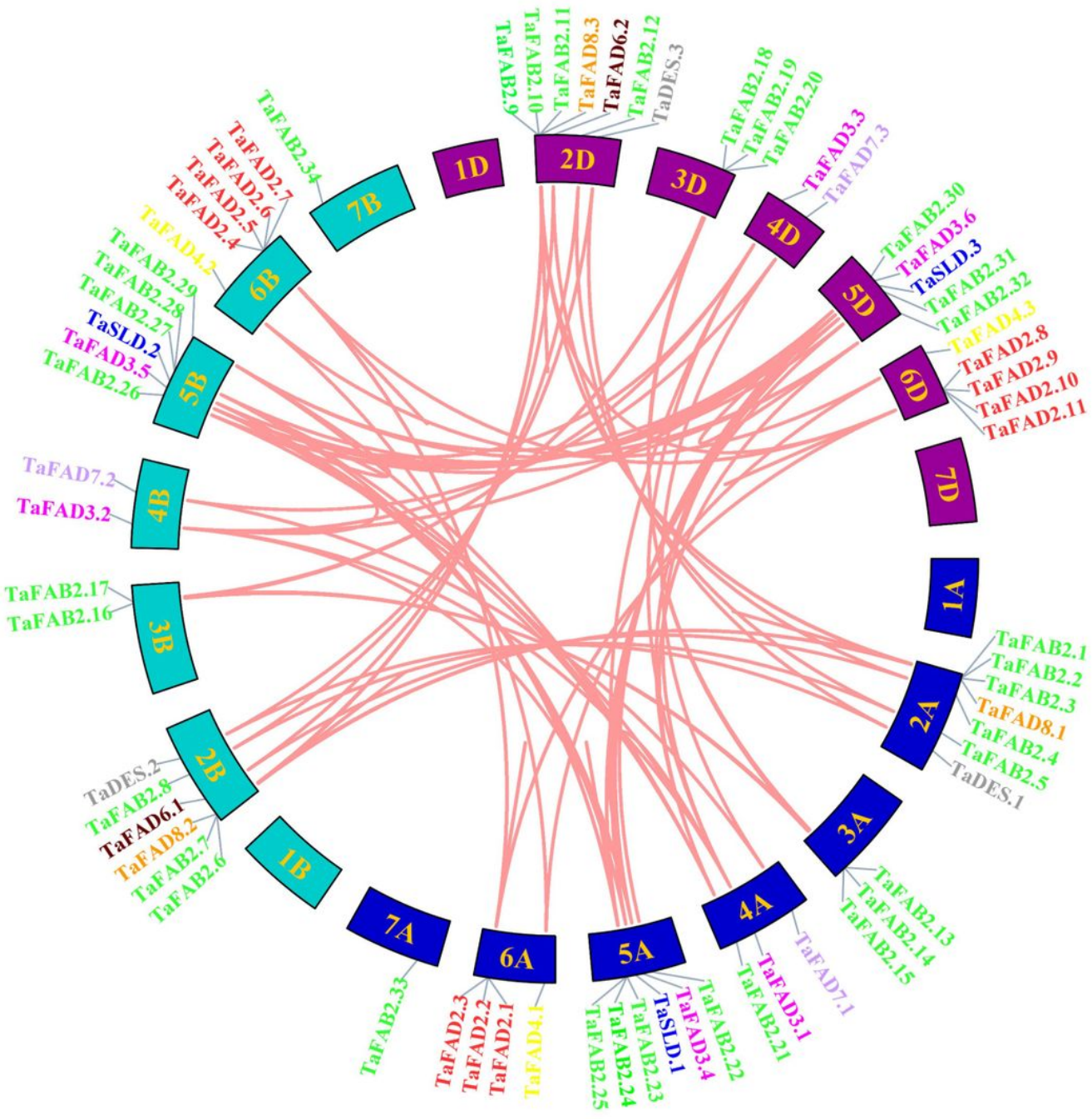

Figure 2

Chromosomal locations of wheat fatty acid desaturase genes. Chromosomes are represented by Colored boxes. Pink curves connecting the genes indicate duplications.

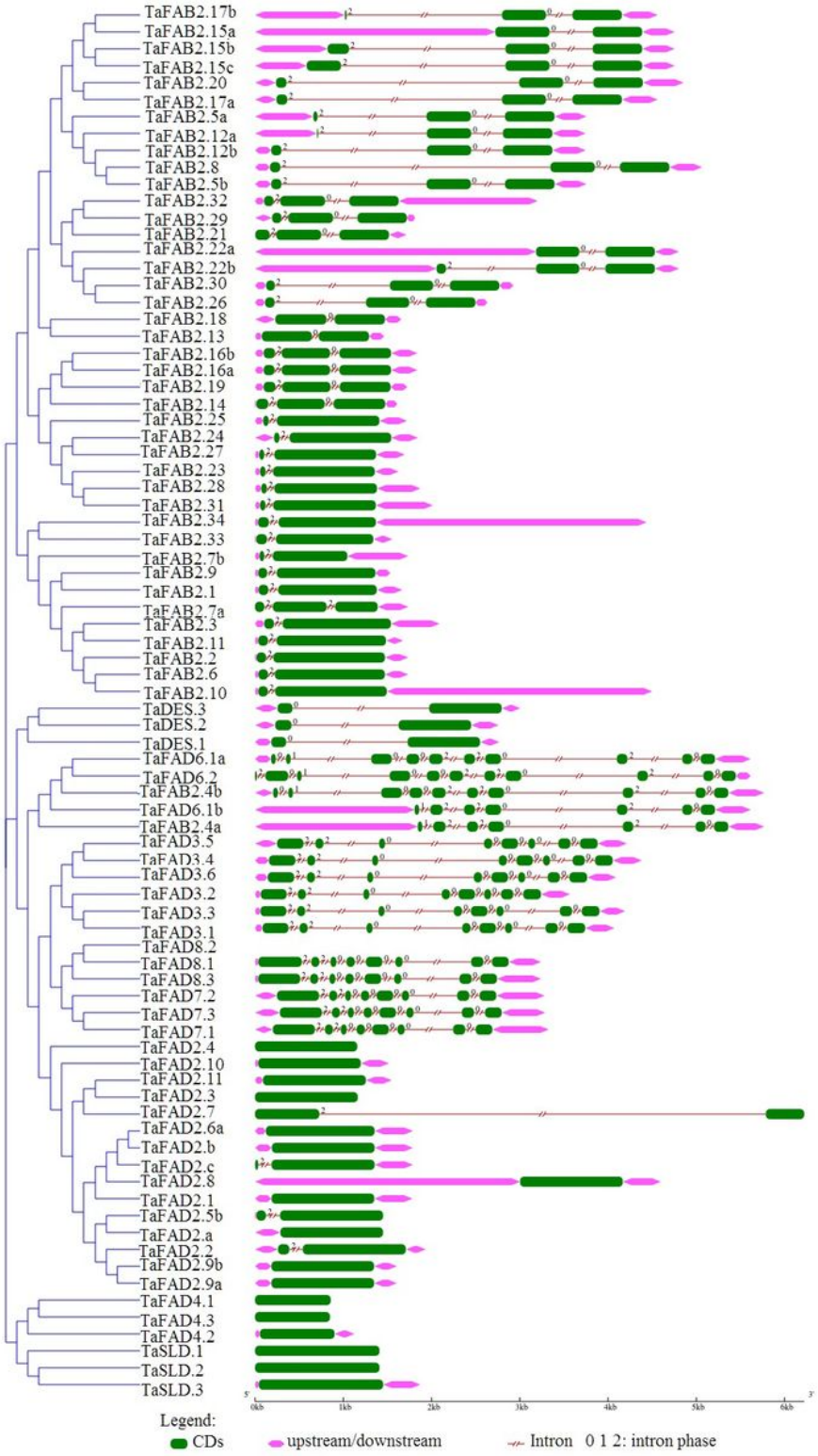

Figure 3

The exon-intron structure of FAD genes in wheat. Exons and introns were represented by green boxes and red lines, respectively.

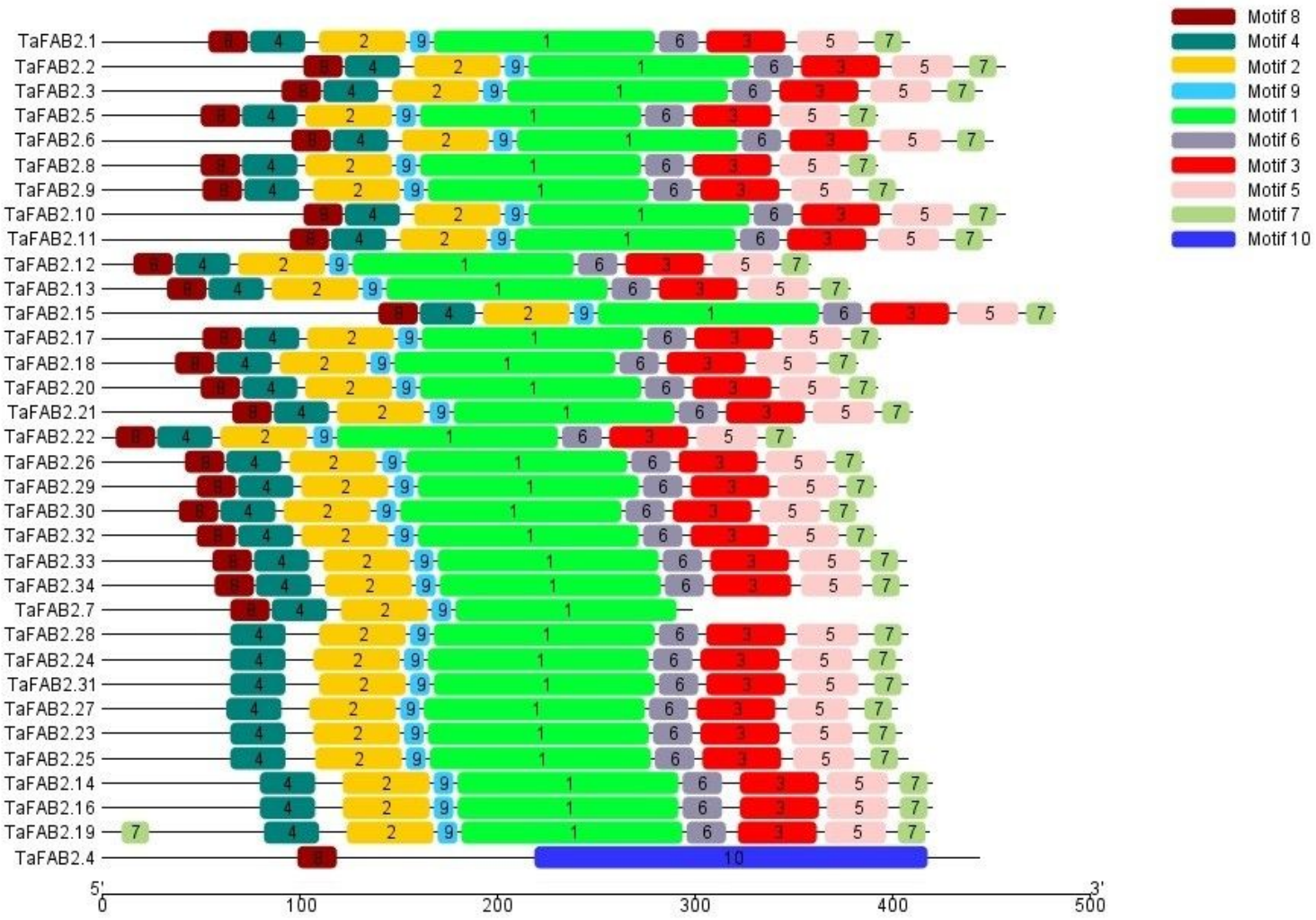

Figure 4

The conserved motifs of the TaFAB2 subfamily. Different motifs are presented in different colors.

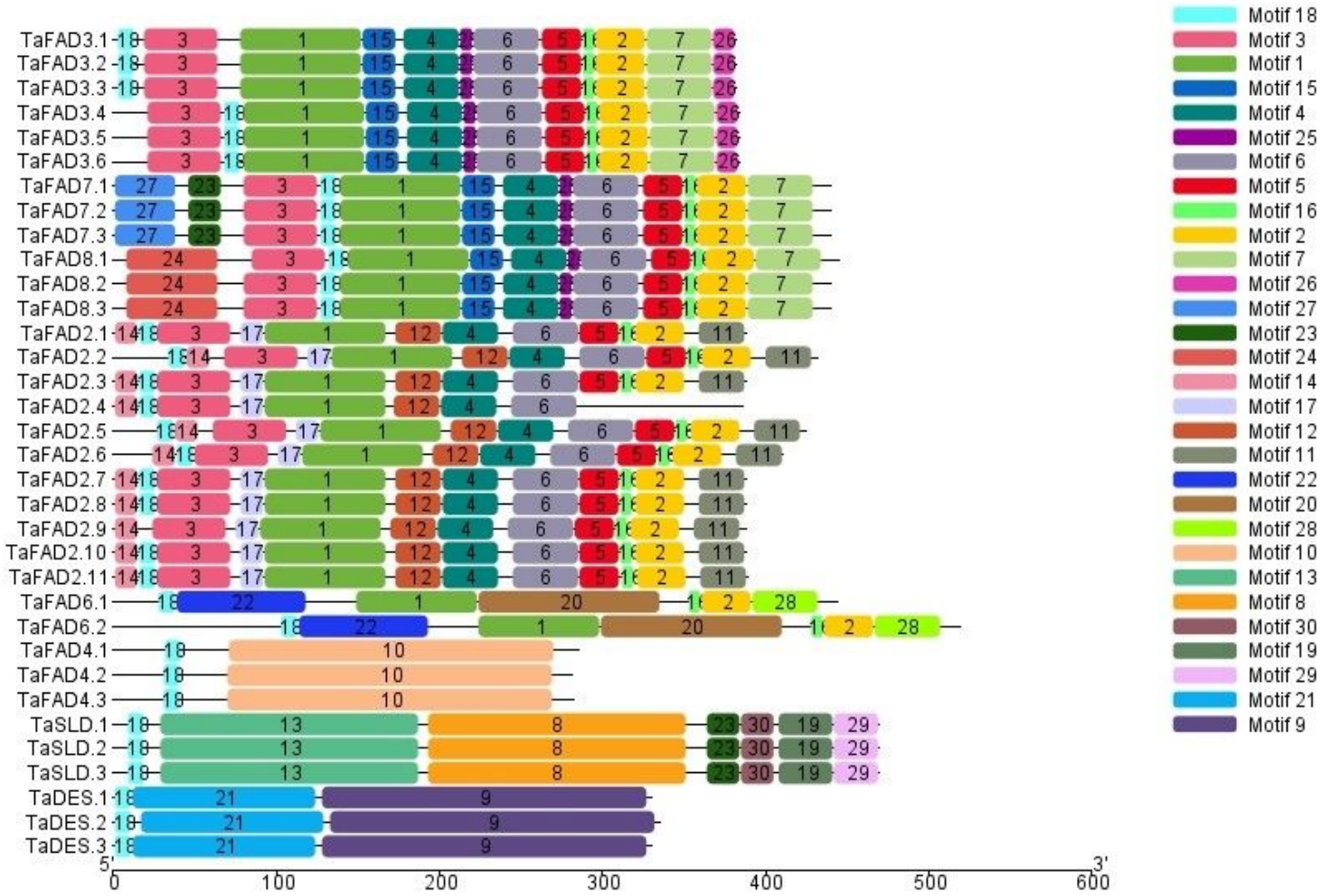

Figure 5

The conserved motifs of the TaFAD subfamily. Different motifs are presented in different colors.

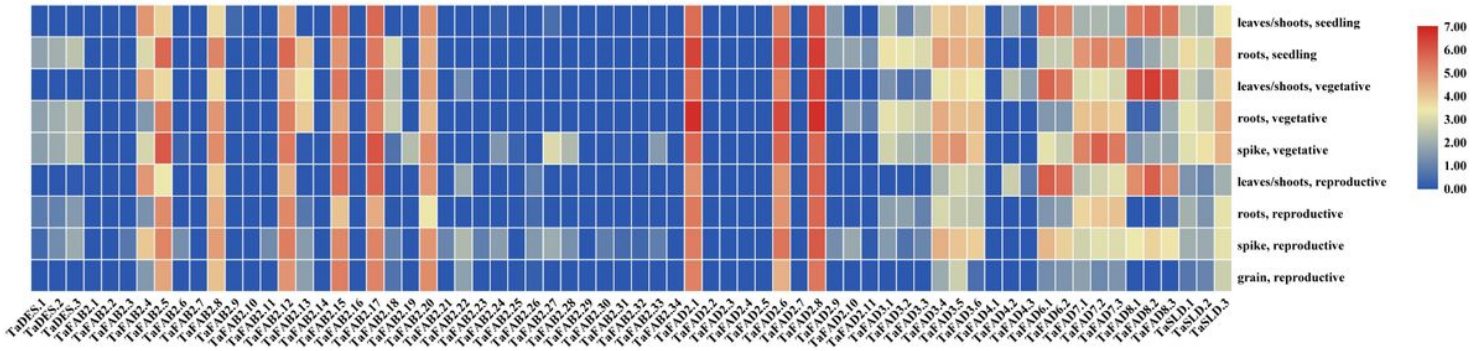

Figure 6

The expression pattern of TaFAD genes at different developmental stages and tissues. The color boxes indicate expression values, the lowest (green), medium (Pale goldenrod) and the highest (red).

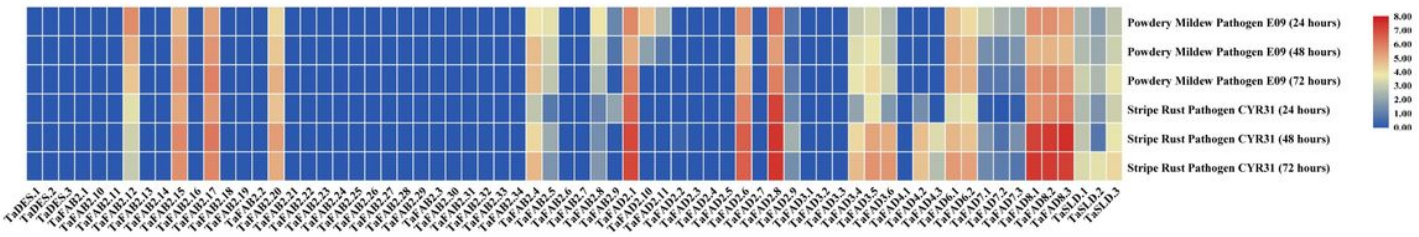

Figure 7

The expression pattern of TaFAD genes under biotic stress. The color boxes indicate expression values, the lowest (blue), medium (pale goldenrod) and the highest (red).

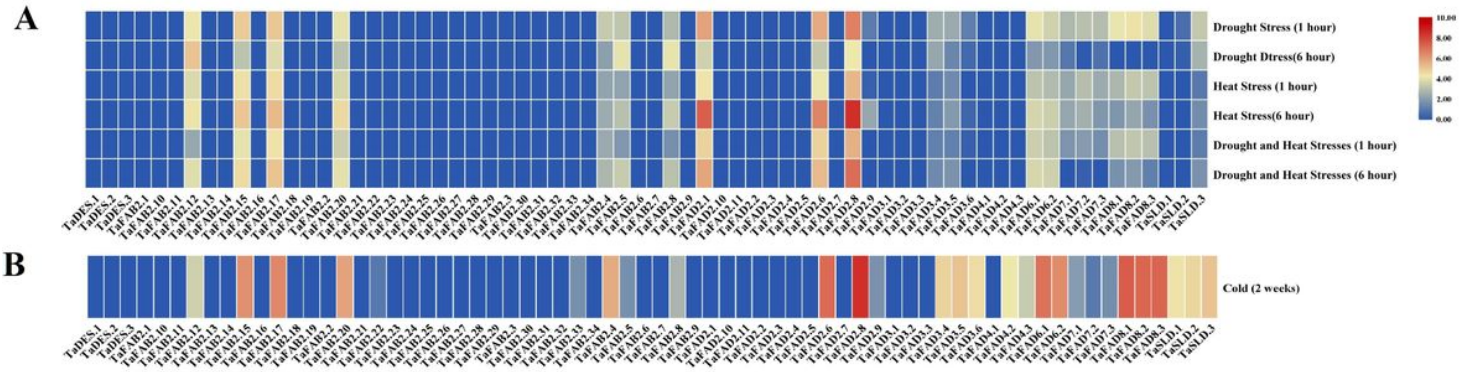

Figure 8

The expression pattern of TaFAD genes under heat, drought (a), and low temperature (b) conditions. The color boxes indicate expression values, the lowest (blue), medium (pale goldenrod) and the highest (red).

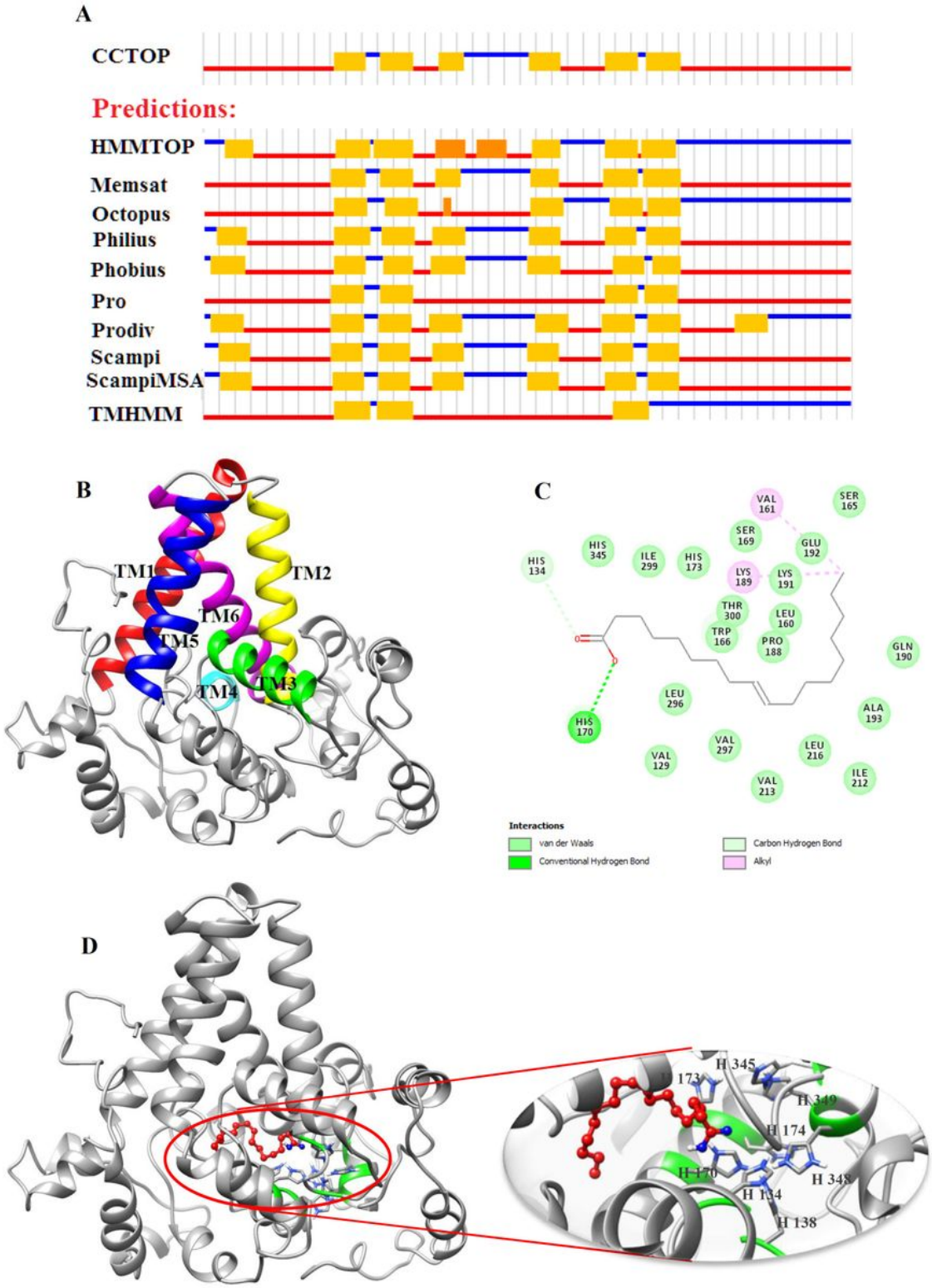

Figure 9

The TaFAD2.6 protein features. The predicted transmembrane (TM) topology from CCTOP prediction (a). Three-dimensional model structure of TaFAD2.6 (b). Six transmembrane domains were found in the protein structure that is shown in red, yellow, green, cyan, blue, and magenta for TM1, TM2, TM3, TM4, TM5, and TM6, respectively. The residues involved in the TaFAD2.6-oleic acid interaction (c). Docking

studies of the three-dimensional structure of oleic acid onto the predicted model of TaFAD2.6 (d). The ligand and histidine boxes are shown in red and green (side chain in blue), respectively.

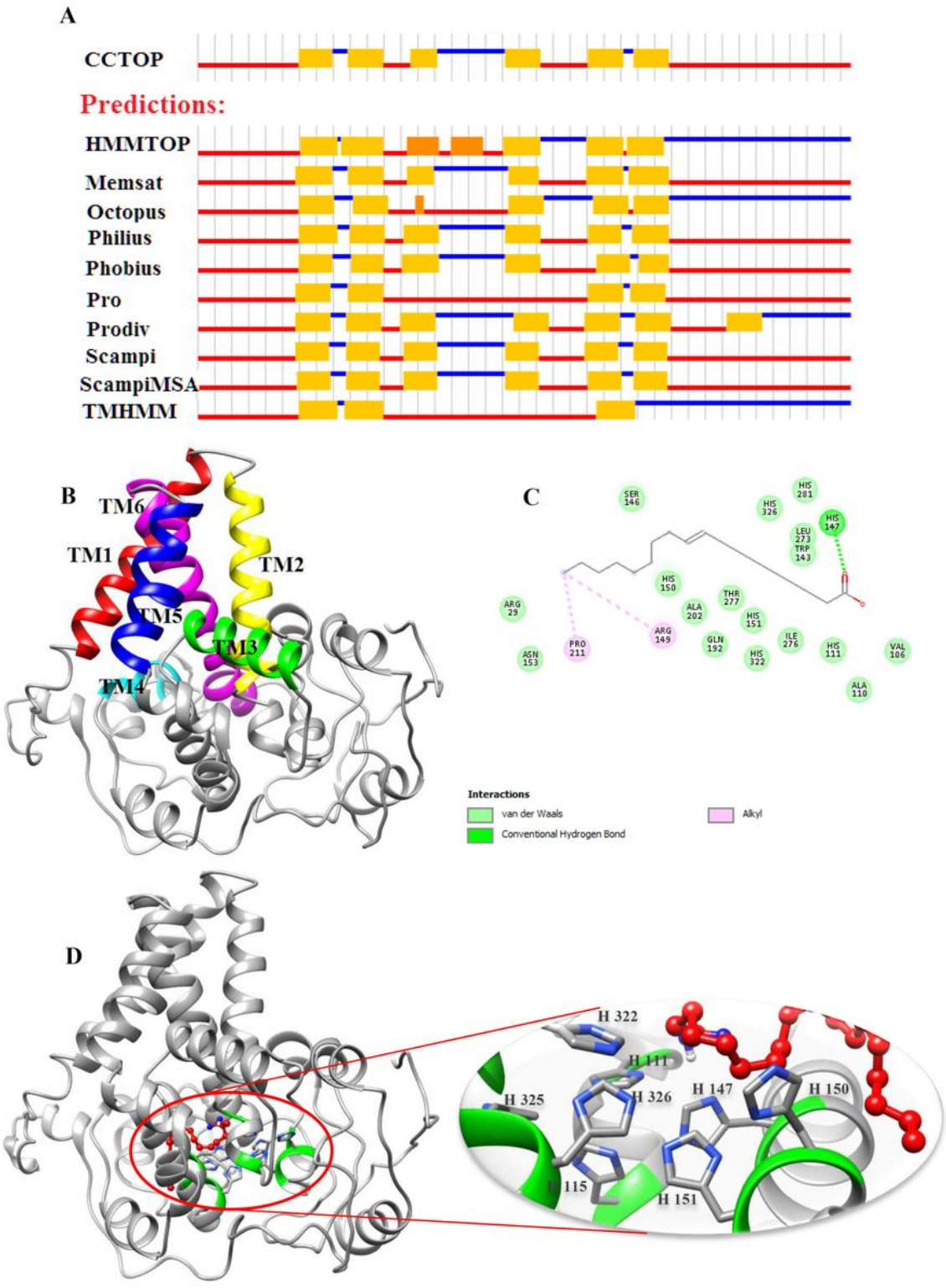

Figure 10

The TaFAD2.8 protein features. The predicted transmembrane (TM) topology from CCTOP prediction (a). Three-dimensional model structure of TaFAD2.8 (b). Six transmembrane domains were found in the protein structure that is shown in red, yellow, green, cyan, blue, and magenta for TM1, TM2, TM3, TM4,

TM5, and TM6, respectively. The residues involved in the TaFAD2.8-oleic acid interaction (c). Docking studies of the three-dimensional structure of oleic acid onto the predicted model of TaFAD2.8 (d). The ligand and histidine boxes are shown in red and green (side chain in blue), respectively.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS5.xlsx
- TableS6.xlsx
- TableS1.xlsx
- TableS2.xlsx
- TableS3.xls
- TableS7.xlsx
- TableS8.xls
- TableS4.xls