Send Orders for Reprints to reprints@benthamscience.net

Current Genomics, 2019, *20*, 371-388

RESEARCH ARTICLE

The TIFY Gene Family in Wheat and its Progenitors: Genome-wide Identification, Evolution and Expression Analysis

Songfeng Xie[1,3,#], Licao Cui[2,#], Xiaole Lei[1], Guang Yang[1], Jun Li[3], Xiaojun Nie[1,*] and Wanquan Ji[1,*]

[1]*State Key Laboratory of Crop Stress Biology in Arid Areas, College of Agronomy and Yangling Branch of China Wheat Improvement Center, Northwest A&F University, Yangling 712100, Shaanxi, China;* [2]*College of Life Science, Jiangxi Agricultural University, Nanchang 330045, Jiangxi, China;* [3]*Key Laboratory of Se-enriched Products Development and Quality Control, Ministry of Agriculture and Rural Affairs, National-Local Joint Engineering Laboratory of Se-enriched Food Development, Ankang R&D Center for Se-enriched Products, Ankang 725000, Shaanxi, China*

ARTICLE HISTORY

Received: June 04, 2019
Revised: August 04, 2019
Accepted: September 27, 2019

DOI:
10.2174/1389202920666191018114557

Abstract: ***Background*:** The TIFY gene family is a group of plant-specific proteins involved in the jasmonate (JA) metabolic process, which plays a vital role in plant growth and development as well as stress response. Although it has been extensively studied in many species, the significance of this family is not well studied in wheat.

***Objective*:** To comprehensively understand the genome organization and evolution of TIFY family in wheat, a genome-wide identification was performed in wheat and its two progenitors using updated genome information provided here.

***Results*:** In total, 63, 13 and 17 TIFY proteins were identified in wheat, *Triticum urartu* and *Aegilops tauschii* respectively. Phylogenetic analysis clustered them into 18 groups with 14 groups possessing A, B and D copies in wheat, demonstrating the completion of the genome as well as the two rounds of allopolyploidization events. Gene structure, conserved protein motif and cis-regulatory element divergence of A, B, D homoeologous copies were also investigated to gain insight into the evolutionary conservation and divergence of homoeologous genes. Furthermore, the expression profiles of the genes were detected using the available RNA-seq and the expression of 4 drought-responsive candidates was further validated through qRT-PCR analysis. Finally, the co-expression network was constructed and a total of 22 nodes with 121 edges of gene pairs were found.

***Conclusion*:** This study systematically reported the characteristics of the wheat TIFY family, which ultimately provided important targets for further functional analysis and also facilitated the elucidation of the evolution mechanism of TIFY genes in wheat and more.

Keywords: Expansion pattern, functional divergence, TIFY gene family, wheat, expression analysis, genome-wide.

1. INTRODUCTION

The TIFY gene family, a plant-specific family of putative transcription factors, is increasingly believed to play vital roles in the control of various biological pathways, including plant growth and development, as well as in response to biotic and abiotic stresses. TIFY proteins possess a conserved core motif (TIF[F/Y]XG) located within an approximately 36-amino acid TIFY domain. Generally, the TIFY gene family can be further divided into four subfamilies, including JAZ, PPD, TIFY and ZML, according to both their distinct domain structures and phylogenetic relationships [1, 2]. In addition to the TIFF domain, the members of the ZML subfamily, including the ZIM and ZML-like (ZML) classes, possess an additional CCT domain and a C2C2-GATA zinc finger domain. Contrastingly, the PPD and JAZ subfamilies contain a conserved JAS motif, while the JAZ proteins have approximately 27 additional amino acids covering the LX2FX2KRX2RX5PY characteristic motif near their C-terminus. At the same time, the PPD subfamily proteins have a unique N-terminal PPD domain and a truncated JAS domain , therefore they did not possess the PY amino acids at the C-terminus [3, 4]. Finally, proteins that only contained the TIFY domain were designated as the TIFY subfamily [3].

Recent studies have shown that Jasmonate (JA) and its bioactive derivatives were the key regulators in diverse aspects of the plant growth and development processes. *ZIM/AtTIFY1* in Arabidopsis was identified as a growth-promoting gene that elongated petioles and hypocotyls [5]. It is reported that *TIFY4a* in Arabidopsis plays a significant role in the synchronization of leaf growth [6]. In rice,

*Address correspondence to these authors at the State Key Laboratory of Crop Stress Biology in Arid Areas, College of Agronomy and Yangling Branch of China Wheat Improvement Center, Northwest A&F University, Yangling 712100, Shaanxi, China; Tel: 086-29-87082984; Fax: 086-29-87082850; E-mails: small@nwsuaf.edu.cn; jiwanquan2008@126.com

#*These authors contributed equally to this work.*

OsTIFY11b/OsJAZ10 was an important regulator in controlling the grain size by taking advantage of the accumulation of carbohydrates in the stem [7]. *OsTIFY3/OsJAZ1* was reported to interact with *OsCOI1b* and *OsMYC2* during the reproductive stage, and was responsible for the regulation of spikelet development [8]. Apart from their function in plant development and hormone responses, TIFY proteins have also been induced by biotic and abiotic stresses [9, 10]. Recent studies reported that *AtTIFY10a*, *10b* and their wild soybean homolog, *GsTIFY10a*, showed positive regulatory roles in plant responses to alkaline stress [11]. It is reported that some of the TIFY genes in rice are regulated by various abiotic stresses, such as drought, salinity, and low-temperature treatment [12]. Most grape TIFY genes are responsive to osmotic, cold, drought, salinity, JA and abscisic acid (ABA) stress [13]. The TIFY genes of the common bean were reported to play a vital role in adaptation to phosphorus (P)-starvation, mediated by JA signaling [14].

As one of the most important cereal crops, wheat (*Tricicum aestivum* L.) provides 20% of the calories consumed worldwide annually and serves as the staple food source for 30% of the entire human population [15-17]. Evolutionarily, wheat has a recent but complex origin, which originated from the hybridization between *Aegilops tauschii* (DD, 2n = 2x = 14) and *Triticum dicoccoides* (AABB, 2n = 4x = 28) approximately 8000 years ago, and the allotetraploid emmer wheat was naturally derived from crossing *Triticum urartu* (genome constitution AA; 2n = 14) and *A. speltoides* (genome constitution SS; 2n = 14) approximately 0.43 MYA (million years ago). As a result, wheat is genetically an allohexaploid species (2n = 6x = 42), with three (A, B and D) homoeologous genomes, making it an ideal model for chromosome interaction and polyploidization studies in plants. To date, extensive studies of the TIFY family have been performed in many plants, such as Arabidopsis [15], rice [12], grape [13], apple [16], wild soybean [17], and *Brachypodium distachyon* [18] as well as rape [19] and Poplar [20]. In wheat, the survey of TIFY family and JAZ subfamily in wheat has also been conducted using the wheat genome version TGACv1 [21, 22]. However, the TGACv1 genome is not the complete and fully annotated wheat reference genome, which could result in the incomplete predictions. The comparative analysis and orthologous pairs of TIFYs in wheat and its diploid progenitors, *T. urartu* (AA) and *A. tauschii*, were also not performed. The newly published fully annotated genome sequences of wheat provided a wealth of information for analyzing the genomic remodeling and evolution dynamics of the TIFY gene family in wheat and its progenitors at the genomic scale [23, 24].

In this study, an *in silico* genome-wide identification and characterization of the TIFY gene family in wheat was performed using the fully updated annotated reference genome (IWGSC v1.1) [23]. A total of 63 TaTIFYs assigned to 18 groups with A, B, and D homoelogous copies were obtained. Then, the gene structure, conserved protein domain, and selection pressure, as well as the phylogenetic relationship were systematically analyzed to reveal the evolutionary and structural significance of these genes. Furthermore, the expansion pattern and functional divergence among the homoeologous copies of the TaTIFY family were also investigated. Taken together, the results detailed in this study shed light on understanding the evolutionary and functional characterization of TaTIFYs and lay the foundation for revealing the regulatory mechanism that TaTIFYs play in wheat development and in response to stresses.

2. MATERIALS AND METHODS

2.1. Identification and Characterization of TIFY Genes from Wheat

The whole-genome sequence of *T. aestivum* was downloaded from the Ensemble plant database(http://plants.ensembl.org/Triticum_aestivum/) (IWGSC v1.1). The hidden Markov model (HMM) profiles of the TIFY domain, JAS, and CCT motifs with the accession numbers PF06200, PF09425 and PF06203, were obtained from the PFAM database (http://pfam.sanger.ac.uk) and HMMER3.0 was further used to search for each domain within the local protein database with an expected value (E-value) of 1e-5. The identified TIFY proteins were subsequently submitted to the web tool SMART and PFAM database to verify the results obtained by the HMMER algorithm. Additionally, all the sequences were confirmed with similarity searches by BLASTN at NCBI against the wheat ESTs (expressing sequence tags). The theoretical isoelectric point (pI), molecular weight (MW) and gravy of *TaTIFYs* were conducted by the ProtParam tool (http://web.expasy.org/protparam/) in the ExPASy server. The cello online tool (http://cello.life.nctu.edu.tw/) was used to predict the subcellular localization. Protein solubility was predicted by PROSOII (http://mips.helmholtz-muenchen.de/prosoII). The genome data of *T. urartu* and *A. tauschii* was downloaded from the *T. urartu* genomic database (http://gigadb.org/dataset/100050) and *A. tauschii* genomic database (http://gigadb.org/dataset/100054), respectively. The TIFY family in these two species was also identified following the method described above.

2.2. Sequence Alignment and Phylogenetic Analysis

Multiple sequence alignment of full-length protein sequences was performed using Clustal X v2.0 [25] with a default parameter. Then, the aligned sequences were visualized using ESPript 3.0 (http://espript.ibcp.fr/ESPript/ESPript/). A phylogenetic tree was created using MEGA6.0 [26] software with the following parameters: A neighbor joining (NJ) statistical method, pairwise deletion and 1000-replicates bootstrap. The homologous copies of wheat the A, B or D sub-genome were defined as genes located in the same branch ends of the phylogenetic tree. Using the same method, the orthologous genes among *T. urartu*, *A. tauschii* and *B. distachyon* were also identified. The genetic variations among and within the species mentioned above were measured using the DnaSP 5.0 tool [27].

2.3. Gene Duplication and Molecular Evolution Analysis

The local reciprocal blast was conducted based on all the *TaTIFY* genes and the gene duplication events were considered using the following criteria: (a) The query coverage > 80% of the longer gene; (b) the similarity of the aligned region was > 80%; (c) only one duplication event was counted for the tightly linked genes [19, 28]. Tandem duplicated genes were defined as adjacent homologous genes on the

same chromosome, with no more than one intervening gene [16]. The rate of Ka (non-synonymous substitution rate)/Ks (synonymous substitution rate) was employed to compare the rates of codon evolution between wheat duplicated gene pairs using the codeml program in the PAML package [29]. The formula T = Ks/2λ was employed to calculate the divergence time, where T referred to the time (MYA) of duplication and divergence time, Ks referred to the synonymous substitutions per site, and λ, for the mutation rate of the divergence of plant's nuclear genes. The λ-value was considered as $\lambda = 6.5\times10^{-9}$ synonymous substitutions per site per year [30, 31].

2.4. Exon-intron Structure, Conserved Motif Composition and Cis-element Analysis

The exon-intron structure and splicing phase of wheat TIFY genes were determined using the Gene Structure Display Server (http://gsds.cbi.pku.edu.cn/) according to the annotation information of the wheat genome. Motifs of the TaTIFY proteins were displayed with the MEME Suite web server [32], using the following parameters: The maximum number of motifs was set at 25 and the optimal width of motifs was set between 5 and 50 residues. To identify the putative cis-regulatory elements in the promoter, the upstream 1.5kb region of the candidate TaTIFY genes were extracted from the *T. aestivum* genome, and subsequently submitted to the PlantCARE database [33].

2.5. Expression Analysis and Interaction Networks Analysis

The available RNA sequencing data was downloaded from the NCBI Sequence Read Archive (SRA: http://www.ncbi.nlm.nih.gov/sra) database. The samples information and accession numbers are listed in Table **S9**. The FPKM value (fragments per kilobase of transcript per million fragments mapped reads) was calculated by the combination of Hisat2 and Stringtie to evaluate the expression profile [34]. The differentially expressed genes were identified with the following threshold values: Fold change (RPKM-tr/RPKM-cont)≥2, FDR-value ≤0.01, and the absolute ratio of log2 (RPKM-tr/ RPKM-cont) ≥1. All FPKM data were finally reported by log2 counts and the heat map was visualized using the R language package. The association and interaction network the putative TIFY genes were involved in were constructed based on the orthologous genes between wheat and Arabidopsis using the online tool STRING 10(http://string-db.org/).

2.6. QRT-PCR Analysis

The plants of wheat cultivar 'Chinese Spring' were grown in a growth chamber at controlled conditions (23 ± 1 °C, 16-h light/8-h dark cycle). One-week-old seedlings which consist with RNA-seq data were used for drought treatments. The plants were incubated in 19.2 % (w/v) PEG-6000 solution for 0, 6, 12, 24 and 48h. Seedlings under the normal condition were used as the control. Leaves of all samples were collected from three to five plants at the time points above under both treated and normal conditions to isolate total RNA by Plant RNA Kit reagent (Omega Bio-Tek, USA) according to the manufacturer's instructions. The RNA integrity was checked by electrophoresis on 1.0% agarose gels stained with ethidium bromide (EB). The first strand cDNAs were synthesized using a Vazyme Reverse Transcription System kit (Vazyme, China) following the manufacturer's protocol. Real-time PCR analyses were performed on the QuantStudio™ 7 Flex System (Thermo Fisher Scientific, USA) with SYBR® Premix Ex Taq™ II (Tli RNaseH Plus) (TaKaRa, Dalian, China). The Elongation Factor 1-Alpha gene was used as internal reference for all the qRT–PCR analysis. Three technical replicates were applied for each treatment. The thermal cycling condition was 95 °C for 30 s followed by 40 cycles of 95 °C for 3 s, 60 °C for 30 s. Primers used in this study are listed in Table **S10**. The expression profile was calculated using the $2^{-\Delta\Delta CT}$ method [35].

3. RESULTS AND DISCUSSION

3.1. Genome-wide Identification of TIFY Genes in Wheat and Two Progenitors

The complete genome information provided the opportunity to understand the genomic composition and organization of the TIFY gene family in wheat. In this study, a total of 63 non-redundant proteins that possessed the complete TIFY domain were identified as TIFY gene family members in wheat using the latest reference genome annotation resource (IWGSCv1.1) (Table **1**). Compared to that 49 TIFYs belonging to 18 groups and only 14 JAZs were identified in wheat using the reference TGACv1 [21, 22], our this study provide a more comprehensive and complete information on the organization and composition of wheat TIFY family. Using the same method, we also detected 13 and 17 TIFY genes from the genome of *T. urartu* and *A. tauschii*, respectively (Tables **S1** and **S2**), suggesting that abundance of TIFY members expanded twice together with some tandem duplications in hexaploide wheat after two allopoly ploidization events occurred in it. Since there is no standard nomenclature, these identified genes were named based on the chromosome localization. Furthermore, the average molecular weight of TIFY proteins in wheat, *A. tauschii* and *T. urartu* was calculated and compared. The results showed that there was no significant difference among them, with an average value of approximately 25 kD. At the same time, the predicted isoelectric point (pI) points showed some variations, ranging from 4.8 to 10.8, 4.6 to 10.6 and 3.9 to 11 for *A. tauschii*, *T. urartu* and *T. aesvitum*, respectively. The strong disparity in pI of TaTIFY gene family demonstrated the rich diversity of the protein structure of this family in wheat, which indicated that they could have differential biological functions. Compared to *T. urartu* (ranging from 82 to 452 with an average of 212 AA) and *A. tauschii* (ranging from 69 to 404 with an average of 241 AA), the amino acid length of wheat TIFY genes ranged from 134 to 409 with an average of 235 AA. Subcellular location prediction was also conducted. The results of this found that 44 *TaTIFYs* were predicted to be located in the nucleus, followed by chloroplast (11), mitochondrion (6), and cytoplasmic space (2) (Table **1**). In addition, all of the wheat TIFY proteins, except *TaJAZ1-B*, *TaJAZ7-B*, *TaJAZ14-A-1* and *TaJAZ14-A-2,* possessed a negative grand average of the hydropathicity (GRAVY) index, suggesting they were most likely hydrophilic. To further confirm the actual existence of these *TaTIFYs*, all the

Table 1. Characteristics of the putative identified TIFY genes in wheat.

Gene ID	Locus_name	Length	pI	MW	Gravy	Sub_cell_location	Soluble	TIFY Subfamily Name	Chromo-some Location	No. of EST Validation
TaJAZ1-A	TraesCS2A02G506500	288	10.19	30711.08	-0.469	Nuclear	soluble/0.679	JAZ	2AL	53
TaJAZ1-B	TraesCS2B02G534800	188	9.14	19723.93	0.187	Nuclear	insoluble/0.450	JAZ	2BL	231
TaJAZ1-D	TraesCS2D02G507200	207	9.14	21852.74	-0.532	Nuclear	soluble/0.855	JAZ	2DL	58
TaJAZ2-A	TraesCS2A02G286700	170	10.3	18304.9	-0.417	Nuclear	insoluble/0.506	JAZ	2AL	18
TaJAZ2-B	TraesCS2B02G303600	169	9.98	18127.64	-0.367	Nuclear	insoluble/0.559	JAZ	2BL	18
TaJAZ2-D	TraesCS2D02G285300	142	10.09	15331.45	-0.368	Nuclear	soluble/0.649	JAZ	2DL	18
TaJAZ3-A	TraesCS2A02G169300	231	9.12	24306.41	-0.477	Nuclear	insoluble/0.598	JAZ	2AS	67
TaJAZ3-B	TraesCS2B02G195600	231	9.28	24274.39	-0.489	Nuclear	soluble/0.608	JAZ	2BS	67
TaJAZ3-D	TraesCS2D02G176800	231	8.9	24274.45	-0.491	Nuclear	insoluble/0.408	JAZ	2DS	67
TaJAZ4-B	TraesCS2B02G264000	159	10.15	16744.64	-0.469	Nuclear	insoluble/0.429	JAZ	2BL	7
TaJAZ4-D	TraesCS2D02G251700	166	11	17416.56	-0.403	Nuclear	insoluble/0.299	JAZ	2DL	4
TaJAZ5-A	TraesCS4A02G008000	175	9.03	17623.04	-0.084	Nuclear	insoluble/0.391	JAZ	4AS	-
TaJAZ5-B	TraesCS4B02G296900	163	9.59	16911.33	-0.201	Nuclear	insoluble/0.297	JAZ	4BL	-
TaJAZ5-D	TraesCS4D02G295800	163	9.2	16919.33	-0.205	Nuclear	insoluble/0.353	JAZ	4DL	1
TaJAZ6-A	TraesCS4A02G007900	198	7.75	20709.48	-0.321	Nuclear	insoluble/0.338	JAZ	4AS	2
TaJAZ6-B	TraesCS4B02G297000	265	9.4	28266.24	-0.388	Mitochondrial	soluble/0.636	JAZ	4BL	2
TaJAZ6-D	TraesCS4D02G295900	210	7.73	21635.51	-0.248	Nuclear	insoluble/0.539	JAZ	4DL	1
TaJAZ7-A	TraesCS4A02G007800	178	9.37	18376.03	-0.075	Chloroplast	insoluble/0.286	JAZ	4AS	4
TaJAZ7-B	TraesCS4B02G297100	189	9.1	19416.28	0.083	Chloroplast	insoluble/0.276	JAZ	4BL	5
TaJAZ7-D	TraesCS4D02G296000	243	10.34	25461.82	-0.311	Chloroplast	insoluble/0.313	JAZ	4DL	1
TaJAZ8-A	TraesCS4A02G249400	187	4.35	20802.94	-0.729	Nuclear	soluble/0.823	JAZ	4AL	-
TaJAZ8-B	TraesCS4B02G065300	185	4.29	20648.81	-0.721	Nuclear	soluble/0.855	JAZ	4BS	-
TaJAZ8-D	TraesCS4D02G064200	242	4.66	26349.22	-0.813	Nuclear	soluble/0.872	JAZ	4DS	2
TaJAZ9-B	TraesCS4B02G364700	360	10.61	38585.36	-0.556	Nuclear	insoluble/0.552	JAZ	4BL	4
TaJAZ9-D-1	TraesCSU02G139000	353	10.64	37668.33	-0.516	Nuclear	insoluble/0.476	JAZ	U(4DL)	4
TaJAZ9-D-2	TraesCSU02G139100	360	9.96	38093.12	-0.402	Nuclear	insoluble/0.381	JAZ	U(4DL)	1
TaJAZ10-A-1	TraesCS5A02G204900	302	10.1	32084.28	-0.443	Nuclear	insoluble/0.443	JAZ	5AL	-
TaJAZ10-A-2	TraesCS5A02G533100	404	10.21	43212.61	-0.492	Nuclear	insoluble/0.576	JAZ	5AL	5
TaJAZ11-A	TraesCS5A02G533000	409	10.09	43254.63	-0.474	Nuclear	insoluble/0.510	JAZ	5AL	61
TaJAZ11-B	TraesCS5B02G203400	371	10.18	39111.04	-0.427	Nuclear	insoluble/0.418	JAZ	5BL	51
TaJAZ11-D	TraesCS5D02G211200	405	10.44	42855.29	-0.446	Nuclear	soluble/0.764	JAZ	5DL	53
TaJAZ12-A	TraesCS5A02G212800	268	9.51	27903.71	-0.383	Nuclear	soluble/0.768	JAZ	5AL	11
TaJAZ12-B	TraesCS5B02G211000	230	7.81	23889.91	-0.404	Nuclear	soluble/0.730	JAZ	5BL	18
TaJAZ12-D	TraesCS5D02G219300	194	6.14	19886.46	-0.325	Nuclear	soluble/0.765	JAZ	5DL	14

(Table 1) contd….

Gene ID	Locus_name	Length	pI	MW	Gravy	Sub_cell_location	Soluble	TIFY Subfamily Name	Chromo-some Location	No. of EST Validation
TaJAZ13-A	TraesCS6A02G293800	330	9.91	36433.3	-0.557	Nuclear	soluble/0.609	JAZ	6AL	2
TaJAZ13-B	TraesCS6B02G324300	331	9.44	36517.39	-0.545	Nuclear	insoluble/0.573	JAZ	6BL	2
TaJAZ13-D	TraesCS6D02G274700	267	9.57	29466.32	-0.594	Nuclear	insoluble/0.481	JAZ	6DL	1
TaJAZ14-A-1	TraesCS7A02G201600	134	9.5	14072.32	0.034	Chloroplast	insoluble/0.366	JAZ	7AS	11
TaJAZ14-A-2	TraesCS7A02G201500	148	9.26	15377.81	0.02	Mitochondrial	insoluble/0.451	JAZ	7AS	27
TaJAZ14-A-3	TraesCS7A02G201300	143	9.5	14996.31	-0.077	Mitochondrial	soluble/0.757	JAZ	7AS	50
TaJAZ14-A-4	TraesCS7A02G201200	143	9.21	15048.2	-0.156	Chloroplast	insoluble/0.491	JAZ	7AS	47
TaJAZ14-B-1	TraesCS7B02G108200	145	9.33	15158.34	-0.199	Chloroplast	insoluble/0.293	JAZ	7BS	43
TaJAZ14-B-2	TraesCS7B02G107700	140	9.84	14821.21	-0.121	Chloroplast	insoluble/0.401	JAZ	7BS	45
TaJAZ14-B-3	TraesCS7B02G107800	147	8.82	15527.89	-0.033	Mitochondrial	insoluble/0.550	JAZ	7BS	59
TaJAZ14-B-4	TraesCS7B02G107900	143	9.13	14977.2	-0.096	Mitochondrial	insoluble/0.412	JAZ	7BS	55
TaJAZ14-B-5	TraesCS7B02G108500	147	9.25	15117.35	-0.02	Chloroplast	insoluble/0.395	JAZ	7BS	51
TaJAZ14-B-6	TraesCS7B02G108000	142	8.98	14856.07	-0.104	Chloroplast	insoluble/0.473	JAZ	7BS	49
TaJAZ14-D-1	TraesCS7D02G204200	147	10.03	15109.38	-0.029	Nuclear	insoluble/0.448	JAZ	7DS	31
TaJAZ14-D-2	TraesCS7D02G204300	134	7.66	14033.17	-0.046	Chloroplast	soluble/0.765	JAZ	7DS	50
TaJAZ14-D-3	TraesCS7D02G204500	147	9.45	15275.64	-0.014	Chloroplast	soluble/0.749	JAZ	7DS	53
TaJAZ14-D-4	TraesCS7D02G204600	140	9.45	14802.2	-0.011	Mitochondrial	insoluble/0.362	JAZ	7DS	42
TaJAZ14-D-5	TraesCS7D02G204700	155	9.22	16114.46	-0.155	Nuclear	insoluble/0.331	JAZ	7DS	9
TaTIFY1-B	TraesCS7B02G323300	261	3.9	28468.86	-0.152	Cytoplasmic	soluble/0.793	TIFY	7BL	-
TaTIFY1-D	TraesCS7D02G415300	329	4.44	36070.85	-0.196	Cytoplasmic	soluble/0.777	TIFY	7DL	1
TaZML1-A	TraesCS7A02G201400	270	6.45	28597.3	-0.47	Nuclear	soluble/0.804	ZML	4AL	2
TaZML1-B	TraesCS7B02G108300	247	8.09	26311.92	-0.385	Nuclear	soluble/0.821	ZML	4BS	1
TaZML1-D	TraesCS7D02G204400	267	6.45	28410.15	-0.478	Nuclear	soluble/0.812	ZML	4DS	3
TaZML2-A	TraesCS6A02G118800	341	5.25	36172.89	-0.729	Nuclear	soluble/0.919	ZML	6AS	3
TaZML2-B	TraesCS6B02G146800	341	5.2	36173.88	-0.729	Nuclear	soluble/0.920	ZML	6BS	3
TaZML2-D	TraesCS6D02G108700	341	5.2	36189.92	-0.713	Nuclear	soluble/0.917	ZML	6DS	3
TaZML3-A	TraesCS7A02G423100	324	4.77	34395.93	-0.661	Nuclear	soluble/0.839	ZML	7AL	5
TaZML3-B	TraesCS7B02G323600	398	4.79	42188.07	-0.791	Nuclear	soluble/0.836	ZML	7BL	5
TaZML3-D	TraesCS7D02G416000	390	4.86	41303.21	-0.77	Nuclear	soluble/0.849	ZML	7DL	5

available wheat ESTs (Expressed Sequence Tag) were used to search them using the BlastN program. Results showed that the 57 out of 63 (90.52%) of the *TaTIFY* genes were supported by the EST hits (Table **1**). Given the limit of available ESTs, the non-supported *TaTIFY* gene might not be detected under specific conditions or low levels of expression that cannot be investigated experimentally [36].

As reported in other plant species, the TIFY gene family generally could be divided into JAZ, TIFY and ZML based on the specific conserved signature domains, namely the TIFY, Jas, GATA and the CCT domains [2]. To support this prediction and classify the obtained wheat TIFY genes, the conserved signature motifs were further investigated. The results showed that all of the TaFIFYs contained the TIFY domain. Out of them, 52 that had both the TIFY domain and Jas motif were assigned as the JAZ subfamily, nine possessed the TIFY domain together with a GATA zinc finger motif and CCT domain and were considered as belonging to the ZML subfamily, and the remaining two proteins containing the TIFY domain motif were considered simply as TIFY subfamily proteins (Fig. **1**). Furthermore, it was found that 100%, 65.1%, 74.6% and 81% of the *TaTIFY* genes have

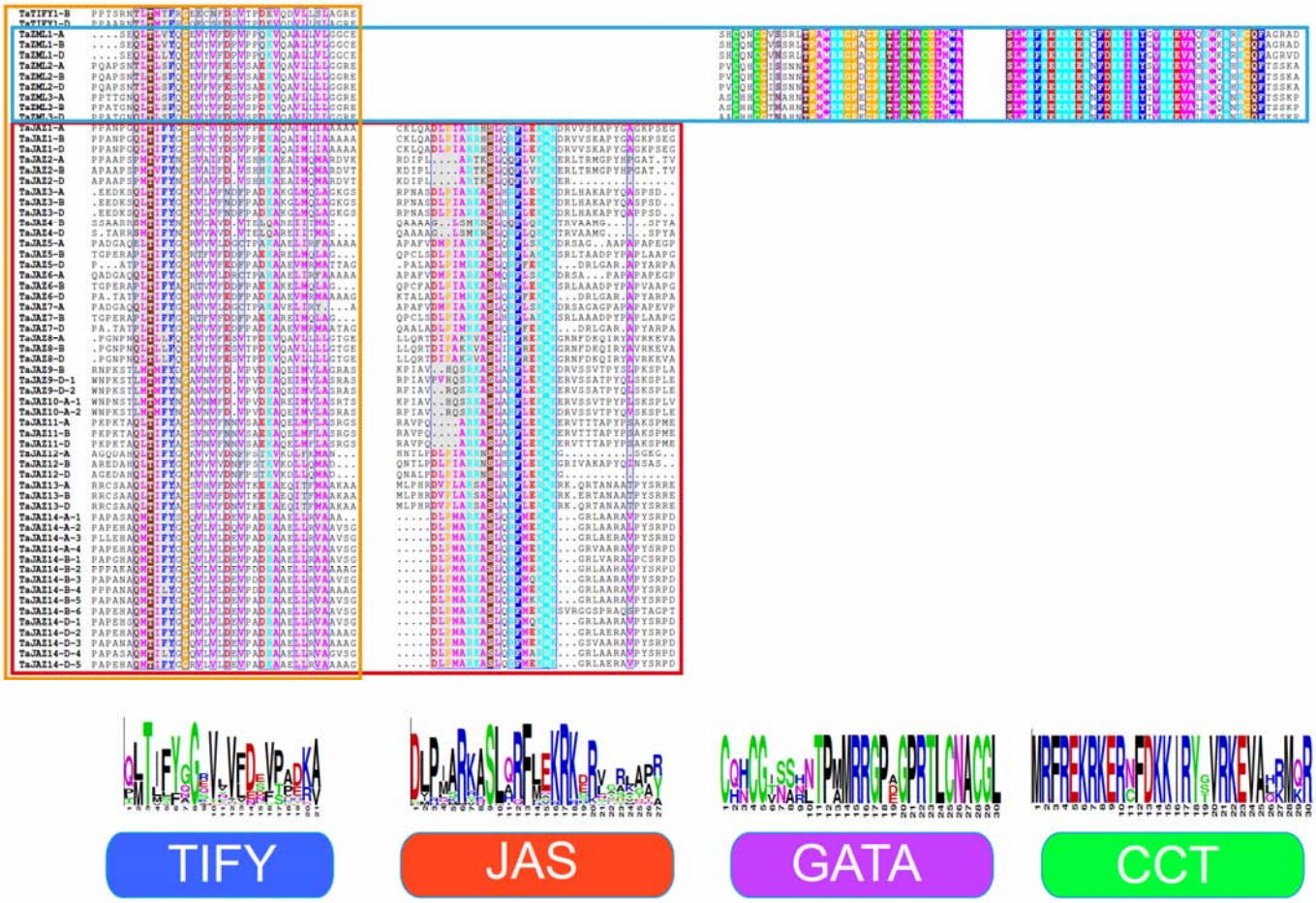

Fig. (1). TIFY domain alignment showing conserved TIFY, JAS, GATA and CCT motif among wheat TIFY proteins. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

characteristic T, I, F and Y residues at the 3rd, 4th, 5th and 6th AA of the TIFY domain, respectively. This means that a variety of TIFY domains combinations could be characteristic such as TIFY[A/G/N/S]G, TLSFQG, TLL[F/Y]QG, TMFY[D/N]G. In particularly, the TIFY domain region was conserved as TMTFRG for the TIFY sub-family. All the isoleucine (I) positions of the TIFY motif in the ZML proteins were replaced by leucine (L) with the consistent sequence conserved as TL[S/V][F/Y]QG. For the Jas motif, the 9th, 10th, 13th, 16th, 17th and 20th AA were S, L, F, K, R and R with percentages, 100%, 98.4%, 100%, 95.1%, 100% and 62.3%, respectively (Fig. **1**). In the C-terminal of the Jas motif, a highly conserved SL(Q/A/H/M)(R/Q)F(L/M/R)(E/Q)KR (K/R)(E/D) R sequence was observed, which was considered as an essential element for the nuclear localization signal of JAZ proteins. The arginine (R) within the Jas motif was supposed for JAZ nuclear localization and interaction with MYC2 [37]. Moreover, the above results revealed that, apart from within the TIF [F/Y] XG region and several other amino acid sites, the TIFY domains possessed diverse variants whereas the Jas motif possessed highly conserved amino acids at the 9th, 10th, 13th, 16th, 17th and 20th AA. Comparatively, the GATA and CCT domain were effectively conserved.

3.2. Phylogenetic Relationships of TaTIFY Proteins

In order to understand the phylogenetic relationships and evolutionary pattern, a phylogenetic tree was generated using the sequence alignments of the 131 full-length TIFY proteins, including 63 from *T. aestivum*, 17 from *A. tauschii*, 13 from *T. urartu*, 18 from Arabidopsis and 20 from rice. All the TIFY proteins were grouped into nine clades, which were named as G1 to G9 based on the Arabidopsis TIFY genes (Fig. **2** and **S1**). Among them, all of the ZML proteins from the five species were clustered together in the G1 clade. G3 and G7 were composed of TIFY proteins that contained only the TIFY domain, and thus were named as group TIFY I and TIFY II. The TIFY I clades included only one wheat protein (*TuTIFY1-AA-1*) grouped with Arabidopsis TIFY proteins while the TIFYII had three proteins (*TaTIFY1-B*, *TuTIFY-AA-2*, *TaTIFY1-D*) grouped with rice TIFY proteins. Notably, the JAZ subfamily can further be divided into six clades (G3, G4, G5, G7, G8 and G9), named as JAZI, JAZII, JAZ-III, JAZIV, JAZV and JAZVI, respectively. Within each clade, two divergent groups related to rice and Arabidopsis were detected, of which the majority of the wheat TIFY genes were more closely related to those of rice than those of Arabidopsis. Among the 52 identified *TaJAZ* proteins, nine of them were clustered within groups JAZ I and JAZ II together with rice JAZ proteins only. The remaining42 *TaJAZ* proteins were grouped into four JAZ subgroups (JAZIII, IV,V and VI), which included JAZ proteins from both rice and Arabidopsis. These results were consistent with the phylogenetic relationship of the species becauseboth wheat and rice were monocots, which had diverged more recently

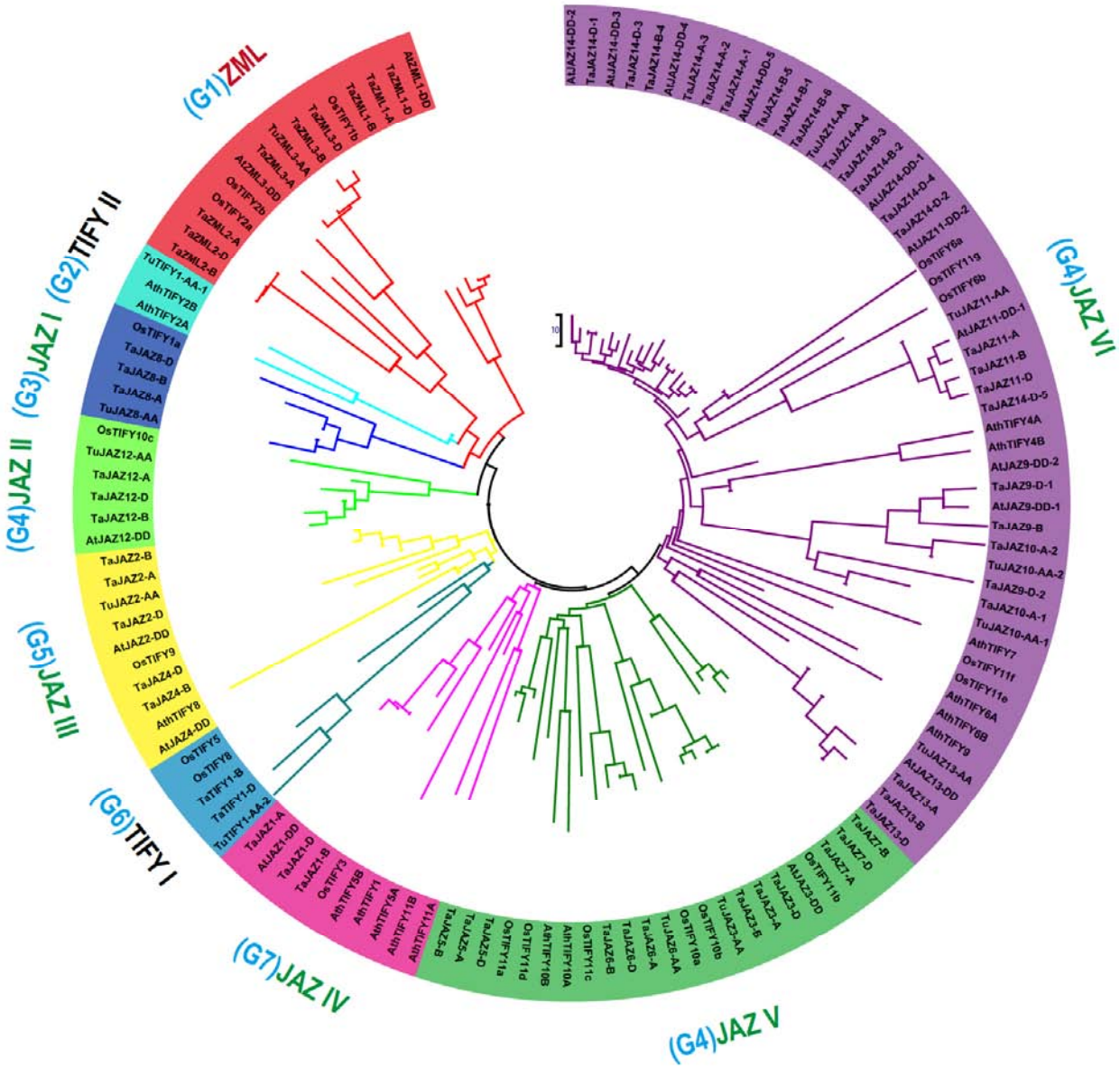

Fig. (2). Phylogenetic relationships among *T.* aestivum*, T. urartu*, *A. tauschii*, *A. thaliana* and *O. sativa* TIFY genes. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

from a common ancestor rather than from the lineage between monocots and eudicots. Additionally, some TIFY subfamily genes from Arabidopsis and rice, for example, *OsTIFY1a*, *OsTIFY8*, *AthTIFY5A* and *AthTIFY5B*, were placed in the JAZ subfamily. In particular, *TaJAZ14* showed less diversity as indicated bya shorter branch length, which suggested that sequence changes tended to be conserved relative to their ancestor [38].

Based on the sequence similarity and phylogenetic tree, the homoeologous group of wheat TIFY genes was further identified. A total of 14 out of 18 groups containing the homoeologous copy on each of the A, B and D homoeologous chromosomes were found in the wheat TIFY family, proving that two rounds of whole genome duplication events had occurred during wheat genome formation [39]. At the same time, the A homoeologous copy of *TaTIFY1* was not found in wheat but was found in *T. urartu*, indicating *TaTIFY1* lost its A homoeologous copy during the allopolyploidization of wheat. For *TaJAZ4*, its A homoeologous copy was not found in wheat, and its orthologous gene in *T. urartu* was also not found, indicating that gene loss events of *TaJAZ4_A* may happen before the formation of wheat. Furthermore, the orthologous pairs between wheat and its progenitors were also identified. Most of the orthologous genes were found, such as the *TaJAZ2, 3, 12, 13* and ZML3. However, gene loss events were also identified, out of which JAZ1, 4, 5 and 7, as well as ZML1 and 2 lost their orthologs in *T. urartu*, while JAZ5-8, ZML2 and TIFY1 were absent from *A. tauschii*. The specific retention and dispersion of homoeologous genes in wheat provided important information forbetter the understanding of the mechanism of chromosome interaction and genetic differentiation as well as gene gain or loss during allopolyploidization [40].

3.3. Expansion Patterns of the Wheat TIFY Gene Family

To determine the effect of Darwinian selection on gene divergence after duplication, the Ka/Ks substitution ratio of the duplicated TIFY genes in wheat and its ancestors was calculated using Brachypodium as the outgroup. The result showed that the Ka/Ks ratios of *TaTIFY*s ranged from 0.0090 to 0.5336 (Table **2**), with an average of 0.2868, suggesting that the wheat TIFY family had undergone purifying selection. The genetic variation of *TaTIFY* genes among the four species mentioned above was also measured using the DnaSP 5.0 software (Table **S3**). The results showed that the Pi value ranged from 0.0613 to 0.3610, with an average value of 0.1806 for *TaTIFY* genes, most of which were lower than 0.3, except for *TaJAZ5*, *TaJAZ7* and *TaJAZ9* with the value of 0.333 (Table **S3**). The nucleotide diversity values of TIFY genes were 0.4873, 0.61944, 0.5420 and 0.5970 in wheat, *T. urartu, A. tauschii* and *B.distachyon*, respectively (Table **S4**). Notably higher genetic variations were observed between species than within species, suggesting that the TIFY paralogous genes tend to be effectively conserved. Furthermore, the *Tajima's* D analysis also found no significant difference from neutral expectations, indicating that the TIFY genes in the same species did not undergo selective pressure, which was consistent with the Ka/Ks analysis.

Segmental and tandem duplications were thought to be two key factors for a gene's family expansion [41]. A total of five JAZ tandem duplication events (*TaJAZ9-D*, *TaJAZ10-A*, *TaJAZ14-A*, *TaJAZ14-B* and *TaJAZ14-D*) were detected (Table **3**), which were mapped to chromosomes 4DL, 5AL, 7AS, 7BS and 7DS, respectively. One round of gene duplication events was observed for TaJAZ9-D and TaJAZ10-A, resulting in two copies for each of them. *TaJAZ14-A*, *TaJAZ14-B* and *TaJAZ14-D*, on the other hand, had more than four copies for each homoeologous gene, suggesting that multiple rounds of tandem duplication had occurred. The Ka/Ks ratios of these 33 tandem duplicated gene pairs ranged from 0.0971 to 0.6914, with an average of 0.3335, whereas the Ka/Ks ratio of the segmental duplication gene pair was 0.6203. Among them, *TaZML1* had a Ka/Ks ratio >1 which indicated that it underwent strong positive selection, while the remaining 32 genes pairs suffered purifying selection pressure as the Ka/Ks ratios of the duplicated repeats were lower than 1. The divergence time was also calculated and the results showed that the average time for the ZML subfamily was about 3.5 MYA, while that for the JAZ subfamily was more than two times larger at 8.13 MYA. During the expansion process of the wheat TIFY gene family, the whole-genome duplication played a significant role in the expansion of the *TaTIFY* gene family, followed by tandem duplication events.

3.4. Gene Structure and Conserved Motif Analysis of TaTIFY Genes

Exon/intron structures of wheat TIFY genes were also investigated. The exon/intron structures of wheat TIFY genes seemed to be relatively conserved within the subfamily but some divergence was observed between different

Table 2. The Ka/Ks ratio of the orthologous pairs of TIFY genes in *T.* aestivum, *T. urartu* and *A. tauschii* between Brachypodium. Brachypodium used as the outgroup.

Gene	A	B	D	AA	DD
TaJAZ1	0.2991	0.4592	0.3983	-	0.3865
TaJAZ2	0.4908	0.4307	0.4355	0.3934	0.2586
TaJAZ3	0.3385	0.3276	0.3643	0.3665	0.3864
TaJAZ4	-	0.2549	0.2162	-	0.1095
TaJAZ5	0.0860	0.0530	0.0432	-	-
TaJAZ6	0.0090	0.0097	0.0094	0.0281	-
TaJAZ7	0.0288	0.0351	0.0186	-	-
TaJAZ8	0.4101	0.4639	0.2542	0.4119	-
TaJAZ9	-	0.0190	0.0196	-	0.0752
TaJAZ10	0.0205	-	-	0.0200	-
TaJAZ11	0.4855	0.4459	0.4468	0.4519	0.2635
TaJAZ12	0.3794	0.4475	0.3621	0.3843	0.3439
TaJAZ13	0.4290	0.4704	0.4961	0.3382	0.3828
ZML1	0.5130	0.4128	0.5336	-	0.5095
ZML2	0.2671	0.2980	0.3031	-	-
ZML3	0.2529	0.2874	0.2943	0.2764	0.4048

Table 3. Ka/Ks value and the divergence time of the duplicated TIFY genes in wheat.

Duplicated Gene Pairs		Ka	Ks	Ka/Ks	Duplication Type	Selection Type	Time (MYA)
TaJAZ9-D	TaJAZ10-A	0.0767	0.1237	0.6203	WGD	Purifying selection	9.52
TaZML1-A	TaZML1-D	0.0145	0.0130	1.1110	WGD	positive selection	1.00
TaZML2-A	TaZML2-B	0.0001	0.0539	0.0010	WGD	Purifying selection	4.15
TaZML2-A	TaZML2-D	0.0043	0.0807	0.0534	WGD	Purifying selection	6.21
TaZML2-B	TaZML2-D	0.0014	0.0318	0.0436	WGD	Purifying selection	2.45
TaJAZ2-A	TaJAZ2-B	0.0186	0.1317	0.1413	WGD	Purifying selection	10.13
TaJAZ2-A	TaJAZ2-D	0.0191	0.0980	0.1951	WGD	Purifying selection	7.54
TaJAZ2-B	TaJAZ2-D	0.0031	0.0678	0.0454	WGD	Purifying selection	5.22
TaJAZ3-A	TaJAZ3-B	0.0137	0.0486	0.2827	WGD	Purifying selection	3.74
TaJAZ3-A	TaJAZ3-D	0.0218	0.0539	0.4051	WGD	Purifying selection	4.15
TaJAZ3-B	TaJAZ3-D	0.0199	0.0916	0.2174	WGD	Purifying selection	7.05
TaJAZ4-B	TaJAZ4-D	0.0623	0.0727	0.8568	WGD	Purifying selection	5.59
TaJAZ5-A	TaJAZ5-B	0.0365	0.2680	0.1362	WGD	Purifying selection	20.62
TaJAZ6-A	TaJAZ6-D	0.0304	0.1106	0.2749	WGD	Purifying selection	8.51
TaJAZ9-B	TaJAZ9-D-1	0.0548	0.0848	0.6459	WGD	Purifying selection	6.52
TaJAZ11-A	TaJAZ11-B	0.0147	0.1185	0.1243	WGD	Purifying selection	9.12
TaJAZ11-A	TaJAZ11-D	0.0459	0.1279	0.3592	WGD	Purifying selection	9.84
TaJAZ11-B	TaJAZ11-D	0.0146	0.0915	0.1597	WGD	Purifying selection	7.04
TaJAZ13-A	TaJAZ13-B	0.0554	0.1035	0.5355	WGD	Purifying selection	7.96
TaJAZ13-A	TaJAZ13-D	0.0106	0.0521	0.2035	WGD	Purifying selection	4.01
TaJAZ13-B	TaJAZ13-D	0.0072	0.0432	0.1675	WGD	Purifying selection	3.32
TaJAZ14-B-2	TaJAZ14-D-4	0.0276	0.0767	0.3602	WGD	Purifying selection	5.90
TaJAZ14-A-3	TaJAZ14-B-4	0.0309	0.1636	0.1886	WGD	Purifying selection	12.58
TaJAZ14-A-4	TaJAZ14-D-2	0.0720	0.2040	0.3528	WGD	Purifying selection	15.69
TaJAZ9-D-1	TaJAZ9-D-2	0.2704	0.3106	0.8705	tandem	Purifying selection	23.89
TaJAZ10-A-1	TaJAZ10-A-2	0.2211	0.3654	0.6051	tandem	Purifying selection	28.11
TaJAZ14-A-1	TaJAZ14-A-2	0.0643	0.2770	0.2323	tandem	Purifying selection	21.31
TaJAZ14-A-1	TaJAZ14-A-3	0.0641	0.3922	0.1633	tandem	Purifying selection	30.17
TaJAZ14-A-1	TaJAZ14-A-4	0.0940	0.3605	0.2608	tandem	Purifying selection	27.73
TaJAZ14-A-2	TaJAZ14-A-3	0.0662	0.1687	0.3925	tandem	Purifying selection	12.98
TaJAZ14-A-2	TaJAZ14-A-4	0.0854	0.1424	0.6000	tandem	Purifying selection	10.95
TaJAZ14-A-3	TaJAZ14-A-4	0.0792	0.2269	0.3490	tandem	Purifying selection	17.45
TaJAZ14-B-1	TaJAZ14-B-2	0.0952	0.1584	0.6013	tandem	Purifying selection	12.18
TaJAZ14-B-1	TaJAZ14-B-3	0.0971	0.2750	0.0971	tandem	Purifying selection	21.15
TaJAZ14-B-1	TaJAZ14-B-4	0.0754	0.2952	0.2553	tandem	Purifying selection	22.71
TaJAZ14-B-1	TaJAZ14-B-5	0.0752	0.1974	0.3809	tandem	Purifying selection	15.18

(Table 3) contd….

Duplicated Gene Pairs		Ka	Ks	Ka/Ks	Duplication Type	Selection Type	Time (MYA)
TaJAZ14-B-1	TaJAZ14-B-6	0.0472	0.0683	0.6914	tandem	Purifying selection	5.25
TaJAZ14-B-2	TaJAZ14-B-3	0.0477	0.2021	0.2361	tandem	Purifying selection	15.55
TaJAZ14-B-2	TaJAZ14-B-4	0.0658	0.2855	0.2304	tandem	Purifying selection	21.96
TaJAZ14-B-2	TaJAZ14-B-5	0.0483	0.1522	0.3176	tandem	Purifying selection	11.71
TaJAZ14-B-2	TaJAZ14-B-6	0.0642	0.1672	0.3842	tandem	Purifying selection	12.86
TaJAZ14-B-3	TaJAZ14-B-4	0.0676	0.1976	0.3422	tandem	Purifying selection	15.20
TaJAZ14-B-3	TaJAZ14-B-5	0.0820	0.2302	0.3561	tandem	Purifying selection	17.71
TaJAZ14-B-3	TaJAZ14-B-6	0.0627	0.2203	0.2848	tandem	Purifying selection	16.95
TaJAZ14-B-4	TaJAZ14-B-5	0.0705	0.2805	0.2511	tandem	Purifying selection	21.58
TaJAZ14-B-4	TaJAZ14-B-6	0.0434	0.2342	0.1852	tandem	Purifying selection	18.02
TaJAZ14-B-5	TaJAZ14-B-6	0.0490	0.1775	0.2760	tandem	Purifying selection	13.65
TaJAZ14-D-1	TaJAZ14-D-2	0.0718	0.2078	0.3458	tandem	Purifying selection	15.98
TaJAZ14-D-1	TaJAZ14-D-3	0.0690	0.1373	0.5023	tandem	Purifying selection	10.56
TaJAZ14-D-1	TaJAZ14-D-4	0.0884	0.2510	0.3520	tandem	Purifying selection	19.31
TaJAZ14-D-1	TaJAZ14-D-5	0.1792	0.4962	0.3612	tandem	Purifying selection	38.17
TaJAZ14-D-2	TaJAZ14-D-3	0.0568	0.1838	0.3090	tandem	Purifying selection	14.14
TaJAZ14-D-2	TaJAZ14-D-4	0.0837	0.3138	0.2667	tandem	Purifying selection	24.14
TaJAZ14-D-2	TaJAZ14-D-5	0.1687	0.4693	0.3594	tandem	Purifying selection	36.10
TaJAZ14-D-3	TaJAZ14-D-4	0.0701	0.2502	0.2803	tandem	Purifying selection	19.25
TaJAZ14-D-3	TaJAZ14-D-5	0.1800	0.4380	0.4109	tandem	Purifying selection	33.69
TaJAZ14-D-4	TaJAZ14-D-5	0.1302	0.4939	0.2637	tandem	Purifying selection	37.99

subfamilies (Fig. **3A** and **3B**). Similar intron/exon patterns and a nearly identical length among each clade was observed. Among TaTIFY genes, the number of introns varied from 0 to 7. The JAZ subfamily presented a more sophisticated structure than the others with a wider variation in intron number. JAZ5 and JAZ7 tended to be intronless, while others had a variable number of introns. The gene structure differentiation between homoeologous genes was also analyzed. Among the 18 homoeologous groups, 10 groups (JAZ3, 4, 5, 6, 7, 9, 12, 13 and ZML1, 2) showed the completely identical exon/intron structure, suggesting they may also have conserved biological functions. The remaining seven groups showed some differences in the exon/intron structure. The structurale differentiation in homoeologous wheat TIFY genes may be related to their differential biological functions. From the perspective of the sub-genome, the A and B sub-genome tend to be more similar compared to the D sub-genome. For example, JAZ2-D lost one exon and JAZ8-D possessedthree more exons compared to the corresponding homoeologous genes in the A and B sub-genome. The increased similarity between the A and B sub-genome as compared to D may be due to their earlier fusion, as well as their longer-lasting genome exchange and recombination [42, 43].

Additionally, the exon-intron structure of orthologous TIFY genes between wheat and its progenitors was analyzed. The results showed that a large difference was found in intron composition and splicing phase of TIFY genes between wheat and its ancestors. For instance, in *A. tauschii*, the first exon and intron were missing in *AtJAZ11* and *AtJAZ13* with regard to *T. urartu*. Large fragments of sequences containing 1 to 3 introns were absent for JAZ3, JAZ6, JAZ11 and JAZ12. Furthermore, the size of the introns also showed huge variations, with more than ten kb of fragments either inserted or missing among them, suggesting that homologous exchange occurred between A, B and D after two rounds of wheat genome fusion. On the other hand, the evolution rate of TIFY introns was much more rapid than that of exons, since intron variations may not significantly influence protein structure and function.

Furthermore, the conserved motifs of these wheat TIFY genes were identified and a total of 15 conserved motifs were identified (Fig. **3C** and Table **S5**). Most of the wheat TIFY proteins were found to possess motif 1 and motif 7, which were thought to be associated with the TIFY domain. Motif 2 and 5 were thought to be relevant to the JAS domain. The other 11 motifs located outside the DNA-binding regions were considered to be either functional factors or

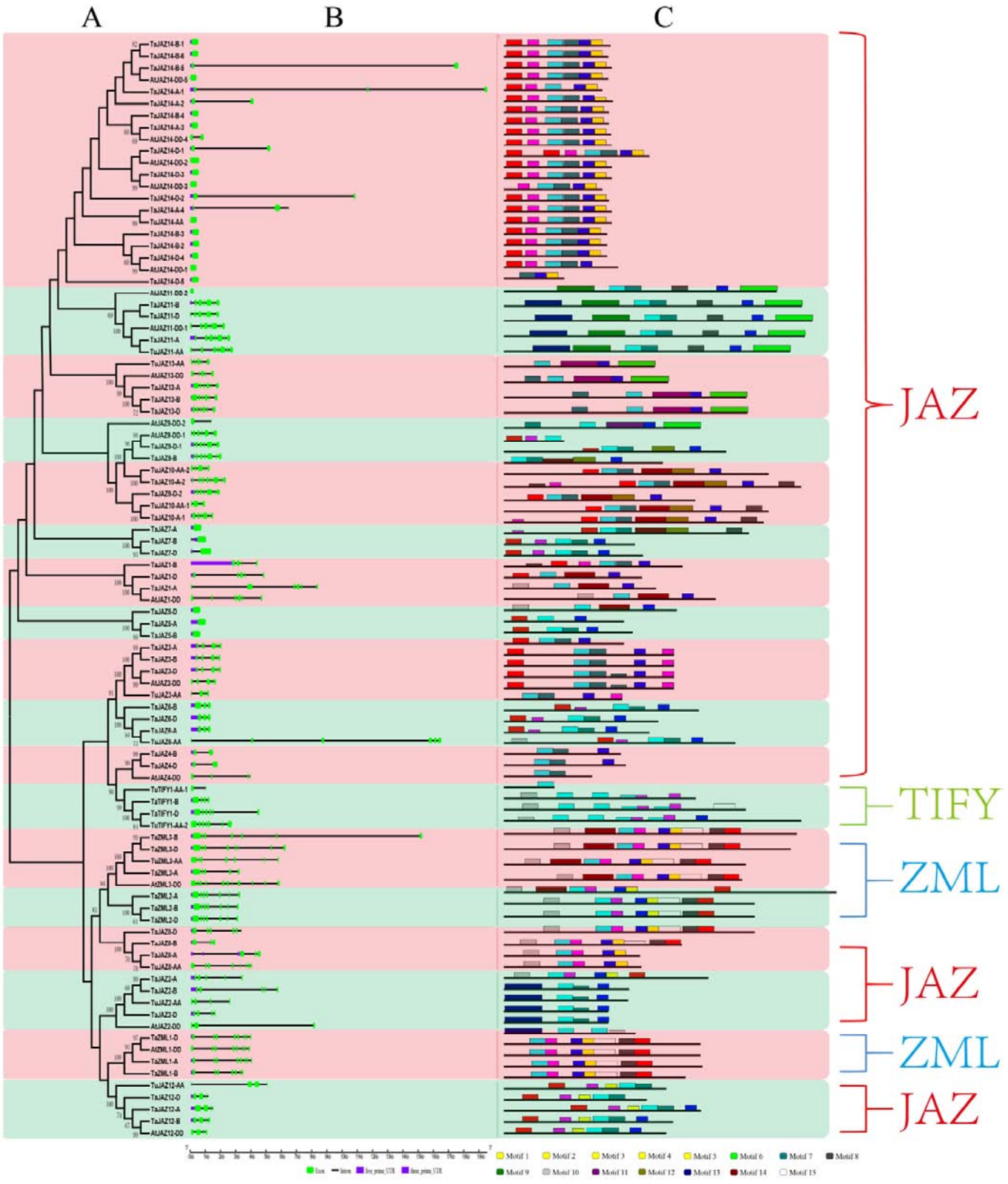

Fig. (3). Phylogenetic relationships (**A**), gene exon-intron structure (**B**) and conserved motif organization (**C**) of TIFY genes in wheat. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

domains related to nuclear localization and transcription regulation. It is noteworthy that proteins clustered into the same group tended to share a similar motif compositions. For instance, Motif 11 was uniquely displayed in the JAZ13 clades and motif 9 was found to be located in JAZ11 clades alone. Motif 12 was shared by JAZ9 and JAZ10. Finally, motif 15 was mostly found in the ZML and TIFY subfamilies. Consistent with the exon/intron structure analysis, motif

analysis revealed that proteins from the A and B sub-genome occupied a higher similarity compared to the D sub-genome. Gain or loss in these motifs, as well as their transformation, may result in a novel function of these genes. In addition, it was found that the similarities of motif composition in wheat homologous groups were higher than those of its progenitors, indicating that motif gain or loss had also occurred between wheat and its progenitors. Compared to wheat, *TuJAZ3-AA*, *TuJAZ8-AA*, and *TuJAZ10-AA* (*T. uratura*) did not possess motif 3, and *AtJAZ2-DD* and *AtJAZ3-DD* (*A. tauschii*) lost motif 7. In addition, motif 15 was lost for *AtZML3-DD*, suggesting that the function or regulation process of ZML3 might have changed during allopolyploidization [44].

3.5. Promoter Analysis of TaTIFYs

Promoter is the DNA region that initiates gene transcription and possessed the transcription factor binding sites, which plays a vital role in regulating spatial and temporal gene expression [45]. In order to understand the transcriptional regulation process of the *TaTIFY* genes, the 1.5 kb upstream sequences from the translation initiation site of these genes were extracted and used to identify the cis-elements. The results showed that three main groups of elements were found in these promoters, namely stress-related elements, hormone-related elements and seed development-related elements (Table **S6**), which was consistent with observations in rice [12] and Brachypodium [18]. Totally, 8 types of motifs were found to be related to stress response, a dehydration responsiveness elements (C-repeat/DRE), heat stress responsiveness element (HSE), low-temperature responsiveness element (LTR), drought-inducibility element (MBS), defense and stress responsiveness element (TC-rich repeats), wound-responsive element (WUN-motif), anaerobic induction element (ARE) and fungal elicitor responsive element (Box-W1). We detected a further 13 motifs that were related to hormone response, including ABA, salicylic acid and gibberellin. Furthermore, six types of plant growth and development motifs were also identified, which comprised of endosperm-specific expression elements (Skn-1 motif, GCN4_motif), circadian control elements (circadian), meristem expression elements (CAT-box, CCGTCC-box) and zein metabolism regulation elements (O2-site). It is noteworthy that most of the *TaTIFY* genes contained the MeJA-responsive elements (52 with TGACG-motif and 51 with CGTCA-motif), suggesting that *TaTIFY*s played a key role in the response to MeJA. The Skn-1 motif, a cis-element required for endosperm expression, was present in most of the *TaTIFY* promoters (60/63). The motifs related to biotic and abiotic stress were found in more than 30 *TaTIFY* promoters whereas the wound-responsive element WUN-motif was present in only eight *TaTIFYs*, suggesting they may play a regulating role in the signaling transduction processes of stress response and tolerance.

3.6. Expression Patterns of TaTIFY Genes at Differential Tissues and Under Stress Conditions

In the present study, the expression profile of 63 *TaTIFYs* in five organs covering the seeding to adult stage was analyzed based on RNA-Seq data mapping (Fig. **4**). All genes in G1 have high expression levels in all of the wheat organs throughout the entire growing period, suggesting the indispensable roles these genes played in regulating growth and development. In total, eight TaTIFY genes in G2 and G4 were predominantly expressed at seed germination. However, 26 genes in group G8 showed very low expression levels in all the tested organs, suggesting that they may express under the special conditions. Furthermore, tissue-specific *TaTIFYs* were also detected (Fig. **4**). A total of five TaTIFYs genes (*TaJAZ8−D*, *TaJAZ11−D*, *TaJAZ12−B* and *TaJAZ14-A-3*) in the G5 clade were specifically expressed in spike. *TaJAZ7−D−1* were uniquely expressed in root, and *TaJAZ12-A* and *TaJAZ12-D* showed preference expression in leaf. Takeing homoeologous genes into consideration, specific patterns were found in different groups. All of the A, B and D homoeologous copies of JAZ3 had higher expression levels in all tested tissues except grain, and A, B, D copies of ZML1 mainly expressed in roots, shoots and leaves. However, JAZ1-D had a high expression level, and JAZ1-B showed a moderate expression level whileJAZ1-A showed no expression in all of the test samples. Meanwhile, the duplicated genes of JAZ14-A were weakly expressed at all stages, while significant differential expression was detected for B and D copies.

To determine the roles that *TaTIFYs* play in response to abiotic stresses, the expression profile of these genes under cold treatment was first investigated (Fig. **5**). A total of 40 *TaTIFYs* genes were detected to express and 11 showed differential expression, of which two genes (*TaJAZ12-D*, *TaJAZ14-B-6*) were down regulated, and the remaining nine showed up-regulated expression. Remarkably, *TaJAZ14-B-6* was down-regulated by more than five times, while *TaJAZ5-D* was up regulated by 27 times. Promoter analysis has shown that the cold-related cis-elements C-repeat/DRE and LTR were found in the promoter regions of these genes, suggesting that cis-elements can regulate the stress-induced expression. Furthermore, the expression patterns of *TaTIFY*s under drought stress were also detected (Fig. **5**). Expression analysis showed that there were two up-regulated and nine down-regulated *TaTIFYs* under drought stress, which could be considered as the drought-responsive candidates. The expression of *TaZML2-B* was six times higher than that of the control, while the expression of *TaJAZ12-B* was 28 times lower than that of control, making them the most up-regulated and down-regulated *TaTIFY* genes, respectively. Regarding the promoter analysis, all of the *TaTIFY* genes responding to drought possessed one or more drought-related cis-elements, such as ABRE and DRE. Apart from *TaJAZ3-A*, which had only an ABRE cis-element, the remaining drought-responsive genes contained at least four cis-elements in the promoter region, suggesting they should be involved in the regulation of the drought response in wheat. Finally, the expression profiles of the *TaTIFY* under heat stress were analyzed (Fig. **5**). A total of 17 *TaTIFYs* were strongly induced by heat stress, of which seven were down-regulated and 10 were up-regulated, respectively. Most of the heat-responsive *TaTIFY* genes had an HSE or TCA-element, but three *TaTIFYs* (*TaJAZ3-A*, *TaJAZ3-D* and *TaZML1-B*) did not contain any of these elements, which suggested they may contain some other unknown elements that play an important role in response to heat stress.

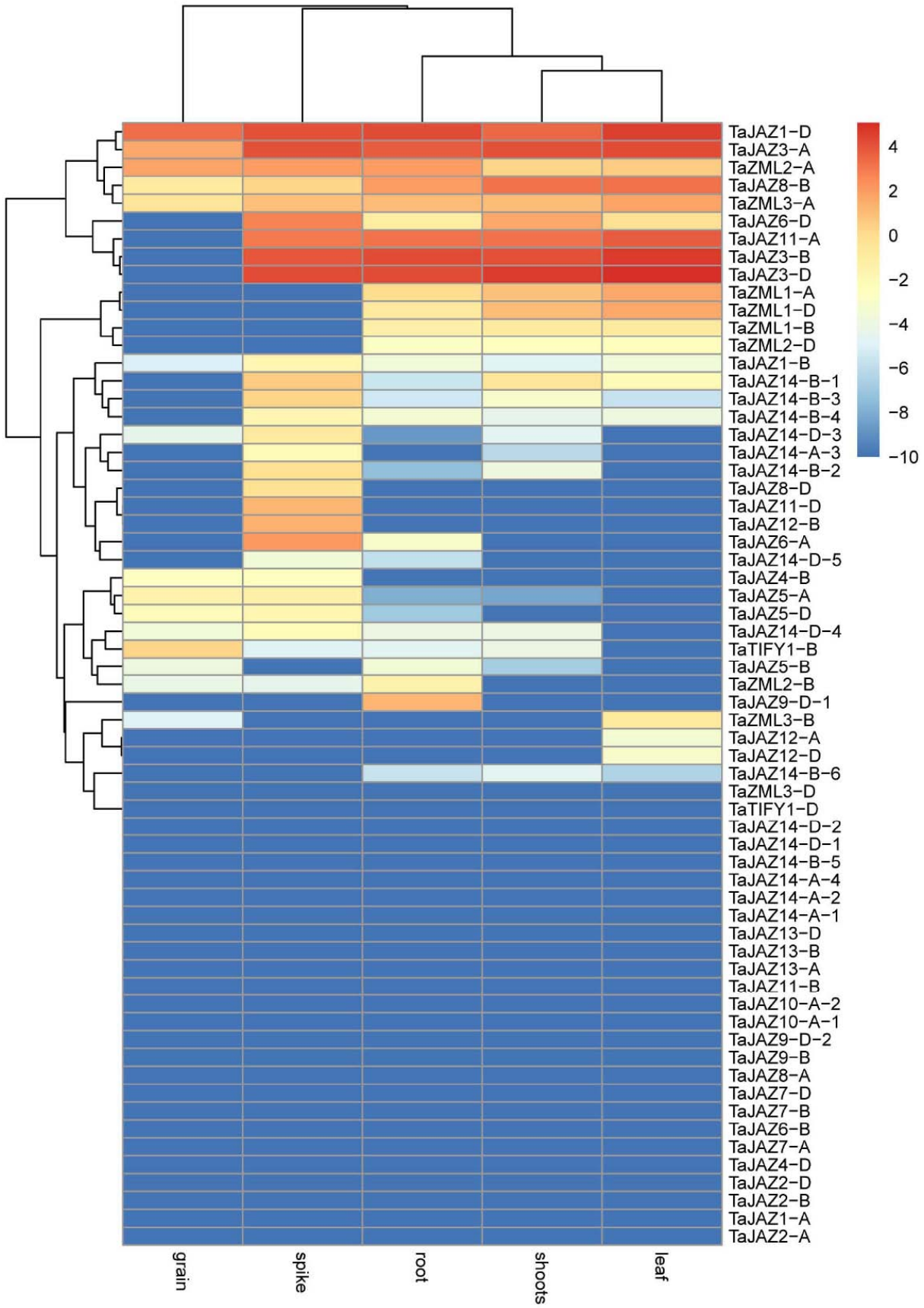

Fig. (4). Hierarchical clustering of expression profiles for *TaTIFY* genes in five tissues (root, shoot, leaf, spike and grain). (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

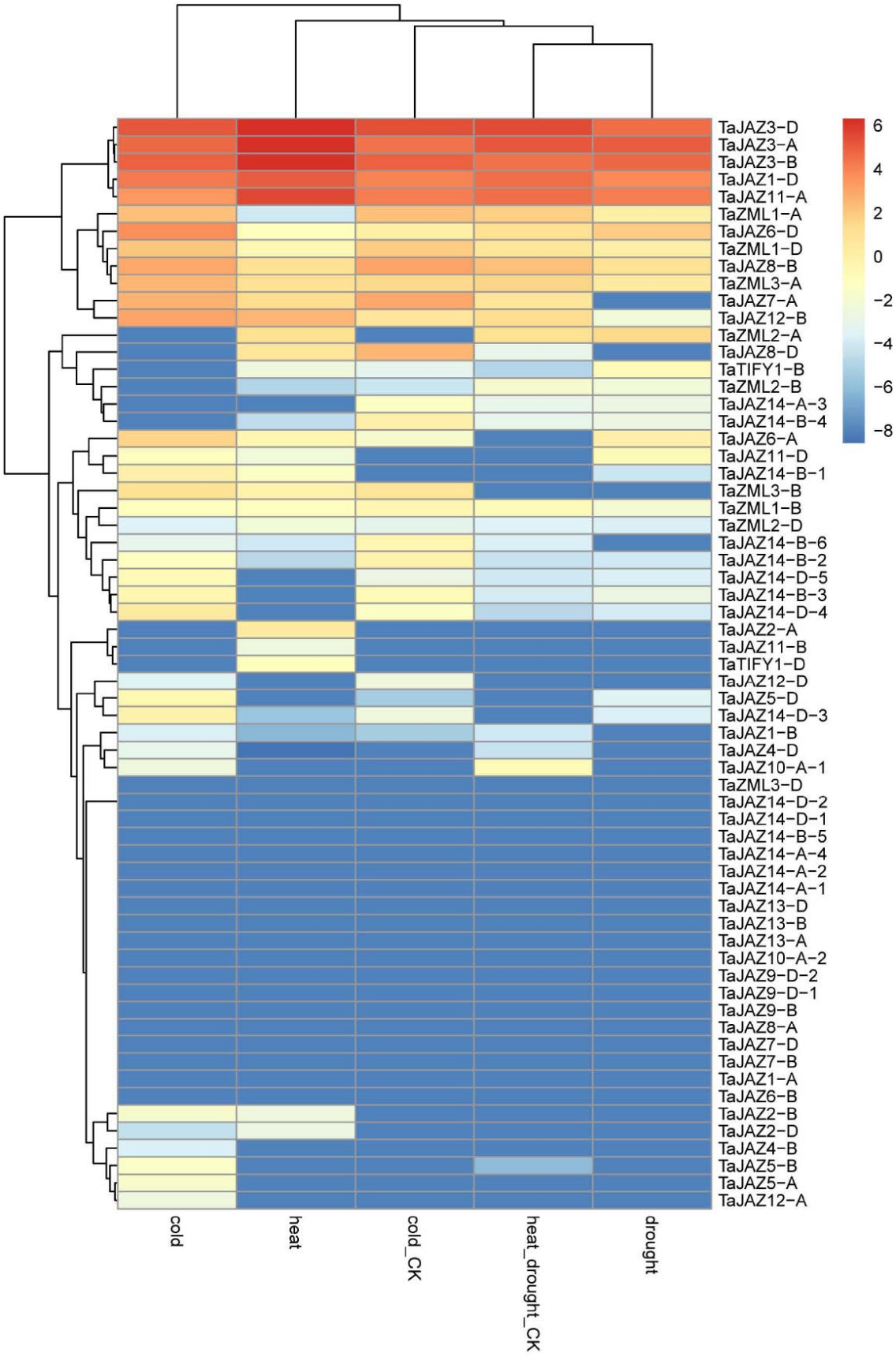

Fig. (5). Hierarchical clustering of expression profiles for *TaTIFY* genes under heat, cold and drought stresses. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

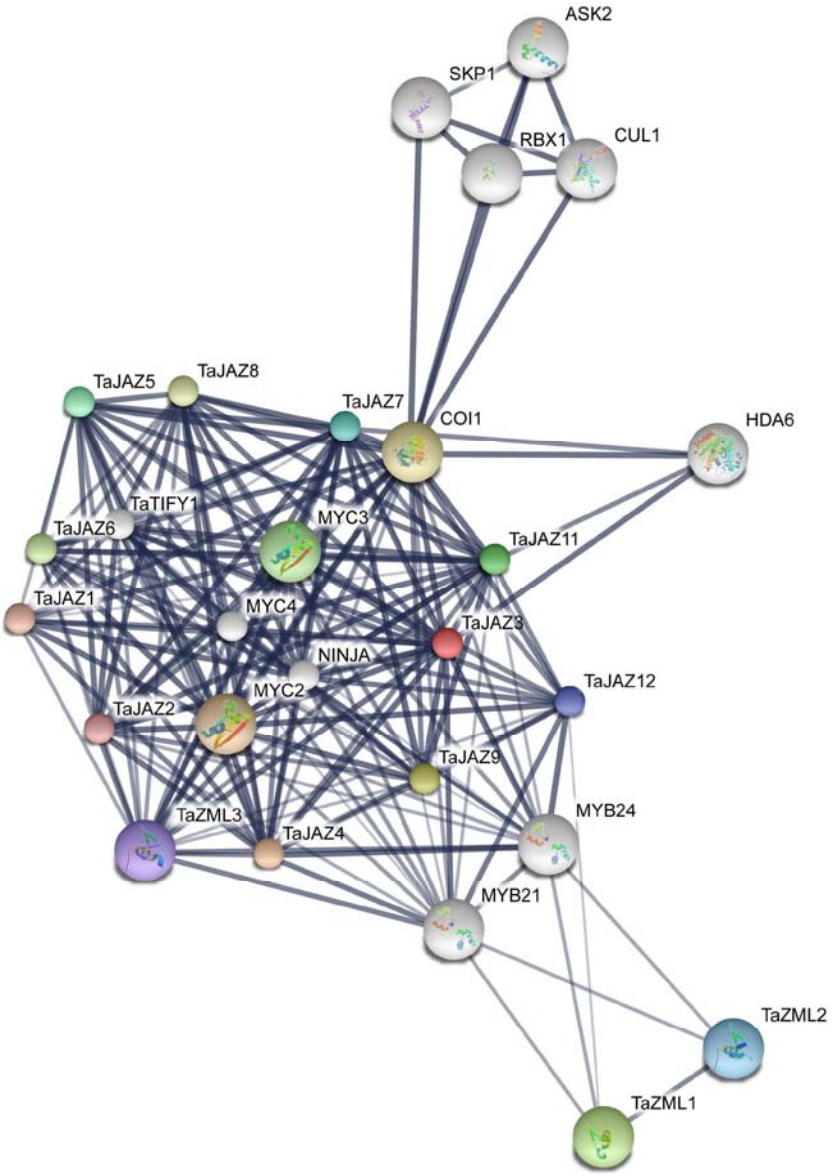

Fig. (6). Interaction network of *TaTIFY* proteins based on the orthologs in Arabidopsis. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

According to the identical expression patterns above, we inferred that function differentiation may have occurred in wheat TIFY genes. Several *TaTIFY* genes were considered to be responsive to multiple stress conditions. *TaJAZ1-B*, *TaJAZ6-D* and *TaJAZ14-B-6* were induced by cold and heat stress. Seven genes (*TaZML1-B*, *TaZML2-B*, *TaJAZ1-D*, *TaJAZ3-A*, *TaJAZ3-B*, *TaJAZ3-D*, *TaJAZ14-B-2*) were induced by heat and drought stress. However, the expression patterns of them were diverse when encountering different stresses. We noticed that all homoeologous genes of *TaJAZ3* were highly activated by heat but significantly repressed by drought. *TaJAZ1-B* and *TaJAZ6-D* were up-regulated under cold treatment and down-regulated under heat treatment. Conversely, *TaJAZ14-B-6* was up-regulated under heat stress but down-regulated under cold stress. We also observed that genes from different subfamilies were preferentially expressed under particular stresses. The ZML subfamily members showed high expression under drought and heat stresses, but displayed almost no expression under cold stress. The expression divergence of A, B, D homoeologous copies was also found. Under heat treatment, *TaZML1-A* and *TaZML1-D* were highly expressed, but the expression of TaZML1-B was repressed. The A and D homoeologous copies of *TaJAZ6* were up-regulated under cold treatment, while *TaJAZ6-B* showed no significant variation. Additionally, three copies of *TaJAZ14* in the D sub-genome (*TaJAZ14-D-3*, *TaJAZ14-D-4* and *TaJAZ14-D-5*) were significantly up-regulated under cold stress, while its B homoeologous gene (*TaJAZ14-B-6*) was down-regulated and no significant differential expression was observed for other duplicated copies of *TaJAZ14*.

To further verify the expression patterns of these stress-responsive TaTIFYs, 4 drought-responsive TaTIFYs were selected for performing QRT-PCR analysis. Results showed that *TaZML2-B* and *TaTIFY1-B* displayed up-regulated expression while TaZML1-A and TaJAZ12-B were down-

regulated under drought stress (Fig. **S2**), which was consistent with the results of RNA-seq data. Among them, *TaZML2-B* expressed continuously with high level when in response to drought stress, showing about 9 times higher than that of control at 6h, and lasting high expression to 48h. *TaJAZ12-B* was found to be low expression at all of four time points. The validated drought-responsive TaTIFYs provided the targets for further functional studies.

3.7. Interactional Network of TaTIFY Proteins

To gain further insight into the interaction and associations between TIFY and other wheat genes, the interaction network of *TaTIFY*s was constructed according to their orthologous with Arabidopsis using the STRING database (Fig. **6**). The *TaTIFY* genes were effectively matched with 12 Arabidopsis TIFY genes. A total of 22 nodes, together with 121 edges of gene pairs of network interactions were detected. GO annotation analysis showed that the interacted genes were involved in various biological processes, cellular components and molecular functions (Table **S7**). For instance, CORONATINEINSENSITIVE1 (COI1), NOVEL INTERACTOROFJAZ (NINJA), CUL1, were identified as the genes that were most similar to *TaTIFY*s, which have been demonstrated to be involved in response to the Jasmonic-acid mediated signaling pathway [46]. KEGG (Kyoto Encyclopedia of Genes and Genomes) analysis also showed that these interacted genes were mostly involved in plant hormone signal transduction and plant-pathogen interaction (Table **S8**). We speculated that these proteins may play an important role in organ development and hormone regulation, as well as in environmental stress response. This is the first study to construct the co-expression network of wheat TIFY genes, which will provide an important foundation to better the understanding of the regulation network and transduction pathway of TIFY genes in wheat and more.

CONCLUSION

In this study, the TIFY gene family was systematically identified and characterized in wheat and its two progenitors to reveal the genome organization, molecular evolution and expression profiles. It was found that *TaTIFYs* had been replicated twice in the two allopolyploidization events of wheat, with tandem and segmental duplications occurring. The Ka/Ks values of the duplicated genes were calculated with the value of lower than 1, suggesting the TIFYs had undergone an evolutionary process of purifying selection and further resulted in a high level of conservation and functional redundancy of homoeologous genes among the sub-genomes. However, functional divergence was also detected among these TaTIFYs as a result of the variations of the gene structure and functional motifs. Finally, the interaction network of the *TaTIFYs* with other wheat genes was constructed to detect the potential metabolic regulatory pathway. To conclude, our current study systematically reported the structure, evolution, and expression, as well as the regulatory network characteristics of the wheat TIFY gene family, which will provide the vital information for a better understanding of the biological functions of wheat TIFYs and shed light on wheat evolution from the TIFY gene family perspective.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No Animals/Humans were used for studies that are the basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

All of the datasets supporting the results of this article are included within the article and its Supplementary files.

FUNDING

This work was mainly funded by the National Key R & D Program of China: chromosome Engineering breeding and New Variety breeding of main crops (Project No. 2016YFD0102004), supported by the Shaanxi Resrarch Station of Crop Gene Resources & Germplasm Enhancement, Ministry of Agriculture, P.R. China, also partially supported by the National Natural Science Foundation of China (Grant No. 31401373) and supported by the Innovation capability support Program Project in Shaanxi Province (Project No. 2018TD-021,2018PT-31).

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

We would like to thank the HPC Center of NWSUAF for providing the compute resource.

SUPPLEMENTARY MATERIAL

Supplementary material is available on the publisher's web site along with the published article.

REFERENCES

[1] Alonso, J.M.; Stepanova, A.N.; Leisse, T.J.; Kim, C.J.; Chen, H.; Shinn, P.; Stevenson, D.K.; Zimmerman, J.; Barajas, P.; Cheuk, R.; Gadrinab, C.; Heller, C.; Jeske, A.; Koesema, E.; Meyers, C.C.; Parker, H.; Prednis, L.; Ansari, Y.; Choy, N.; Deen, H.; Geralt, M.; Hazari, N.; Hom, E.; Karnes, M.; Mulholland, C.; Ndubaku, R.; Schmidt, I.; Guzman, P.; Aguilar-Henonin, L.; Schmid, M.; Weigel, D.; Carter, D.E.; Marchand, T.; Risseeuw, E.; Brogden, D.; Zeko, A.; Crosby, W.L.; Berry, C.C.; Ecker, J.R. Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science,* **2003**, *301*(5633), 653-657. [http://dx.doi.org/10.1126/science.1086391] [PMID: 12893945]

[2] Bai, Y.; Meng, Y.; Huang, D.; Qi, Y.; Chen, M. Origin and evolutionary analysis of the plant-specific TIFY transcription factor family. *Genomics,* **2011**, *98*(2), 128-136. [http://dx.doi.org/10.1016/j.ygeno.2011.05.002] [PMID: 21616136]

[3] Staswick, P.E. JAZing up jasmonate signaling. *Trends Plant Sci.,* **2008**, *13*(2), 66-71. [http://dx.doi.org/10.1016/j.tplants.2007.11.011] [PMID: 18261950]

[4] Chung, H.S.; Niu, Y.; Browse, J.; Howe, G.A. Top hits in contem-

porary JAZ: An update on jasmonate signaling. *Phytochemistry,* **2009**, *70*(13-14), 1547-1559. [http://dx.doi.org/10.1016/j.phytochem.2009.08.022] [PMID: 19800644]

[5] Shikata, M.; Matsuda, Y.; Ando, K.; Nishii, A.; Takemura, M.; Yokota, A.; Kohchi, T. Characterization of Arabidopsis ZIM, a member of a novel plant-specific GATA factor gene family. *J. Exp. Bot.,* **2004**, *55*(397), 631-639. [http://dx.doi.org/10.1093/jxb/erh078] [PMID: 14966217]

[6] White, D.W.R. PEAPOD regulates lamina size and curvature in Arabidopsis. *Proc. Natl. Acad. Sci. USA,* **2006**, *103*(35), 13238-13243. [http://dx.doi.org/10.1073/pnas.0604349103] [PMID: 16916932]

[7] Hakata, M.; Kuroda, M.; Ohsumi, A.; Hirose, T.; Nakamura, H.; Muramatsu, M.; Ichikawa, H.; Yamakawa, H. Overexpression of a rice TIFY gene increases grain size through enhanced accumulation of carbohydrates in the stem. *Biosci. Biotechnol. Biochem.,* **2012**, *76*(11), 2129-2134. [http://dx.doi.org/10.1271/bbb.120545] [PMID: 23132589]

[8] Cai, Q.; Yuan, Z.; Chen, M.; Yin, C.; Luo, Z.; Zhao, X.; Liang, W.; Hu, J.; Zhang, D. Jasmonic acid regulates spikelet development in rice. *Nat. Commun.,* **2014**, *5*(3), 3476. [http://dx.doi.org/10.1038/ncomms4476] [PMID: 24647160]

[9] Devoto, A.; Turner, J.G. Jasmonate-regulated Arabidopsis stress signalling network. *Physiol. Plant.,* **2005**, *123*, 161-172. [http://dx.doi.org/10.1111/j.1399-3054.2004.00418.x]

[10] Wasternack, C. Jasmonates: An update on biosynthesis, signal transduction and action in plant stress response, growth and development. *Ann. Bot.,* **2007**, *100*(4), 681-697. [http://dx.doi.org/10.1093/aob/mcm079] [PMID: 17513307]

[11] Zhu, D.; Li, R.; Liu, X.; Sun, M.; Wu, J.; Zhang, N.; Zhu, Y. The positive regulatory roles of the TIFY10 proteins in plant responses to alkaline stress. *PLoS One,* **2014**, *9*(11), e111984. [http://dx.doi.org/10.1371/journal.pone.0111984] [PMID: 25375909]

[12] Ye, H.; Du, H.; Tang, N.; Li, X.; Xiong, L. Identification and expression profiling analysis of TIFY family genes involved in stress and phytohormone responses in rice. *Plant Mol. Biol.,* **2009**, *71*(3), 291-305. [http://dx.doi.org/10.1007/s11103-009-9524-8] [PMID: 19618278]

[13] Zhang, Y.; Gao, M.; Singer, S.D.; Fei, Z.; Wang, H.; Wang, X. Genome-wide identification and analysis of the TIFY gene family in grape. *PLoS One,* **2012**, *7*(9), e44465. [http://dx.doi.org/10.1371/journal.pone.0044465] [PMID: 22984514]

[14] Aparicio-Fabre, R.; Guillén, G.; Loredo, M.; Arellano, J.; Valdés-López, O.; Ramírez, M.; Iñiguez, L.P.; Panzeri, D.; Castiglioni, B.; Cremonesi, P.; Strozzi, F.; Stella, A.; Girard, L.; Sparvoli, F.; Hernández, G. Common bean (*Phaseolus vulgaris* L.) PvTIFY orchestrates global changes in transcript profile response to jasmonate and phosphorus deficiency. *BMC Plant Biol.,* **2013**, *13*, 26. [http://dx.doi.org/10.1186/1471-2229-13-26] [PMID: 23402340]

[15] Vanholme, B.; Grunewald, W.; Bateman, A.; Kohchi, T.; Gheysen, G. The TIFY family previously known as ZIM. *Trends Plant Sci.,* **2007**, *12*(6), 239-244. [http://dx.doi.org/10.1016/j.tplants.2007.04.004] [PMID: 17499004]

[16] Li, X.; Yin, X.; Wang, H.; Li, J.; Guo, C.; Gao, H.; Zheng, Y.; Fan, C.; Wang, X. Genome-wide identification and analysis of the apple (Malus × domestica Borkh.) TIFY gene family. *Tree Genet. Genomes,* **2015**, *11*, 1-13. [http://dx.doi.org/10.1007/s11295-014-0808-z]

[17] Zhu, D.; Bai, X.; Luo, X.; Chen, Q.; Cai, H.; Ji, W.; Zhu, Y. Identification of wild soybean (*Glycine soja*) TIFY family genes and their expression profiling analysis under bicarbonate stress. *Plant Cell Rep.,* **2013**, *32*(2), 263-272. [http://dx.doi.org/10.1007/s00299-012-1360-7] [PMID: 23090726]

[18] Zhang, L.; You, J.; Chan, Z. Identification and characterization of TIFY family genes in *Brachypodium distachyon*. *J. Plant Res.,* **2015**, *128*(6), 995-1005. [http://dx.doi.org/10.1007/s10265-015-0755-2] [PMID: 26423998]

[19] Saha, G.; Park, J.I.; Kayum, M.A.; Nou, I.S. A genome-wide analysis reveals stress and hormone responsive patterns of TIFY family genes in *Brassica rapa*. *Front. Plant Sci.,* **2016**, *7*, 936. [http://dx.doi.org/10.3389/fpls.2016.00936] [PMID: 27446164]

[20] Xia, W.; Yu, H.; Cao, P.; Luo, J.; Wang, N. Identification of TIFY family genes and analysis of their expression profiles in response to phytohormone treatments and *Melampsora larici-populina* infection in poplar. *Front. Plant Sci.,* **2017**, *8*, 493. [http://dx.doi.org/10.3389/fpls.2017.00493] [PMID: 28424731]

[21] Ebel, C.; BenFeki, A.; Hanin, M.; Solano, R.; Chini, A. Characterization of wheat (*Triticum aestivum*) TIFY family and role of *Triticum Durum* TdTIFY11a in salt stress tolerance. *PLoS One,* **2018**, *13*(7), e0200566. [http://dx.doi.org/10.1371/journal.pone.0200566] [PMID: 30021005]

[22] Wang, Y.; Qiao, L.; Bai, J.; Wang, P.; Duan, W.; Yuan, S.; Yuan, G.; Zhang, F.; Zhang, L.; Zhao, C. Genome-wide characterization of JASMONATE-ZIM DOMAIN transcription repressors in wheat (*Triticum aestivum* L.). *BMC Genomics,* **2017**, *18*(1), 152. [http://dx.doi.org/10.1186/s12864-017-3582-0] [PMID: 28193162]

[23] Appels, R.; Eversole, K.; Feuillet, C.; Keller, B.; Rogers, J.; Stein, N.; Pozniak, C.J.; Stein, N.; Choulet, F.; Distelfeld, A.; Eversole, K.; Poland, J.; Rogers, J.; Ronen, G.; Sharpe, A.G.; Pozniak, C.; Ronen, G.; Stein, N.; Barad, O.; Baruch, K.; Choulet, F.; Keeble-Gagnère, G.; Mascher, M.; Sharpe, A.G.; Ben-Zvi, G.; Josselin, A.A.; Stein, N.; Mascher, M.; Himmelbach, A.; Choulet, F.; Keeble-Gagnère, G.; Mascher, M.; Rogers, J.; Balfourier, F.; Gutierrez-Gonzalez, J.; Hayden, M.; Josselin, A.A.; Koh, C.; Muehlbauer, G.; Pasam, R.K.; Paux, E.; Pozniak, C.J.; Rigault, P.; Sharpe, A.G.; Tibbits, J.; Tiwari, V.; Choulet, F.; Keeble-Gagnère, G.; Mascher, M.; Josselin, A.A.; Rogers, J.; Spannagl, M.; Choulet, F.; Lang, D.; Gundlach, H.; Haberer, G.; Keeble-Gagnère, G.; Mayer, K.F.X.; Ormanbekova, D.; Paux, E.; Prade, V.; Šimková, H.; Wicker, T.; Choulet, F.; Spannagl, M.; Swarbreck, D.; Rimbert, H.; Felder, M.; Guilhot, N.; Gundlach, H.; Haberer, G.; Kaithakottil, G.; Keilwagen, J.; Lang, D.; Leroy, P.; Lux, T.; Mayer, K.F.X.; Twardziok, S.; Venturini, L.; Appels, R.; Rimbert, H.; Choulet, F.; Juhász, A.; Keeble-Gagnère, G.; Choulet, F.; Spannagl, M.; Lang, D.; Abrouk, M.; Haberer, G.; Keeble-Gagnère, G.; Mayer, K.F.X.; Wicker, T.; Choulet, F.; Wicker, T.; Gundlach, H.; Lang, D.; Spannagl, M.; Lang, D.; Spannagl, M.; Appels, R.; Fischer, I.; Uauy, C.; Borrill, P.; Ramirez-Gonzalez, R.H.; Appels, R.; Arnaud, D.; Chalabi, S.; Chalhoub, B.; Choulet, F.; Cory, A.; Datla, R.; Davey, M.W.; Hayden, M.; Jacobs, J.; Lang, D.; Robinson, S.J.; Spannagl, M.; Steuernagel, B.; Tibbits, J.; Tiwari, V.; van Ex, F.; Wulff, B.B.H.; Pozniak, C.J.; Robinson, S.J.; Sharpe, A.G.; Cory, A.; Benhamed, M.; Paux, E.; Bendahmane, A.; Concia, L.; Latrasse, D.; Rogers, J.; Jacobs, J.; Alaux, M.; Appels, R.; Bartoš, J.; Bellec, A.; Berges, H.; Doležel, J.; Feuillet, C.; Frenkel, Z.; Gill, B.; Korol, A.; Letellier, T.; Olsen, O.A.; Šimková, H.; Singh, K.; Valárik, M.; van der Vossen, E.; Vautrin, S.; Weining, S.; Korol, A.; Frenkel, Z.; Fahima, T.; Glikson, V.; Raats, D.; Rogers, J.; Tiwari, V.; Gill, B.; Paux, E.; Poland, J.; Doležel, J.; Číhalíková, J.; Šimková, H.; Toegelová, H.; Vrána, J.; Sourdille, P.; Darrier, B.; Appels, R.; Spannagl, M.; Lang, D.; Fischer, I.; Ormanbekova, D.; Prade, V.; Barabaschi, D.; Cattivelli, L.; Hernandez, P.; Galvez, S.; Budak, H.; Steuernagel, B.; Jones, J.D.G.; Witek, K.; Wulff, B.B.H.; Yu, G.; Small, I.; Melonek, J.; Zhou, R.; Juhász, A.; Belova, T.; Appels, R.; Olsen, O.A.; Kanyuka, K.; King, R.; Nilsen, K.; Walkowiak, S.; Pozniak, C.J.; Cuthbert, R.; Datla, R.; Knox, R.; Wiebe, K.; Xiang, D.; Rohde, A.; Golds, T.; Doležel, J.; Čížková, J.; Tibbits, J.; Budak, H.; Akpinar, B.A.; Biyiklioglu, S.; Muehlbauer, G.; Poland, J.; Gao, L.; Gutierrez-Gonzalez, J.; N'Daiye, A.; Doležel, J.; Šimková, H.; Číhalíková, J.; Kubaláková, M.; Šafář, J.; Vrána, J.; Berges, H.; Bellec, A.; Vautrin, S.; Alaux, M.; Alfama, F.; Adam-Blondon, A.F.; Flores, R.; Guerche, C.; Letellier, T.; Loaec, M.; Quesneville, H.; Pozniak, C.J.; Sharpe, A.G.; Walkowiak, S.; Budak, H.; Condie, J.; Ens, J.; Koh, C.; Maclachlan, R.; Tan, Y.; Wicker, T.; Choulet, F.; Paux, E.; Alberti, A.; Aury, J.M.; Balfourier, F.; Barbe, V.; Couloux, A.; Cruaud, C.; Labadie, K.; Mangenot, S.; Wincker, P.; Gill, B.; Kaur, G.; Luo, M.; Sehgal, S.; Singh, K.; Chhuneja, P.; Gupta, O.P.; Jindal, S.; Kaur, P.; Malik, P.; Sharma, P.; Yadav, B.; Singh, N.K.; Khurana, J.; Chaudhary, C.; Khurana, P.; Kumar, V.; Mahato, A.; Mathur, S.; Sevanthi, A.; Sharma, N.; Tomar, R.S.; Rogers, J.; Jacobs, J.; Alaux, M.; Bellec, A.; Berges, H.; Doležel, J.; Feuillet, C.; Frenkel, Z.; Gill, B.; Korol, A.; van der Vossen, E.; Vautrin, S.; Gill, B.; Kaur, G.; Luo, M.; Sehgal, S.; Bartoš, J.; Holušová, K.; Plíhal, O.; Clark, M.D.; Heavens, D.; Kettleborough, G.; Wright, J.; Valárik, M.; Abrouk, M.; Balcárková, B.; Holušová, K.; Hu, Y.; Luo, M.; Salina, E.; Ravin, N.; Skryabin, K.; Beletsky, A.; Kadnikov, V.; Mardanov, A.;

Nesterov, M.; Rakitin, A.; Sergeeva, E.; Handa, H.; Kanamori, H.; Katagiri, S.; Kobayashi, F.; Nasuda, S.; Tanaka, T.; Wu, J.; Appels, R.; Hayden, M.; Keeble-Gagnère, G.; Rigault, P.; Tibbits, J.; Olsen, O.A.; Belova, T.; Cattonaro, F.; Jiumeng, M.; Kugler, K.; Mayer, K.F.X.; Pfeifer, M.; Sandve, S.; Xun, X.; Zhan, B.; Šimková, H.; Abrouk, M.; Batley, J.; Bayer, P.E.; Edwards, D.; Hayashi, S.; Toegelová, H.; Tulpová, Z.; Visendi, P.; Weining, S.; Cui, L.; Du, X.; Feng, K.; Nie, X.; Tong, W.; Wang, L.; Borrill, P.; Gundlach, H.; Galvez, S.; Kaithakottil, G.; Lang, D.; Lux, T.; Mascher, M.; Ormanbekova, D.; Prade, V.; Ramirez-Gonzalez, R.H.; Spannagl, M.; Stein, N.; Uauy, C.; Venturini, L.; Stein, N.; Appels, R.; Eversole, K.; Rogers, J.; Borrill, P.; Cattivelli, L.; Choulet, F.; Hernandez, P.; Kanyuka, K.; Lang, D.; Mascher, M.; Nilsen, K.; Paux, E.; Pozniak, C.J.; Ramirez-Gonzalez, R.H.; Šimková, H.; Small, I.; Spannagl, M.; Swarbreck, D.; Uauy, C. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science,* **2018**, *361*(6403), 1-13.
[PMID: 30115783]

[24] Luo, M.C.; Gu, Y.Q.; Puiu, D.; Wang, H.; Twardziok, S.O.; Deal, K.R.; Huo, N.; Zhu, T.; Wang, L.; Wang, Y.; McGuire, P.E.; Liu, S.; Long, H.; Ramasamy, R.K.; Rodriguez, J.C.; Van, S.L.; Yuan, L.; Wang, Z.; Xia, Z.; Xiao, L.; Anderson, O.D.; Ouyang, S.; Liang, Y.; Zimin, A.V.; Pertea, G.; Qi, P.; Bennetzen, J.L.; Dai, X.; Dawson, M.W.; Müller, H.G.; Kugler, K.; Rivarola-Duarte, L.; Spannagl, M.; Mayer, K.F.X.; Lu, F.H.; Bevan, M.W.; Leroy, P.; Li, P.; You, F.M.; Sun, Q.; Liu, Z.; Lyons, E.; Wicker, T.; Salzberg, S.L.; Devos, K.M.; Dvořák, J. Genome sequence of the progenitor of the wheat D genome *Aegilops tauschii*. *Nature,* **2017**, *551*(7681), 498-502.
[http://dx.doi.org/10.1038/nature24486] [PMID: 29143815]

[25] Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; Thompson, J.D.; Gibson, T.J.; Higgins, D.G. Clustal W and Clustal X version 2.0. *Bioinformatics,* **2007**, *23*(21), 2947-2948.
[http://dx.doi.org/10.1093/bioinformatics/btm404] [PMID: 17846036]

[26] Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A.; Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.,* **2013**, *30*(12), 2725-2729.
[http://dx.doi.org/10.1093/molbev/mst197] [PMID: 24132122]

[27] Librado, P.; Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics,* **2009**, *25*(11), 1451-1452.
[http://dx.doi.org/10.1093/bioinformatics/btp187] [PMID: 19346325]

[28] Kong, X.; Lv, W.; Jiang, S.; Zhang, D.; Cai, G.; Pan, J.; Li, D. Genome-wide identification and expression analysis of calcium-dependent protein kinase in maize. *BMC Genomics,* **2013**, *14*, 433.
[http://dx.doi.org/10.1186/1471-2164-14-433] [PMID: 23815483]

[29] Goldman, N.; Yang, Z. A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Mol. Biol. Evol.,* **1994**, *11*(5), 725-736.
[PMID: 7968486]

[30] Lynch, M.; Conery, J.S. The evolutionary fate and consequences of duplicate genes. *Science,* **2000**, *290*(5494), 1151-1155.
[http://dx.doi.org/10.1126/science.290.5494.1151] [PMID: 11073452]

[31] Yang, Z.; Gu, S.; Wang, X.; Li, W.; Tang, Z.; Xu, C. Molecular evolution of the CPP-like gene family in plants: Insights from comparative genomics of Arabidopsis and rice. *J. Mol. Evol.,* **2008**, *67*(3), 266-277.
[http://dx.doi.org/10.1007/s00239-008-9143-z] [PMID: 18696028]

[32] Bailey, T.L.; Johnson, J.; Grant, C.E.; Noble, W.S. The MEME Suite. *Nucleic Acids Res.,* **2015**, *43*(W1), W39-49.
[http://dx.doi.org/10.1093/nar/gkv416] [PMID: 25953851]

[33] Lescot, M.; Déhais, P.; Thijs, G.; Marchal, K.; Moreau, Y.; Van de Peer, Y.; Rouzé, P.; Rombauts, S. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. *Nucleic Acids Res.,* **2002**, *30*(1), 325-327.
[http://dx.doi.org/10.1093/nar/30.1.325] [PMID: 11752327]

[34] Pertea, M.; Kim, D.; Pertea, G.M.; Leek, J.T.; Salzberg, S.L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.,* **2016**, *11*(9), 1650-1667.
[http://dx.doi.org/10.1038/nprot.2016.095] [PMID: 27560171]

[35] Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods,* **2001**, *25*(4), 402-408.
[http://dx.doi.org/10.1006/meth.2001.1262] [PMID: 11846609]

[36] Wang, M.; Yue, H.; Feng, K.; Deng, P.; Song, W.; Nie, X. Genome-wide identification, phylogeny and expressional profiles of mitogen activated protein kinase kinase kinase (MAPKKK) gene family in bread wheat (*Triticum aestivum* L.). *BMC Genomics,* **2016**, *17*, 668.
[http://dx.doi.org/10.1186/s12864-016-2993-7] [PMID: 27549916]

[37] Withers, J.; Yao, J.; Mecey, C.; Howe, G.A.; Melotto, M.; He, S.Y. Transcription factor-dependent nuclear localization of a transcriptional repressor in jasmonate hormone signaling. *Proc. Natl. Acad. Sci. USA,* **2012**, *109*(49), 20148-20153.
[http://dx.doi.org/10.1073/pnas.1210054109] [PMID: 23169619]

[38] Altenhoff, A.M.; Studer, R.A.; Robinson-Rechavi, M.; Dessimoz, C. Resolving the ortholog conjecture: Orthologs tend to be weakly, but significantly, more similar in function than paralogs. *PLOS Comput. Biol.,* **2012**, *8*(5), e1002514.
[http://dx.doi.org/10.1371/journal.pcbi.1002514] [PMID: 22615551]

[39] Shewry, P.R. Wheat. *J. Exp. Bot.,* **2009**, *60*(6), 1537-1553.
[http://dx.doi.org/10.1093/jxb/erp058] [PMID: 19386614]

[40] Brenchley, R.; Spannagl, M.; Pfeifer, M.; Barker, G.L.; D'Amore, R.; Allen, A.M.; McKenzie, N.; Kramer, M.; Kerhornou, A.; Bolser, D.; Kay, S.; Waite, D.; Trick, M.; Bancroft, I.; Gu, Y.; Huo, N.; Luo, M.C.; Sehgal, S.; Gill, B.; Kianian, S.; Anderson, O.; Kersey, P.; Dvorak, J.; McCombie, W.R.; Hall, A.; Mayer, K.F.; Edwards, K.J.; Bevan, M.W.; Hall, N. Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature,* **2012**, *491*(7426), 705-710.

[41] Cannon, S.B.; Mitra, A.; Baumgarten, A.; Young, N.D.; May, G. The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biol.,* **2004**, *4*, 10.
[http://dx.doi.org/10.1186/1471-2229-4-10] [PMID: 15171794]

[42] Deng, P.C.; Nie, X.J.; Wang, L.; Cui, L.; Liu, P.; Tong, W.; Biradar, S.S.; Edwards, D.; Berkman, P.; Šimková, H.; Doležel, J.; Luo, M.; You, F.; Batley, J.; Fleury, D.; Appels, R.; Weining, S. Computational identification and comparative analysis of wheat miRNAs in group 7 chromosomes. *Plant Mol. Biol. Report.,* **2014**, *32*, 487-500.
[http://dx.doi.org/10.1007/s11105-013-0669-x]

[43] Akhunov, E.D.; Sehgal, S.; Liang, H.; Wang, S.; Akhunova, A.R.; Kaur, G.; Li, W.; Forrest, K.L.; See, D.; Simková, H.; Ma, Y.; Hayden, M.J.; Luo, M.; Faris, J.D.; Dolezel, J.; Gill, B.S. Comparative analysis of syntenic genes in grass genomes reveals accelerated rates of gene structure and coding sequence evolution in polyploid wheat. *Plant Physiol.,* **2013**, *161*(1), 252-265.
[http://dx.doi.org/10.1104/pp.112.205161] [PMID: 23124323]

[44] Lai, K.; Berkman, P.J.; Lorenc, M.T.; Duran, C.; Smits, L.; Manoli, S.; Stiller, J.; Edwards, D. WheatGenome.info: An integrated database and portal for wheat genome information. *Plant Cell Physiol.,* **2012**, *53*(2), e2.
[http://dx.doi.org/10.1093/pcp/pcr141] [PMID: 22009731]

[45] Kong, F.; Wang, J.; Cheng, L.; Liu, S.; Wu, J.; Peng, Z.; Lu, G. Genome-wide analysis of the mitogen-activated protein kinase gene family in *Solanum lycopersicum*. *Gene,* **2012**, *499*(1), 108-120.
[http://dx.doi.org/10.1016/j.gene.2012.01.048] [PMID: 22306326]

[46] Fernández-Calvo, P.; Chini, A.; Fernández-Barbero, G.; Chico, J.M.; Gimenez-Ibanez, S.; Geerinck, J.; Eeckhout, D.; Schweizer, F.; Godoy, M.; Franco-Zorrilla, J.M.; Pauwels, L.; Witters, E.; Puga, M.I.; Paz-Ares, J.; Goossens, A.; Reymond, P.; De Jaeger, G.; Solano, R. The Arabidopsis bHLH transcription factors MYC3 and MYC4 are targets of JAZ repressors and act additively with MYC2 in the activation of jasmonate responses. *Plant Cell,* **2011**, *23*(2), 701-715.
[http://dx.doi.org/10.1105/tpc.110.080788] [PMID: 21335373]