Genome-wide identification and characterization of Fibrillin gene family in wheat

Yaoyao Jiang[#1], Haichao Hu[#1], Yuhua Ma[2], Junliang Zhou[2]*

1, Zhejiang Agriculture and Forestry University, School of Forestry and Biotechnology, Hangzhou, Zhejiang, China

2, Zhejiang Agriculture and Forestry University, College of Agriculture and Food Science, Hangzhou , Zhejiang, China

3, Guizhou Academy of Agricultural Sciences, Guizhou Institute of Pomological Sciences, Guiyan, Guizhou, China

*Corresponding: Zhou Junliang, gsszjl2008@163.com (JL.Z)

[#] YY.J. and HC.H. contributed equally to this work.

ABSTRACT

Fibrillin is a highly conserved family of gene that is widely distributed in photosynthetic organs of plants. Members of this gene family are involved in the growth and development of plants and in response to biotic and abiotic stresses. As an important food crop, wheat, has a complex genetic background and less progress in molecular mechanism. In this study, we identified 26 *FBN* genes in the whole genome of wheat through bioinformatics. These genes were divided into 11 subgroups distributed on 11 chromosomes of wheat. The genetic structure of each subgroup of gene family members and the location and number of motifs were highly similar. At least 25 pairs of wheat *FBN* genes (*TaFBNs*) were genetically replicated by tandem replication. The results of evolutionary analysis indicated that the affinities of *FBNs* in monocots were closer. Tissue-specific analysis revealed that the *TaFBN* gene was expressed in different tissues and developmental stages. In addition, some *TaFBNs* were involved in one or more biotic and abiotic stresses. These results provide a basis for further study of the biological function of the *FBNs*.

Subjects: Bioinformatics, Genomics, Plant Science

Keywords: Fibrillin, *Triticum aestivum*, Abiotic stress, Gene duplication, Phylogenetic tree, *Cis*-regulatory elements

INTRODUCTION

Fibrillins (FBN, FIB) are named after fibrils because these proteins were first detected in fibrils in chromoplasts of dog rose (*Rosa rugosa*) and bell pepper (*Capsicum annuum*) fruit (*Deruère et al., 1994; Newman et al., 1989; Kim et al., 2015*). Since then, more and more fibrillin proteins have been found in different organelles, including plastoglobules (PGs) in chloroplasts and algal eyespots. As a result, members of the fibrillin protein family have been given many different names, including plastid lipid associated protein (PAP), plastoglobule (PGL), chloroplastic drought-induced stress protein (CDSP 34), and chromoplast-specific carotenoid-associated protein (ChrC) (*Pozueta-Romero et al., 1997; Ting et al., 1998; Kim et al., 2015*). Fibrillin proteins are found in the photosynthetic organs of cyanobacteria and some higher plants (Kim et al., 2015; Kim et al., 2017). A proteomic analysis performed by Lundquist *et al*. (2012) identified 14 *FBN* genes in *Arabidopsis*, 50% of which were located in PGs; the others were mainly distributed throughout the stroma and thylakoid membranes (*Lundquist et al., 2012; Kim et al., 2015*).

The fibrillin protein family comprises 12 subfamilies: 11 of these have been found in higher plants and one has been identified in algae (*Kim et al., 2017; Lohscheider and Río Bártulos, 2016*). However, the biophysical properties of fibrillin proteins are quite diverse. All these proteins contain hydrophobic structures, with molecular weights ranging between 20 and 42 kDa and isoelectric point (pI) values of 4 to 9 (*Vidi et al., 2006; Lundquist et al., 2012*). These findings suggest that each fibrillin protein may have specific biological functions. In *Arabidopsis thaliana*, fibrillin proteins contain a conserved hydrophobic domain (lipocalin motif 1) in the N-terminus and amino acid residues near the C terminus, including aspartic acid (*Singh et al., 2010*). Furthermore, Lohscheider and Río Bártulos (2016) predicted that the three-dimensional structure of FBNs is similar to that of lipocalin (*Lohscheider and Río Bártulos, 2016*), which suggests that the fibrillin family are involved in binding and transporting small hydrophobic molecules (*Kim et al., 2015, Singh et al., 2010*).

Fibrillin proteins have important biological functions, including participation in photosynthesis, the formation of lipoprotein structures, and responses to abiotic and biotic stresses (*Kim et al., 2015*). Immunogold electron microscopy revealed that fibrillin is located on the outer surface of red pepper chromoplast fibrils (*Deruère et al., 1994*). Furthermore, fibril-like structures can be reconstituted *in vitro* from a mixture of fibrillin protein, lipids, and bicyclic carotenoids (*Deruère et al., 1994; Kim et al., 2015*). Compared with wild type, RNAi-transgenic tomato plants with suppressed *LeCHRC* (FBN1) accumulate 30% less carotenoid (*Leitner-Dagan et al., 2006; Singh et al., 2010*). These results suggest that fibrillins play a major role in the formation of chromoplast fibrils and in the accumulation of carotenoid. In addition to structural roles, fibrillins are also involved in abiotic stress tolerance, especially oxidative stress (*Youssef et al., 2010*). For example, knockdown of *FBN4* expression in apple and mutation of *FBN4* in *Arabidopsis* caused plants be sensitivity to ozone. (*Singh et al., 2010*). Interestingly, similar results have been reported for cyanobacteria: *Synechocystis pgl1* and *pgl2* mutants were more sensitive to high light stress than the wild type (*Cunningham et al., 2010*). Furthermore, when the *FBN* gene in tomato, apple and *Arabidopsis* was knockdown, plants were more susceptible to the phytopathogenic fungus *Botrytis cinerea*, and pathogenic bacteria *Erwinia amylovora* and *Pseudomonas syringae* pv. *tomato*, respectively (*Leitner-Dagan et al., 2006; Cooper et al., 2003; Singh et al., 2010*). Moreover, the expression of fibrillin is regulated by hormones, including gibberellic acid (GA), jasmonate and abscisic acid, during plant growth and developmental stages and when plants are subjected to stresses (*Yang et al., 2006; Youssef et al., 2010; Kim et al., 2017*). The *FBN1* and *FBN2* gene families can respond to light and cold stress by modulating jasmonate biosynthesis (*Youssef et al., 2010*). By contrast, *FBN1* mRNA and protein levels declined in red pepper fruit when treated with GA (*Deruère et al., 1994*).

Wheat (*Triticum aestivum* L.) is an important food crop that is widely grown around the world. Approximately 40% of the global population depends on wheat as their staple food (*Paux et al., 2008; Han et al., 2019*). Common wheat is a heterogenous hexaploid containing A, B, and D genomes and, therefore, the genome information is large and complex (*Han et al., 2019; Glover et al., 2015; Ling et al., 2013*). At present, the study of wheat functional genomics is lagging far behind that of rice and corn. Moreover, owing to the complex genetic background of wheat (*Triticum aestivum* L.), studies was limited to the gene cloned for regulated important agronomic traits and molecular breeding. At present, the sequencing of the wheat genome has been completed, would play an important role in elucidating molecular mechanisms involved in growth and development, resistance and high yield..

Although there is more and more evidence that fibrillins play major roles in photosynthetic organisms, to date, fibrillins have only been identified and characterized in a few plant species. In addition ,functional annotation information about wheat FBNs is limited. The identification and functional characterization of the FBN family in wheat will contribute to elucidating the mechanism of stress response.. In this study, we performed a genome-wide survey using the reported FBN protein sequences in the wheat database. We identified 26 *FBN* genes in wheat and used bioinformatics methods to analyze their biophysical properties, including gene structures, conserved motifs, as well as the chromosome distribution and gene duplication of *FBN* genes. In addition, we analyzed the expression profiles of *TaFBN* genes in different tissues and at different developmental stages and in response to abiotic and biotic stresses using the GENEVESTIGATOR database. These results may provide a basis for studying the biological function of the protein encoded by the *FBN* gene in the future, and play an important role in understanding the function of FBN in different growth and development stages of plants.

MATERIALS AND METHODS

Identification of *TaFBN* genes

FBN sequences of *A. thaliana* were downloaded from The *Arabidopsis* Information Resource database (https://www.arabidopsis.org/). Previously identified protein sequences of *TaFBN* genes were used as queries to perform a BLAST (E-value le-10) search against *T. aestivum* genome sequences to obtain a dataset of *TaFBN* sequences (http://ensembl.gramene.org/). Redundant sequences were filtered out. The online tools SMART (http://smart.embl-heidelberg.de/) and Interproscan (http://www.ebi.ac.uk/interpro/) were used to predict the functional domains of potential FBN proteins. A typical FBN protein is reported to contain a conserved PAP_fibrillin domain To verify our results, all of the proteins were compared with the PAP_fibrillin domain (PF04755) using the HMMER 3.0 program with the default E-value. Proteins without the PAP_fibrillin domain were removed. The biophysical properties of the final protein sequences of TaFBNs were calculated using the ExPASy ProtParam tool (https://web.expasy.org). The subcellular localization of TaFBNs was analyzed using ProComp (http://linux1.softberry.com) and WoLF PSORT II (https://www.genscript.com/wolf-psort.html). In addition, the signal peptide and chloroplast transit peptides of *TaFBN* genes were predicted using the SignalP 4.1 server (http://www.cbs.dtu.dk/services/SignalP-4.1/) and ChloroP 1.1 server (http://www.cbs.dtu.dk/services/ChloroP/). The theoretical values of the pI, relative molecular mass and the grand average of hydrophobicity (GRAVY) of TaFBNs were predicted using the ProtParam tool (https://web.expasy.org).

Multiple sequence alignments and phylogenetic analysis

Full-length protein sequences of *FBN* gene family members identified in seven plant species, including monocotyledons and dicotyledons, were downloaded from the NCBI database (https://www.ncbi.nlm.nih.gov/) and the phytozome v12.1 database (https://phytozome.jgi.doe.gov/pz/portal.html). Full-length protein sequences of these *FBN* genes were aligned using MAFFT software (https://mafft.cbrc.jp/alignment/server/). Based on fasta files, a neighbor-joining phylogenetic tree was constructed using Molecular Evolutionary Genetics Analysis (MEGA) version 7.0 software with bootstrap values of 1000 replicates.

Analysis of gene structures and conserved motifs

To investigate the structure of *TaFBN* genes, we used the Gene Structure Display Server 2.0 database (http://gsds.cbi.pku.edu.cn/) to analyze the distribution of exons and introns in *TaFBN* genes. Conserved motifs were predicted using the MEME database (http://alternate.meme-suite.org/; the number of motifs was set to 10 and the motif width was set to 6 to 50).

Chromosome distribution, gene duplication, and the calculation of Ka/Ks

Information about the position of the *TaFBN* gene family on chromosomes was obtained from the phytozome v12.0 database (https://phytozome.jgi.doe.gov/pz/portal.html). The *TaFBN* gene distribution map was analyzed based on sequencing genome information obtained from the phytozome v12.1 database. *TaFBN* coding sequences (CDS) were blasted against each other; when the identity exceeded 90%, we inferred that genes were produced by gene duplication (*Song et al., 2016; Ning et al., 2017*). Two or more adjacent homologous genes located on a single chromosome were defined as tandem duplicated; non-adjacent homologous genes were defined as segmental duplicated genes. We used MCScan X software and the Plant Duplicate Gene Database (http://pdgd.njau.edu.cn) to examine the duplication of genes among the *T. aestivum* genomes. The distribution of *FBN* genes on chromosomes and gene duplication of TaFBNs were plotted using TBtools software (https://github.com/CJ-Chen/TBtools). In order to analyze whether positive selection has driven the evolution of the *FBN* genes, we calculated the synonymous substitution rate (Ks) and non-synonymous substitution rate (Ka) values of orthologous genes using the DnaSP 6 software (*Rozas, 2017*). The formula, $T = Ks/(2 \times 1.5 \times 10^{-8}) \times 10^{-6}$ million years ago (MYA) was used to calculate the divergence time (*Koch et al., 2000*).

Analysis of the *cis*-regulatory element of *FBN* gene promoters

In this study, 2000-bp sequences upstream of translational start sites of Tagene*FBN* genes were set as promoter sequences. PlantCARE software (http://bioinformatics.psb.ugent.be/webtools/plant care/html/) was used to predict the *cis*-regulatory elements based on these promoter sequences.

Analysis of *TaFBN* gene expression patterns

Expression profile data used in this study were obtained via the GENEVESTIGATOR tool (*Li et al., 2015; Jangam et al., 2016*). We searched for *FBN* genes on the website using keywords, gene ID and probe ID numbers as query terms. The expression of *TaFBN*s in different tissues, at different developmental stages and under different abiotic and biotic stress conditions (i.e., included drought, heat, salt, and hormones) were analyzed. The results were presented as heatmaps, with different colors representing the absolute signal values. The color scale of the heatmap was given in $\log_2$ ratio values. The cultivar used in the gene expression profiles analysis was Chinese spring (*Zimmermann et al., 2004*).

RESULTS

Identification and characterisation of *FBN* genes in *Triticum aestivum*

In this study, a total of 26 *FBN* genes were identified in wheat, which we named *TaFBN*-1–*TaFBN*-26 (Table.1). *TaFBN* characteristics, including chromosomal localization, intron number, gene length, number of amino acids, molecular mass, CDS, subcellular localization, signal peptide, and instability index are listed in Table 1. The results are as follows, *Ta*FBN protein sequences ranged from 219 to 402 amino acids and molecular weights ranged from 23.75 to 43.59 kDa. The prediction of subcellular location indicated that 18 *Ta*FBNs were located in chloroplasts and eight were located extracellularly. As we know, GRAVY values reflects the hydrophobicity of the protein, almost all *Ta*FBN proteins GRAVY values less than 0, except *Ta*FBN2, *Ta*FBN22 and *Ta*FBN26. Meanwhile, the predict results showed that all *Ta*FBN proteins don't contain signal peptides, but chloroplast transit peptides was found in all *Ta*FBN proteins.

Gene structure analysis of *TaFBN* genes

To gain insight into the evolution of the *TaFBN* gene family, a diagram of the *TaFBN* exon–intron gene structure was constructed based on cDNA and genomic DNA sequence information (Supplementary txt.1) using the Gene Structure Display Server (Figure. 1b). A neighbor-joining phylogenetic tree was also constructed to explore the relationship of exon–intron distribution patterns and the phylogenetic classification. Gene structure analyses indicated that homologous genes had similar exon-intron distribution pattern (Figure. 1b). However, the number of introns in different *TaFBN* gene family members varied greatly (ranging from 2 to 10 introns), while less difference between the same subfamily members. These results suggested that the same subfamily members may have similar biological functions. In addition, we used the MEME online tool to analysis the conserved motifs of *Ta*FBNs, the results showed that all *Ta*FBN members contain five to nine conserved motifs (Figure. 2). Logo representation of the ten conserved motifs identified for proteins encoded by *TaFBN* genes as (Supplementary Figure.1) described. Figure. 2 showed that motif 1, motif 2, motif 3, motif 4, and motif 5motif 6 were highly conserved and widely distributed in all *Ta*FBN proteins. Motif-domain analysis revealed that motif 1 contains conserved amino acid residues in the C-terminal and motif 3 contains a conserved lipocalin motif (Supplementary Figure.2). The types and distribution of conserved motifs may be the reason for the functional diversity of *Ta*FBNs.

Phylogenetic and evolutionary analysis of *TaFBN*

An unrooted phylogenetic tree was constructed for 26 *FBNs* from wheat, 9 *FBNs* from rice, 14 *FBNs* from *A. thaliana*, 10 *FBNs* from *Sorghum bicolor*, 14 *FBNs* from *Zea mays*, 17 *FBNs* from *Brassica oleracea* var. *capitata*, and 13 *FBNs* from *Nicotiana tabacum* to study the evolutionary relationships of *TaFBN* members (Figure. 3). These 103 FBNs were classified into eleven subfamilies (Group 1 to Group 11). This classification method was based on the *FBN* gene characteristics of *A. thaliana*. Interestingly, nine subfamilies of *FBN* genes were identified in most species, each subfamily contain three *FBN* gene. The analysis also revealed that the *FBN* genes of the monocotyledons plants (i.e. *T. aestivum*, *O. sativa*, *Z. mays* and *S. bicolor*) were more closely related than those of the dicotyledons plants (i.e., *A. thaliana*, *B. oleracea* var. *capitata*, and *N. tabacum*).

Analysis of *TaFBN cis*-regulatory elements

In order to further identify the *cis*-regulatory elements located upstream of *TaFBN* genes, 2000-bp sequences upstream from translational start sites of putative *TaFBN* gene families were analyzed using the PlantCARE tool. As shown in Figure.4, many *cis*-regulatory elements were identified in promoters of *TaFBN* genes. There were mainly three types of *cis*-regulatory elements: hormone response elements, stress response-related reactive elements, and light-induced reactive-related elements. Hormone response elements were widely distributed in promoters of the *TaFBN*, including the methyl jasmonate (Me-JA)-responsive element, abscisic acid-responsive element, gibberellin-responsive element, salicylic acid-responsive element, and auxin-responsive element. The main components associated with abiotic stress responses were the light-response element, the low-temperature response element and the drought-stress response element. These results suggested that *TaFBN* genes may play a major role in photosynthesis, stress responses, and in maintaining the hormone balance in plants, thereby improving the chances for organisms to escape or better cope with the damaging effects of adverse environmental conditions.

Chromosomal location and duplication of *TaFBN* in wheat genome

In order to clarify the distribution of *TaFBN* genes on wheat chromosomes, the 26 *TaFBN* genes were mapped onto 11 wheat chromosomes (Figure. 5). As shown in Figure. 5, we can found *TaFBN* genes were not randomly distributed on the chromosome, chromosome 2A, 2B, and 2D contain six *TaFBN* genes, respectively, however, the other chromosomes (i.e. chromosomes 1A, 1B, 1D, 4A, 4B, 4D, 5A and 5B) have only one *TaFBN* gene. The unequal distribution of *TaFBN* genes in chromosomes are caused by gene duplication. In order to research potential gene duplication within the *TaFBN* gene family, segmental duplications and tandem duplications during evolution were analyzed. Gene sequence homology, positional information, and chromosomal position detection of gene duplication analyses revealed that at least 25 pairs of *TaFBN* genes underwent gene duplication arising from segmental duplications. The values of synonymous (Ks) and non synonymous (Ka) could be driven the evolution process of the *TaFBN* genes. The ratio of Ka/Ks was less than 0.3 in all duplicated gene pairs showed that *TaFBN* gene family in wheat has evolved under purifying selection (Table. 2). In addition, duplication events was occured on as old as 5.233 MYA.

Tiussse special expression patterns of *TaFBNs* at different developmental stages

To explore the tissue-specific expression patterns of *TaFBN* genes at different growth and development stages in wheat, publicly available microarray data sets for three *TaFBNs* were analyzed to examine transcription levels in various wheat tissues, including the root, anther, spikelet and leaf. Most of the *TaFBN* genes can be detected at least two or more different tissues. In total, *TaFBN* genes were constitutively expressed in seven different tissues (Figure. 6a). Moreover, the expression levels of *TaFBN* genes varied among the sevendifferent tissues. The results suggested that the expression levels of TaFBNs in leaves were strongly high than others tiusses. As shown in Figure. 6b, the expression levels of *TaFBN* were notabllydifferent at different developmental stages. High levels of expression were also detected for all *TaFBN* genes during seedling growth and stem elongation, but low levels of expression were detected during the germination stage. These data indicate that *TaFBN* genes have tissue-specific expression patterns in wheat.

Expression profiles of *TaFBN* genes in response to abiotic stress

To further clarify the potential functions of *TaFBN* genes under abiotic stress responses, the expression levels of *TaFBN* genes were analyzed under drought, salinity, cold, nitrogen, and heat conditions. Almost all *TaFBN* genes could be induced by various abiotic stresses, including salt and nitrogen stress (Figure. 7). The transcripts of *TaFBN4*, *TaFBN5*, and *TaFBN6* were slightly downregulated by cold and heat treatment. Interestingly, the transcript levels of the tested *TaFBN* genes were significantly downregulated during drought stress conditions. These results indicated that *TaFBN* genes might be involved in wheat responses to abiotic stress, especially salt and nitrogen stress.

DISCUSSION

In this study, we identified 26 *FBN* genes in the wheat genome. These genes were distributed on 11 chromosomes and molecular masses ranging from 23.75 to 43.59 kDa and pI values ranging from 4.59 to 9.61. This diversity suggests that these gene may have specific biological functions in different metabolic processes. Furthermore, study results indicate that most of *TaFBN* genes located on chloroplast, and contained chloroplast transit peptides. This suggests that the various fibrillins might participate in photosynthesis. GRAVY was used to calculate the overall hydrophobicity of the protein sequence, with a higher positive GRAVY values indicating a greater level of hydrophobicity (*Faya et al., 2015*). Most all of the *TaFBN* genes' GRAVY values were less than 0, this means that most proteins are hydrophilic. By contrast, previous studies have reported that the fibrillin family can bind to and transport small hydrophobic molecules in *A. thaliana* (*Kim et al., 2015; Singh and McNellis, 2011*). However, the spatial structure and the percentage of hydrophobic residues may affect the hydrophobicity of proteins (*Dyson et al., 2004*). Therefore, these different results may reflect the diversity of TaFBN protein structures.

Phylogenetic tree analysis showed that there were more *FBN* genes in *T. aestivum* than there were in the other monocotyledons and dicotyledons plants analyzed in this study. This is because common wheat (*T. aestivum*) is an allohexaploid species with three genomes A, B, and D. Indeed, some reports have suggested that more than 85% of sequences are repeated in the wheat genome (*Glover et al., 2015; Ling et al., 2013; Han et al., 2019*). To analyze the evolutionary relationships of *FBN* genes, we constructed a phylogenetic tree with 103 FBNs from *T. aestivum*, *O. sativa*, *S. bicolor*, *Z. mays*, *A. thaliana*, *B. oleracea* var. *capitata* and *N. tabacum*. These *FBN* genes were divided into 11 subfamilies using the classification method described for FBN in *Arabidopsis thaliana* (*Singh and McNellis, 2011*). Interestingly, similar exon–intron structures and similar numbers of conserved motifs were found in the same subgroups. This phenomenon suggests that *FBN* genes in the same subgroup have similar functions. However, many gene duplication events have occurred and these are one of the primary driving forces in wheat evolution (*Han et al., 2019*). The gene duplication and chromosomal localization analysis showed that most *TaFBN* genes originated from segmental duplication events. The member of *TaFBN* gene family can increase by duplication of chromosomal segments, and duplicated genes will have also played a role in the evolution of wheat (*Moore and Purugganan, 2005*).

Gene expression in different tissues and at different developmental stages may be determined by gene function. Previous studies have shown that fibrillins are regulated by a variety of biological and environmental factors at different growth and developmental stages (*Singh and McNellis, 2011*). We analyzed the expression patterns of TaFBN gene family at different growth, development stages and respond to abiotic stress in wheat though publicly available microarray data. Unfortunately, in these studies, we only obtained three TaFBN genes expression profiles. This may be due to wheat genome have a complex genetic background, the sequencing work of wheat genome was not completed. In addition, there are a few reported about *TaFBN* gene. Expression profile data show that the highest expression levels of most *TaFBN* genes were found in leaves and at stem elongation stages. Similar results have been reported in potato, *Arabidopsis* and *Brassica rapa* (*Monte et al., 1999; Yang et al., 2006; Kim et al., 2001*). Furthermore, the expression profile data suggested that *TaFBN* expression was strongly induced under salt and nitrogen stresses but was only slightly changed under cold and heat stresses. Interestingly, *TaFBN* expression was downregulated under drought stress. As we know, transcription factors participated in various biological processes by regulating the expression of downstream genes *cis*-regulatory elements. (*Ning et al, 2017*) In this study, a large number of *cis*-regulatory elements were detected in the promoter sequence of *TaFBN* genes. These elements contained light responsiveness elements, drought responsiveness elements and hormone responsiveness elements, such as MeJA, abscisic acid, GA, salicylic acid, and auxin. Indeed, all the *TaFBN* genes include many light-responsive elements. Such as Rey *et al* found that overexpressing FBN1 can promoted plant height and flowering under high light levels in tobacco (*Rey et al., 2000; Singh and McNellis, 2011*). Although the expression patterns of *TaFBN* genes was varied and complexed, overall, these genes had similar functions in plant stress resistance and chromoplast development (*Singh and McNellis, 2011*).

CONCLUSION

In this study, we identified 26 *FBN* genes in wheat using a genome-wide screening approach. Based on their phylogenetic relationships, these *FBN* genes were classified into 11 subfamilies. Analysis of *TaFBN* gene structures and conserved motifs revealed that *TaFBN* genes in the same subgroup were highly conserved. Analysis of *cis*-regulatory elements of *TaFBN* genes showed that the expression of *TaFBN* genes was regulated by various hormones and environmental factors. Tissue-specific expression analysis revealed that the highest levels of *TaFBN* gene expression were found in leaves and stem elongation stages. The expression profiling data suggest that *TaFBN* genes are involved in biotic and abiotic stress responses. These works can helped us to clarify the structural and functional relationships among *TaFBN* gene family member.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This study was supported by National Natural Science Foundation of China (No. 31560547), Talents of Guizhou science and technology cooperation platform ([2017]5603), Scientific and Technological Innovation Platform of Guizhou Province (No. 2014-4003), Research institutions Enterprise Action Program of Guizhou Province (No. 2014-4007). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

National Natural Science Foundation of China (No. 31560547).

Talents of Guizhou science and technology cooperation platform ([2017]5603).

Scientific and Technological Innovation Platform of Guizhou Province (No. 2014-4003).

Research institutions Enterprise Action Program of Guizhou Province (No. 2014-4007).

Competing Interests

The authors declare that they have no competing interests.

Data Availability.

Cooper B, Clarke JD, Budworth P, Kreps J, Hutchison D, Park S. 2003. A network of rice genes associated with stress response and seed development. *Proceedings of the National Academy of Sciences of the United States of America* **100**(8):4945-4950 DOI:10.1073/pnas.0737574100.

Cunningham FX, Tice AB, Pham C, Gantt E. 2010. Inactivation of genes encoding plastoglobulin-like proteins in Synechocystis sp. pcc 6803 leads to a light-sensitive phenotype. *Journal of Bacteriology* **192**(6):1700-1709 DOI:10.1128/JB.01434-09.

Deruère J, Römer S, d'Harlingue A, Backhaus RA, Kuntz M, Camara B. 1994. Fibril assembly and carotenoid overaccumulation in chromoplasts: a model for supramolecular lipoprotein structures. *Plant Cell* **6**: 119–133 DOI:10.2307/3869680.

Dyson MR, Shadbolt SP, Vincent KJ, Perera RL, McCafferty J. 2004. Production of soluble mammalian proteins in Escherichia coli: identification of protein features that correlate with successful expression. *BMC Biotechnology* **4**:32 DOI:10.1186/1472-6750-4-32.

Faya N, Penkler DL, Tastan Bishop Ö. 2015. Human, vector and parasite hsp90 proteins: a comparative bioinformatics analysis. *FEBS Open Bio* **5**(1):916-927 DOI:10.1016/j.fob.2015.11.003.

Glover NM, Daron J, Pingault L, Vandepoele K, Paux E, Feuillet C, Choulet F. 2015. Small-scale gene duplications played a major role in the recent evolution of wheat chromosome 3B. *Genome Biology* **16**(1):188 DOI:10.1186/s13059-015-0754-6.

Han Z, Liu Y, Deng X, Liu D, Liu Y, Hu Y. 2019. Genome-wide identification and expression analysis of expansin gene family in common wheat (*Triticum aestivum* L.). *BMC Genomics* **20**(1):101 DOI:10.1186/s12864-019-5455-1.

Jangam AP, Pathak RR, Nandula R. 2016. Microarray analysis of rice d1 (rga1) mutant reveals the potential role of g-protein alpha subunit in regulating multiple abiotic stresses such as drought, salinity, heat, and cold. *Frontiers in Plant Science* **7**:11 DOI:10.3389/fpls.2016.00011.

Kim EH, Lee Y, Kim HU. 2015. Fibrillin 5 is essential for plastoquinone-9 biosynthesis by binding to solanesyl diphosphate synthases in Arabidopsis. *The Plant Cell* **27**: 2956–2971 DOI:10.1105/tpc.15.00707.

Kim HU, Wu SS, Ratnayake C, Huang AH. 2001. *Brassica rapa* has three genes that encode proteins associated with different neutral lipids in plastids of specific tissues. *Plant Physiology* **126:**330–341 DOI:0.2307/4279891.

Kim EH, Lee DW, Lee KR, Jung SJ, Jeon JS, Kim HU. 2017. Conserved function of fibrillin5 in the plastoquinone-9 biosynthetic pathway in Arabidopsis and rice. *Frontiers in Plant Science* **8**:1197 DOI: 10.3389/fpls.2017.01197.

Koch MA, Haubold B, Mitchell-Olds T. 2000. Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabis*, and related genera (Brassicaceae). *Molecular Biology & Evolution* **17**(10):1483–1498 DOI:10.3389/fpls.2017.01197.

Laizet Y, Pontier D, Mache, Régis, Kuntz M. 2004. Subfamily organization and phylogenetic origin of genes encoding plastid lipid-associated proteins of the fibrillin type. *Journal of Genome Science & Technology* **3**(1):19-28 DOI:info:doi/10.1166/gl.2004.038.

Leitner-Dagan, Ovadis M, Shklarman E, Elad Y, Rav David D, Vainstein A. 2006. Expression and functional analyses of the plastid lipid-associated protein CHRC suggest its role in chromoplastogenesis and stress. *Plant Physiology* **142**(1):233-244 DOI:10.2307/20205918.

Li DY, Fu FY, Zhang HJ, Song FM. 2015. Genome-wide systematic characterization of the bzip transcriptional factor family in tomato (*Solanum lycopersicum* L.). *BMC Genomics* **16**(1): 771 DOI:10.1186/s12864-015-1990-6.

Ling HQ, Zhao S, Liu D, Wang J, Sun H, Zhang C. 2013. Draft genome of the wheat A-genome progenitor *Triticum urartu*. *Nature* **496**(2):87-90 DOI:10.1038/nature11997.

Lohscheider JN, Río Bártulos C. 2016. Plastoglobules in algae: a comprehensive comparative study of the presence of major structural and functional components in complex plastids. *Marine Genomics* **28**:127-136 DOI:10.1016/j.margen.2016.06.005.

Lundquist PK, Poliakov A, Bhuiyan NH, Zybailov B, Sun Q, van Wijk KJ. 2012. The functional network of the Arabidopsis plastoglobule proteome based on quantitative proteomics and genome-wide coexpression analysis. *Plant Physiology* **158**: 1172– 1192 DOI:10.1104/pp.111.193144.

Monte E, Ludevid D, Prat S. 1999. Leaf C40. 4: a carotenoid-associated protein involved in the modulation of photosynthetic efficiency? *The Plant Journal* **19**:399–410 DOI: 10.1046/j.1365-313x.1999.00537.x.

Moore RC, Purugganan MD. 2004. The evolutionary dynamics of plant duplicate genes. *Current Opinion in Plant Biology* **8**(2): 122-128 DOI:10.1016/j.pbi.2004.12.001.

Newman LA, Hadjeb N, Price CA. 1989. Synthesis of two chromoplast-specific proteins during fruit development in capsicum annuum. *Plant Physiology* **91**(2):455-458 DOI:10.1104/pp.91.2.455.

Ning P, Liu C, Kang J. 2017. Genome-wide analysis of wrky transcription factors in wheat (*Triticum aestivum* L.) and differential expression under water deficit condition. *Peerj* **5**(5):e3232 DOI:10.7717/peerj.3232.

Paux E, Legeai F, Guilhot N, Adam-Blondon AF, Alaux M. 2008. Physical mapping in large genomes: accelerating anchoring of bac contigs to genetic maps through in silico analysis. *Functional & Integrative Genomics* **8**(1):29-32 DOI:10.1007/s10142-007-0068-1.

Pozueta-Romero J, Rafia F, Houlné G, Cheniclet C, Carde JP, Schantz ML, Schantz R. 1997. A ubiquitous plant housekeeping gene, PAP, encodes a major protein component of bell pepper chromoplasts. *Plant Physiology* **115**: 1185–1194 DOI:10.1104/pp.115.3.1185.

Rey P, Gillet B, Römer S, Eymery F, Massimino J, Peltier G, Kuntz M. 2000. Over-expression of a pepper plastid lipid-associated protein in tobacco leads to changes in plastid ultrastructure and plant development upon stress. *The Plant Journal* **21**:483–494 DOI:10.1046/j.1365-313x.2000.00699.x.

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A. 2017. Dnasp 6: dna sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution* **34**:3299–3302 DOI:10.1093/molbev/msx248.

Singh DK, McNellis TW. 2011. Fibrillin protein function: the tip of the iceberg? *Trends in Plant Science* **16**(8):432-441 DOI:10.1016/j.tplants.2011.03.014.

Singh DK, Maximova SN, Jensen PJ, Lehman BL, McNellis TW. 2010. FIBRILLIN4 is required for plastoglobule development and stress resistance in apple and Arabidopsis. *Plant Physiology* **154**(3):1281-1293 DOI:10.1104/pp.110.164095.

Song H, Wang P, Hou L, Zhao S, Zhao C, Xia H. 2016. Global analysis of wrky genes and their response to dehydration and salt stress in soybean. *Frontiers in Plant Science* **7**:9 DOI:10.3389/fpls.2016.00009.

Ting JTL, Wu SS, Ratnayake C, Huang AH. 1998. Constituents of the tapetosomes and elaioplasts in *Brassica campestris* tapetum and their degradation and retention during microsporogenesis. *The Plant Journal* **16**:541–551 DOI:10.1046/j.1365-313x.1998.00325.x.

Vidi PA, Kanwischer M, Baginsky S, Austin JR, Csucs G, Dörmann P, Kessler F, Bréhélin C. 2006. Tocopherol cyclase (VTE1) localization and vitamin E accumulation in chloroplast plastoglobule lipoprotein particles. *Journal of Biological Chemistry* **281**:11225–11234 DOI:10.1074/jbc.M511939200.

Yang Y, Sulpice R, Himmelbach A, Meinhard M, Christmann A, Grill E. 2006. Fibrillin expression is regulated by abscisic acid response regulators and is involved in abscisic acid-mediated photoprotection. *Proceedings of the National Academy of Sciences of the United States of America* **103**(15): 6061-6066 DOI:www.pnas.org/cgi/doi/10.1073/pnas.0501720103.

Youssef A1, Laizet Y, Block MA, Maréchal E, Alcaraz JP, Larson TR, Pontier D, Gaffé J, Kuntz M.. 2010. Plant lipid-associated fibrillin proteins condition jasmonate production under photosynthetic stress. *The Plant Journal* **61**(3):436-445 DOI:10.1111/j.1365-313X.2009.04067.x.

Ytterberg AJ, Wijk PKJV. 2006. Protein profiling of plastoglobules in chloroplasts and chromoplasts a surprising site for differential accumulation of metabolic enzymes. *Plant Physiology* **140**(3):984-997 DOI:10.2307/20205663.

Zimmermann P, Hirsch-Hoffmann M, Hennig L, Gruissem W. 2004. GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. *Plant Physiology* **136**(1):2621-2632 DOI:10.1104/pp.104.046367.