Characterization of wheat homeodomain-leucine zipper family genes and functional analysis of TaHDZ5-6A in drought tolerance in transgenic Arabidopsis

Shumin Li
Northwest Agriculture and Forestry University
Nan Chen
Northwest Agriculture and Forestry University
Fangfang Li
Northwest Agriculture and Forestry University
Fangming Mei
Northwest Agriculture and Forestry University
Zhongxue Wang
Northwest Agriculture and Forestry University
Xinxiu Cheng
Northwest Agriculture and Forestry University
Zhensheng Kang
Northwest Agriculture and Forestry University
Hude Mao (mhd163com@163.com)
Northwest Agriculture and Forestry University https://orcid.org/0000-0002-2484-9952

Research article

Keywords: HD-Zip gene family, wheat, phylogenetic relationships, expression profiles, TaHDZ5-6A, drought tolerance

Posted Date: January 3rd, 2020

DOI: https://doi.org/10.21203/rs.2.14322/v2

Version of Record: A version of this preprint was published on January 31st, 2020. See the published version at https://doi.org/10.1186/s12870-020-2252-6.

Abstract

Background: Many studies in Arabidopsis and rice have demonstrated that HD-Zip transcription factors play important roles in plant development and responses to abiotic stresses. Although common wheat (Triticum aestivum L.) is one of the most widely cultivated and consumed food crops in the world, the function of the HD-Zip proteins in wheat is still largely unknown. Results: To explore the potential biological functions of HD-Zip genes in wheat, we performed a bioinformatics and gene expression analysis of the HD-Zip family. We identified 113 HD-Zip members from wheat and classified them into four subfamilies (I-IV) based on phylogenic analysis against proteins from Arabidopsis, rice, and maize. Most HD-Zip genes are represented by two to three homeoalleles in wheat, which are named as TaHDZX_ZA, TaHDZX_ZB, or TaHDZX_ZD, where X denotes the gene number and Z the wheat chromosome on which it is located. TaHDZs in the same subfamily have similar protein motifs and intron/exon structures. The expression profiles of TaHDZ genes were analysed in different tissues, at different stages of vegetative growth, during seed development, and under drought stress. We found that most TaHDZ genes, especially those in subfamilies I and II, were induced by drought stress, suggesting the potential importance of subfamily I and II TaHDZ members in the responses to abiotic stress. Compared with wild-type (WT) plants, transgenic Arabidopsis plants overexpressing TaHDZ5-6A displayed enhanced drought tolerance, lower water loss rates, higher survival rates, and higher proline content under drought conditions. Additionally, the transcriptome analysis identified a number of differentially expressed genes between 35S::TaHDZ5-6A transgenic and wild-type plants, many of which are involved in stress response. Conclusions: Our results will facilitate further functional analysis of wheat HD-Zip genes, and also indicate that TaHDZ5-6A may participate in regulating the plant response to drought stress. Our experiments show that TaHDZ5-6A holds great potential for genetic improvement of abiotic stress tolerance in crops.

Background

Changes to the transcriptome are achieved through the action of transcription factors (TFs), which repress or activate suites of genes to modulate plant growth and respond to environmental stimuli [1]. The HD-Zip family consists of a large number of transcription factors that seem be unique to the plant kingdom. HD-Zip proteins contain a Homeobox domain (HD) and an adjacent Leucine Zipper (LZ) motif [2]. The HD domain is responsible for specific DNA binding, whereas the LZ motif acts as a mediator to protein dimerization [2]. Based on the additional conserved motifs and their phylogenetic relationships, HD-Zip genes can be classified into four subfamilies (HD-Zip I, II, III, and IV) [2, 3-5]. All subfamilies cotnain the LZ domain and are characterized by differences in the regions downstream of this domain. HD-Zip II subfamily proteins contain a conserved “CPSCE” motif located in the C-terminus, which is not found in HD-Zip I subfamily proteins [2]. HD-Zip III and IV subfamily proteins uniquely contain the extra conserved START and HD-SAD domains [2]. The HD-Zip III subfamily proteins are distinguished from those of HD-Zip IV by the presence and absence, respectively, of a C-terminal MEKHLA domain [2, 6].

In recent years, many efforts have been made to elucidate the functions of HD-Zip genes. Members of the HD-Zip family have been found to play pivotal roles in plant development and the adaption to environmental stresses. HD-Zip I subfamily proteins are mainly involved in the regulation of organ growth and development, de-etiolation, blue light signaling, and also in regulating the response to abiotic stresses [7-11]. For example, *ATHB7* and *ATHB12* are both sensitive to abscisic acid (ABA) and water deficit, and negatively regulate the ABA response in *Arabidopsis* [6]. *ATHB1* acts as a positive regulator to promote hypocotyl elongation [8] and to mediate the determination of leaf cell fate [9]. The *TaHDZipI-2* gene was shown to regulate flowering and spike development and improve frost tolerance in transgenic barley lines [10]. Additionally, wheat *TaHDZipI-3*, *-4* and *-5* genes are differentially expressed in response to abscisic acid (ABA), cold and drought treatment through binding to specific *cis*-elements [11]. HD-Zip II subfamily proteins participate in embryonic apical development, auxin signaling, and are also involved in light and abiotic stress responses [12-15]. In *Arabidopsis*, both *ATHB2/HAT4* and *HAT2* regulate auxin-mediated morphogenesis, and *ATHB2/HAT4* also mediates the effects of red/far-red light on leaf cell expansion and shade avoidance [13, 14]. *OsHOX11* and *OsHOX27* are two rice HD-Zip II genes, and their expression is dramatically decreased upon exposure to drought in a drought-resistant cultivar [12]. Additionally, a sunflower HD-ZIP II gene, *HAHB10*, participates in the response to biotic stress [15].

HD-Zip III subfamily proteins have been reported to control embryogenesis, apical meristem development, vascular bundle development, morpho-physiological changes in roots and auxin transport, and leaf polarity [16-20]. *ATHB8* and *ATHB15* are thought to direct vascular development [17, 18]. CLV3 has been shown to interact with HD-Zip III members to regulate floral meristem activities [19], and KANADI interacts with HD-Zip III genes to control lateral root development [20]. *PopREVOLUTA(PRE)*, a class III HD-Zip gene in poplar, is involved in the growth of cambia and secondary vascular tissues [16]. HD-Zip IV subfamily proteins are integral to growth and development of trichome, cuticle, and root tissues, as well as epidermal cell differentiation [21-24]. In *Arabidopsis*, *GL2* regulates trichome expansion and root hair differentiation [22], and *PDF2* plays a vital role in epidermal cells to control normal development of the floral organs [21]. *OCL4* (*OUTER CELL LAYER4*) encodes a maize HD-Zip IV transcription factor that inhibits trichome development and influences anther cell division in maize [23]. In addition, recent studies have demonstrated that overexpression of *AtHDG11*, an HD-Zip IV gene, increases drought tolerance in *Arabidopsis*, tobacco, rice, sweet potato, and cotton [24].

Bread wheat (*Triticum aestivum*; 2n = 6x = 42; AABBDD) is an integral global food crop [25, 26]. The modern bread wheat genome is the result of two allopolyploidization events with three genomes. First, the A genome donor (*T. urartu*, AA; 2n = 14) hybridized with the B genome donor (*A. speltoides*, SS; 2n = 14). This event, which occurred ~ 0.2 Mya, produced the allotetraploid *T. turgidum* L. (AABB). Second, this AABB donor hybridized with the D genome donor (*A. tauschii* ~9,000 ya. This resulted in the allohexaploid wheat *T. aestivum* (AABBDD) [27, 28], which has a large (> 17 Gb) and composite genome, making genomic studies difficult. Because of wheat's importance globally, extensive research has been conducted to sequence and annotate its genome [25, 26, 29-33]. Recent efforts have sequenced isolated chromosome arms and constructed a draft sequence of the hexaploid wheat genome (IWGSC, 2018). However, compared with *Arabidopsis* and rice, there are fewer studies of the HD-Zip family in wheat. To

date, only five genes encoding HD-Zip subfamily I members (*TaHDZipI-1* to *TaHDZipI-5*) have been isolated and partially characterized from wheat [10, 11, 34]. Although some studies have focused on the function of wheat *HD-Zip* genes, their genome organization, structure and evolutionary features are not well-understood, especially for those genes involved in the regulation of drought stress.

In a previous study, 46 wheat *HD-Zip* genes were identified [35], which is not consistent with the large genome of wheat. Thus, a further survey of the *HD-Zip* gene family should be conducted using the most current version of the wheat genome. Here, we present a genome-wide identification and analysis of the *HD-Zip* genes from wheat and show the phylogenetic relationships among the wheat genes and to those from *Arabidopsis* and other plants. We performed gene expression analyses to characterize the expression profiles of *HD-Zip* genes in various organs/tissues and in response to drought stress. We then performed functional analysis of a drought- induced *HD-Zip I* gene, *TaHDZ5-6A*, by investigating drought stress tolerance and physiological traits in transgenic *Arabidopsis* plants. Finally, we propose a putative mechanism by which *TaHDZ5-6A* enhances drought tolerance in transgenic *Arabidopsis* plants. Our results provide a basis for the further functional analysis of the wheat *HD-Zip* gene family.

Results

Identification of the HD-Zip gene family in wheat

Wheat genome data used in this study were downloaded from the Chinese Spring IWGSC RefSeq v1.1 reference genome assembly (https://wheat-urgi.versailles.inra. fr/). We firstly converted the wheat genome into a local BLAST database using the UNIX pipeline. Then, we used 90 *Arabidopsis* and rice HD-Zip protein sequences to perform a BLAST search (BLASTP) against this local blast database using cut-off *E*-value < 1e-10. After remove the all redundant sequences using CD-hit program, the rest of protein sequences were further subjected to identify the HD domain and LZ motif using the Simple Modular Architecture Research Tool (SMART; http://smart. embl-heidelberg.de/smart/set_mode.cgi? NORMAL=1). In a recently study, a total of 46 *HD-Zip* genes were identified in wheat by a genome-wide bioinformatic survey [35]. In this study, we further identified 67 additional *HD-Zip* genes in wheat latest genome and extended the total member to 113. Based on the genomic position information, 113 *HD-Zip* genes were located across all the 21 wheat chromosomes, ranging from 3 to 8 per chromosome. Chromosome 5A/B/D have the most HD-Zip genes (24 total, 8 per chromosome), followed by chromosome 4A/B/D (18 total, 6 per chromosome) (Table 1; Additional file 1: Figure S1). The 113 HD-Zip proteins were grouped into 40 clusters based on their phylogenetic relationship. Among them, 39 clusters were assigned to different A, B or D sub-genomes, which were considered as the homoeologous copies of one *HD-Zip* gene. Finally, We designated wheat *HD-Zip* genes as *TaHDZX_ZA*, *TaHDZX_ZB*, or *TaHDZX_ZD*, where X denotes the gene number and Z the wheat chromosome where it is located. The detailed information of HD-Zip family genes in wheat, including nomenclature proposed in the previous study [35] was listed in Table 1. As shown in Table 1, the identified *HD-Zip* genes in wheat encode proteins ranging from 192 (TaHDZ12-6D) to 890 (TaHDZ35-1B) amino acids (aa) in length with an average of 501 aa. Furthermore, the computed molecular weights of these HD-Zip proteins ranged from 20.88 (TaHDZ12-6D) to 96.02

(TaHDZ35-1B) kDa. The theoretical pI of the deduced HD-Zip proteins ranged from 4.59 (TaHDZ5-6A) to 9.79 (TaHDZ12-6D).

Phylogenetic analysis of HD-Zip gene family

Our study aimed to understand the phylogenetic relationships between plant HD-Zip proteins. We began by identification of *HD-Zip* genes from seven other plant species with varying levels of complexity for which entire genomes were accessible: an algae (*Chlamydomonas reinhardtii*), a moss (*Physcomitrella patens*), the monocotyledonous angiosperms *Oryza sativa*, *Zea mays*, and *Brachypodium distachyon*, and the dicotyledonous angiosperms *Arabidopsis thaliana*, *Populus trichocarpa*, and *Vitis vinifera*. From this analysis, we found that the *HD-Zip* gene family seems to be restricted to land plants; all genomes except that of the algae contained genes for HD-Zip proteins. We then used the neighbour-joining phylogenetic tree method with full-length HD-Zip protein sequence alignments from eight plant species to describe their evolutionary relationships. All HD-Zip proteins sorted into four well-conserved subfamilies, HD-Zip I to IV (Fig. 1A). The phylogenetic tree revealed that the plant HD-Zip sequence distribution predominates with species bias (Fig. 1B). HD-Zip I genes generally consisted of the largest subfamilies in the plant species except for *Brachypodium distachyon* and wheat, where HD-Zip II and IV were the largest respectively. In contrast, HD-Zip III genes composed of the fewest numbers of HD-Zip members except for moss (Fig. 1C). Subfamily I included 31 *TaHDZ* genes, grouped into 11 clusters (*TaHDZ1-4A/B/D*, *TaHDZ2-5A/B/D*, *TaHDZ3-4A/B/D*, *TaHDZ4-5A/B/D*, *TaHDZ5-6A/D*, *TaHDZ6-5A/B/D*, *TaHDZ7-2A/B/D*, *TaHDZ8-6A/B/D*, *TaHDZ9-4A/B/D*, *TaHDZ10-2B/D*, and *TaHDZ11-2A/B/D*); Similarly, subfamily II embraces 31 *TaHDZs*, grouped into 12 clusters (*TaHDZ12-6A/B/D*, *TaHDZ13-6A/B/D*, *TaHDZ14-7A/B/D*, *TaHDZ15-1A/B/D*, *TaHDZ16-4B/D*, *TaHDZ17-3B/D*, *TaHDZ18-5A/B/D*, *TaHDZ19-3A/B/D*, *TaHDZ20-1A/B/D*, *TaHDZ21-2A/B/D*, *TaHDZ22-4A*, and *TaHDZ23-7A/D*); While subfamily III is the smallest, and contained 14 *TaHDZs*, which grouped into 5 clusters (*TaHDZ24-3A/B/D*, *TaHDZ25-1A/B/D*, *TaHDZ26-4B/D*, *TaHDZ27-5A/B/D*, and *TaHDZ28-5A/B/D*); subfamily IV contained 36 *TaHDZs*, and grouped into 12 clusters (*TaHDZ29-3A/B/D*, *TaHDZ30-4A/B/D*, *TaHDZ31-5A/B/D*, *TaHDZ32-3A/B/D*, *TaHDZ33-6A/B/D*, *TaHDZ34-7A/B/D*, *TaHDZ35-1A/B/D*, *TaHDZ36-6A/B/D*, *TaHDZ37-2A/B/D*, *TaHDZ38-5A/B/D*, *TaHDZ39-7A/B/D*, and *TaHDZ40-2A/B/D*) (Table 1).

To clarify the paralog and ortholog relationships among this family, we further divided the subfamilies into subclasses. According to the phylogenic tree (Fig. 2), each subfamily contain rice, *Arabidopsis* and wheat *HD-Zip* genes, suggesting that this subfamily in plants were generated before the dicot-monocot split. Consistent with the nomenclature in previous studies of *Arabidopsis* and rice [36], HD-Zip I subfamily was divided into seven subclasses, i.e., α, β, γ, δ, ε, φ and ζ (Fig. 2). Clade ε and φ contains only sequences from *Arabidopsis*. Clade ζ contains sequences from both rice and wheat, while no members of *Arabidopsis* were detected in this subclades, suggesting that *Arabidopsis* lost its members of this group during the long period of evolution. The HD-Zip II subfamily was divided into ten subclasses, from α to κ, according to Hu et al. (2012) [37]. Clade β contains only sequences from *Arabidopsis*. Clade α and γ contains sequences from both rice, wheat, and *Arabidopsis*. While the other clades only contains sequences from rice and wheat. The HD-Zip III subfamily was classified into three subclades, designated

as α, β and γ according to the previous studies [37]. Each clade contains sequences from both rice, wheat, and *Arabidopsis*. The HD-Zip IV subfamily was also divided into six subclades, designated clade α, β, γ, δ, ε and ζ as in a previous study [37]. Clade δ excluded genes from rice and *Arabidopsis*, while clade ζ included only sequences from rice and wheat. Eudicot- and monocot-specific clustering patterns of *HD-Zip genes* emerged when tree topology was examined. This pattern may reflect evolutionary history of these subgroups: *HD-Zip* genes in eudicots were likely retained after they diverged from monocots and then expanded.

Gene structure and motif composition analysis

Exon-intron structural divergence can play an important role in the evolution of multiple gene families [38]. We constructed a phylogenetic tree using only the 113 full-length wheat HD-Zip protein sequences fo further examine patterns in wheat. We found that wheat HD-Zip proteins also fell into the four subfamilies described previously (Fig. 3A). We further mapped the exon/intron organization in the coding regions of each *TaHDZ* gene. Specifically, 21 *TaHDZ* genes had two introns, 28 had three introns, 15 had four introns, two had five introns, two had seven introns, 11 had eight introns, 12 had nine introns, three had 10 introns, five had 11 introns, two had 15 introns, two had 16 introns, and 10 had 18 introns (Fig. 3B, C). In general, orthologous genes are highly conserved with respect to gene structure, and this conservation is sufficient to reveal their evolutionary relationships [38]. In wheat, *HD-Zip* genes within the same subfamily shared similar gene structures (intron number and exon length), especially the members of the HD-Zip I and HD-Zip III subfamilies, i.e, most wheat HD-Zip I genes had two or three introns, and most HD-Zip III genes had 18 introns in their coding regions. Nonetheless, the gene structures in wheat HD-Zip subfamilies II and IV appear to be more variable and display the largest number of exon/intron structural variations, i.e., HD-Zip II members possessed two to four introns, and the number of introns in HD-Zip IV family members varied from 4 to 11 (Fig. 3B, C).

The allohexaploid bread wheat genome is known to have formed by fusion of the *T. urartu* (subgenome A), *Aegilops speltoides* (subgenome B), and *A. tauschii* (subgenome D) genomes prior to several hundred thousand years ago. A majority (60.1-61.3%) of genes in the A, B, and D sub-genomes have orthologs in all the related diploid genomes. To obtain intron gain/loss information for all of the *TaHDZ* genes in the A, B, and D sub-genomes, we also compared the intron/exon structures of the genes that clustered together based on the phylogenetic tree. Among these, fourteen clusters showed changes in their intron/exon structure, including *TaHDZ1-4A/B/D*, *TaHDZ3-4A/B/D*, *TaHDZ5-6A/D*, *TaHDZ10-2B/D*, *TaHDZ12-6A/B/D*, *TaHDZ20-1A/B/D*, *TaHDZ24-3A/B/D*, *TaHDZ25-1A/B/D*, *TaHDZ30-4A/B/D*, *TaHDZ32-3A/B/D*, *TaHDZ35-1A/B/D*, *TaHDZ38-5A/B/D*, *TaHDZ39-7A/B/D*, and *TaHDZ40-2A/B/D* (Fig. 3B). Because there are many orthologs in the wheat A, B, and D sub-genomes, intron gain/loss of these orthologs significantly increases the transcriptome and proteome complexity in wheat.

To further examine the diverse structurse of wheat HD-Zip proteins, the conserved motifs were identified by searching the SALAD database along with subsequent annotation with InterPro (Additional file 2: Figure S2). Seven of these motifs were found to be associated with the functionally defined domains.

Motifs 1 and 2 were referred to the HD domain, which is the typical conserved domain found in the middle of all the TaHDZ proteins, and motif 5 was associated with the adjacent LZ domain. Motifs 17 and 34 were referred to the MEKHLA domain, which is specific to subfamily III, and was found only in subfamily III proteins in wheat (14 members). Motifs 3 and 4 were associated with the START region, which has been identified in subfamily III and IV proteins (Additional file 2: Figure S2). Similar motif compositions are shared by TaHDZ proteins which cluster together, and this indicates that members of a given group possess similar functionalities.

Tissue-specific expression profile of *TaHDZ* genes

Gene family members can exhibit different expression patterns in different tissues to accommodate various physiological processes. To gain insight into the temporal and spatial expression patterns and putative functions of *HD-Zip* genes in wheat growth and development, the tissue-specific expression patterns of the 113 *TaHDZ* genes were investigated using RNA-seq data from 10 different tissues. All *TaHDZ* genes were found to be expressed in at least one of the tissues examined (Fig. 4; Additional file 3: Table S1). Subfamily I *TaHDZ* genes were found to be much more highly expressed in seedling roots, stems, leaves, flag leaves, young spikes, and 5-day-old grains; for example, *TaHDZ1-4A/B/D* are highly expressed in leaves and 5-day-old grains, *TaHDZ8-6A/B/D* are highly expressed in leaves and young spikes (15-days-old), and *TaHDZ11-2A/B/D* are highly expressed in leaves and 5-day-old spikes (Fig. 4; Additional file 4: Figure S3). Subfamily II *TaHDZ* genes are more highly expressed in seedling roots, stems, leaves, flag leaves, and young spikes; for example, *TaHDZ19-3A/B/D* are highly expressed in young spikes, while *TaHDZ20-1A/B/D* are highly expressed in seedling stems, leaves, and 5-day-old spikes (Fig. 4; Additional file 5: Figure S4). Subfamily III *TaHDZ* genes showed relatively higher expression levels in seedling stems, leaves, and young spikes; *TaHDZ24-3A/B/D* are highly expressed in seedling leaves, and *TaHDZ27-5A/B/D* are highly expressed in seedling stems and leaves (Fig. 4; Additional file 6: Figure S5). Subfamily IV *TaHDZ* genes are highly expressed in seedling stems, young spikes, and grains; *TaHDZ29-3A/B/D* are highly expressed in 10-day-old grains, *TaHDZ32-3A/B/D* are highly expressed in 5-20 day-old grains, and *TaHDZ38-5A/B/D* are highly expressed in seedling stems and young spikes (Fig. 4; Additional file 7: Figure S6). Thus, genes in the four wheat HD-Zip subfamilies display obvious differences in expression patterns and levels, which indicates that these genes have undergone functional differentiation and redundancy. It is worth mentioning that most homologous genes show similar expression patterns during development. However, it should also be noted that many clustered expression profiles do not reflect gene similarities, and this includes the copies of individual *HD-Zip* gene types from the sub-genomes. Some of them even show the opposite expression patterns. For instance, *TaHDZ7*, which is located on chromosome 2D, is preferentially expressed in the seedling leaves and flag leaves, whereas the homologous *TaHDZ7* gene from 2A is only expressed in the flag leaves, and the *TaHDZ7* homolog from 2B is preferentially expressed in flag leaves and 5-day-old spikes (Fig. 4; Additional file 4: Figure S3). *TaHDZ37* on 2A shows relatively higher expression in 10-15 day-old grains, while its homologous *TaHDZ37* from 2B is preferentially expressed in seedling leaves and 20-day-old grains, and the homologous from 2D is highly expressed in 15-days-old grains (Fig. 4; Additional file 7: Figure S6). The divergences in expression profiles between homologous genes from the different

subgenomes reveals that some of them may have lost their function or acquired a new function after polyploidization during the evolution of wheat.

Expression patterns of *TaHDZ* genes in response to drought stress

Wheat productivity is severely affected by drought stress, and therefore the study of drought responsive genes is important to increase wheat yield. Many studies have shown that the *HD-Zip* genes play a crucial role in the response to abiotic stresses in plants. To gain more insight into the roles of wheat *HD-Zip* genes in stress tolerance, we first identified the *cis*-elements within 2 kb promoter region using online program PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/). We found a number of *cis*-acting elements related to stress response in the promoter of *TaHDZ*s. They included DRE (Dehydration-responsive element), ABRE (ABA-responsive element), MBS (MYB binding site involved in drought-inducibility), MYC (MYC recognition site), MYB (MYB recognition site), and LTR (low temperature responsive element) (Additional file 8: Table S2). To further understand the potential role of *TaABFs* in the drought stress response, we reanalyzed the expression profiles of all wheat *HD-Zip* genes using RNA-seq data from roots and leaves that were subjected to drought treatment. We found that the wheat *HD-Zip* genes could be mainly classified into two groups based on their expression patterns (Fig. 5A, B; Fig. 6A, B). In leaves, the expression levels of 45 *TaHDZ* genes were up-regulated at one or more time point during drought stress treatment; this included 20 genes from the HD-Zip I subfamily (*TaHDZ2-5A/B/D*, *TaHDZ4-5A/B/D*, *TaHDZ5-6A/D*, *TaHDZ6-5A/B/D*, *TaHDZ7-2A/B/D*, *TaHDZ8-6A/B/D*, *TaHDZ9-4B/D*, and *TaHDZ11-2D*), 19 genes from the HD-Zip II subfamily (*TaHDZ18-5A/B/D*, *TaHDZ20-1A/B*, *TaHDZ16-4A/B/D*, *TaHDZ12-6A/D*, *TaHDZ13-6A/B/D*, *TaHDZ14-7A/B*, *TaHDZ15-1A/B/D*, and *TaHDZ17-3D*), one gene from the HD-Zip III subfamily (*TaHDZ24-3A*), and five genes from the HD-Zip IV subfamily (*TaHDZ29-3A*, *TaHDZ30-4B*, *TaHDZ31-5D*, *TaHDZ37-2A/B*) (Fig. 5A, C, and D). In contrast, 50 *TaHDZ* genes showed down-regulated expression under drought stress, including seven genes from subfamily I, six genes from subfamily II, 12 genes from subfamily III, and 25 genes from subfamily IV (Fig. 5A, C, D). In roots, 34 *TaHDZ* genes were found to be up-regulated in response to drought stress, including 16 genes from subfamily I (*TaHDZ4-5A/B/D*, *TaHDZ6-5A/B*, *TaHDZ7-2A/B/D*, *TaHDZ8-6A/B/D*, *TaHDZ9-4B/D*, and *TaHDZ11-2A/B/D*), 16 genes from subfamily II (*TaHDZ15-1A/B/D*, *TaHDZ16-4A/B/D*, *TaHDZ17-3B*, *TaHDZ19-3A/B/D*, *TaHDZ20-1A/B/D*, *TaHDZ21-2A/B*, and *TaHDZ22-4A*) and two genes from subfamily IV (*TaHDZ37-2B* and *TaHDZ40-2B*) (Fig. 6A, C, D). In contrast, 51 *TaHDZ* genes were down-regulated under drought stress in roots, including 12 genes from subfamily I, 8 genes from subfamily II, 13 genes from subfamily III, and 18 genes from subfamily IV (Fig. 6A, C, D). These results indicate that most *TaHDZ* genes in subfamilies I and II may play important roles in the response to drought stress.

TaHDZ5-6A confers drought tolerance in *Arabidopsis*

The phylogenetic analysis and gene expression profiles suggest that *TaHDZ5-6A/D* may participate in regulating the drought stress response in wheat. Protein sequence analysis revealed that TaHDZ5-6A and TaHDZ5-6D share 95% sequence similarity (Additional file 9: Figure S7). In order to further comfirm the potential role of *TaHDZ5* in the drought stress response, we performed quantitative real-time PCR (qRT-

PCR) using RNA isolated from different tissues and drought conditions. The PCR primers were designed to amplify the homologous alleles of *TaHDZ5*. The results showed that *TaHDZ5* is expressed at higher levels in the seedling leaves, flag leaves and young spikes, with the highest expression detected in the seedling leaves, and *TaHDZ5* was upregulated throughout the testing period by drought stress (Additional file 10: Figure S8). To further investigate the role of *TaHDZ5* in the drought stress response, we generated *35S::TaHDZ5-6A* transgenic *Arabidopsis* lines. Three independent transgenic lines (*OE1*, *OE2*, and *OE3*) were chosen for analysis based on their *TaHDZ5-6A* expression levels (Fig. 7A). WT and *35S::TaHDZ5-6A* transgenic plants were grown for 3 weeks in soil before water was withheld for 14 d. At the early stages of drought treatment, *TaHDZ5-6A* transgenic and WT plants grew normally, with no notable phenotypic differences between them (Fig. 7C). After the drought treatment and six days of rewatering, 72-88% of the *35S::TaHDZ5-6A* plants had survived, whereas only ~ 8% of the WT plants were alive (Fig. 7B, C). Thus, the ectopic of *TaHDZ5-6A* greatly improved drought tolerance in transgenic *Arabidopsis*.

The stomatal apertures of leaves from *35S::TaHDZ5-6A* and WT plants grown in soil were measured. The stomatal aperture indices of of the *OE1*, *OE2*, and *OE3* plants were 0.41, 0.42 and 0.41, respectively, while that of the WT plants was 0.40, when grown under normal conditions (Fig. 7D, E). After being subjected to 10 d of drought stress, the stomatal aperture indices of the *OE1*, *OE2*, and *OE3* plants decreased to 0.22, 0.18, and 0.22, respectively, significantly reduced as compared to that of the WT (Fig. 7D, E). In keeping with these results, after dehydration, detached leaves of *35S::TaHDZ5-6A* transgenic plants lost water much more slowly than those of WT plants (Fig. 7F). These results indicate that the *35S::TaHDZ5-6A* transgenic plants removed water from the soil more slowly than did the WT plants, reducing the rate of wilting. To explore whether *TaHDZ5-6A* ectopic expression influences proline accumulation, we measured the free proline contents in the transgenic and WT plants. After drought treatment, the proline contents of the transgenic lines were much higher than those of the WT plants (Fig. 7G). These findings collectively indicate that *TaHDZ5-6A* can enhance drought tolerance in transgenic *Arabidopsis*.

Global gene expression changes in *35S::TaHDZ5-6A* transgenic *Arabidopsis*

RNA sequencing allowed us to understand how drought tolerance was conferred by the ectopic of *TaHDZ5-6A*. The transcriptome of the *35S::TaHDZ5-6A* transgenic plants was compared to that of WT plants under normal, non-stress conditions. In transgenic plants, a total of 495 and 111 genes were upregulated and downregulated by at least 2-fold ($P < 0.001$, FDR < 0.05) as compared with the WT (Fig. 8A, B; Additional file 11: Table S3). The upregulated genes included genes related to water deprivation, abscisic acid, hormones, and abiotic stimuli, and downregulated pathways included those responsive to auxin stimuli, oxidative stress, and defense responses (Fig. 8C). We then chose 10 genes upregulated in transgenic plants and known to be involved in response to drought: *DREB2A* [39], *RD29A* [40], *RD29B* [40], *RD26* [41], *RD17* [42], *PP2CA* [43], *RAB18* [42], *ANAC019* [44], *NCED3* [45], and *RD20* [46]. We used qRT-PCR to measure their relative expression levels under normal and drought conditions in transgenic and WT plants (Fig. 8D). The results of qRT-PCR were in alignment with those of RNA-seq, indicating that *TaHDZ5-6A* may positively regulate the transcription of these 10 genes, and thereby play a role in the

response, including rapid stomatal closure and reduction of water loss, of transgenic *Arabidopsis* plants under drought conditions.

Discussion

The advent of whole genome sequencing and the availability of global genomic databases have made it possible to examine complex genomes such as wheat in much greater detail [47]. The identification of wheat *HD-Zip* genes is an essential step towards the further functional characterization of these genes. Although the *HD-Zip* gene family has been widely studied in both monocots and dicots, their functions remain obscure in wheat. Previous studies have reported the identification of a few individual *HD-Zip* gene families in wheat, including the *TaHDZipI-1* to *TaHDZipI-5* genes [10, 11, 34]. However, a systematic identification and characterization of wheat *HD-Zip* genes has not been performed until the present. To address this knowledge gap, we performed a comprehensive identification and analysis of wheat *HD-Zip* genes in this study.

We identified 113 putative *HD-Zip* genes in wheat based on the Chinese Spring IWGSC RefSeq v1.1 reference genome (https://wheat-urgi.versailles. inra.fr/) (Table 1). The number of *HD-Zip* genes is twice than that found in *Arabidopsis*, rice, maize, or poplar [3, 4, 37], because wheat is an allohexaploid crop. The 113 HD-Zip protein sequences were further subjected to multiple sequence alignment and phylogenetic analysis. The multiple sequence alignment of wheat HD-Zip proteins clearly showed a high degree of sequence divergence, especially at the C-terminus, revealing the diverse roles that HD-Zip genes play in plant growth and development [2, 48]. The wheat HD-Zip protein family can be further subdivided into four subfamilies (I-IV) based on their relationships with homologous HD-Zip transcription factors in other species (Fig. 1; Fig. 2), gene structures (Fig. 2), and motif arrangements (Additional file 2: Figure S2). Our results are consistent with earlier reports [3, 4, 37]. The HD-Zip III subfamily was found to be the smallest of the four subfamilies in our study (Fig. 1; Fig. 2), which is consistent with previous reports that HD-Zip III is the most conserved subfamily among various species [3, 37]. Also, the number of HD-Zip II and IV subfamily members vary in different species, which is the main reason that there are different numbers of HD-Zip family genes in various species. The gene structure analysis revealed that genes in each HD-Zip subfamily have similar exon-intron structures with respect to numbers and positions (Fig. 3B). However, the HD-Zip II and HD-Zip IV subfamilies were found to be more divergent (Fig. 3C), indicating that these genes may have different functions in wheat development. In addition, wheat HD-Zip proteins contain specific conserved domains in each subfamily (Additional file 2: Figure S2). The HD and LZ domains, which have been reported to be responsible for protein-DNA and for protein-protein interactions, are conserved in all HD-Zip proteins [36]. Except for subfamily-biased conserved motifs, the HD-Zip proteins target different sequences; for example, the HD-Zip I proteins target the CAAT(A/T)ATTG sequence, HD-Zip II proteins interact with the CAAT(C/G)ATTG sequence, and HD-Zip III and IV proteins recognize GTAAT(G/C)ATTAC and TAAATG(C/T)A, respectively [10, 36].

Recently, there have been efforts to understand the functions of *HD-Zip* genes in *Arabidopsis*. To further elucidate the potential functions of wheat *HD-Zip* genes, the orthologous relationships between wheat

and *Arabidopsis* proteins have been examined in depth. Subfamily I is divided into seven subclasses or clades, i.e., α, β, γ, δ, ε, φ, and ζ (Fig. 2). Clade α includes *Arabidopsis ATHB13*, a positive regulator of drought, salinity, and cold stresses [49, 50], that is a ortholog of the wheat *TaHDZ3-4A/B/D* genes. Expression of the β-clade members *ATHB5* and *ATHB6* is also affected by water deficit, and both genes appear to regulate growth in response to ABA and/or drought treatment [51, 52], but these genes have no orthologs in wheat. γ-Clade members are typically induced by abiotic stresses, and include *Arabidopsis ATHB7* and *ATHB12* [7], which are the orthologs of the wheat *TaHDZ5-6A/B/D*, *TaHDZ7-2A/B/D*, and *TaHDZ8-6A/B/D* genes. Furthermore, the δ-clade genes *ATHB21*, *ATHB40*, and *ATHB53* are induced by ABA treatment and salt stress; these three TFs are involved in controlling axillary bud development [53], and are the orthologs of the wheat *TaHDZ9-4A/B/D*, *TaHDZ10-2B/D*, and *TaHDZ11-2A/B/D* genes. The HD-Zip II subfamily is divided into ten subclasses, from α to κ (Fig. 2). Clade γ consists of *ATHB2/HAT4* and *HAT2*, genes that regulate auxin-mediated morphogenesis in *Arabidopsis* [13, 14] and that are the orthologs to wheat *TaHDZ21-2A/B/D*. The HD-Zip III subfamily is classified into three subclades, designated α, β, and γ (Fig. 2). Clade α corresponds to the *REV* clade described in previous studies [54], which are orthologs of the wheat *TaHDZ25-1A/B/D* and *TaHDZ26-4B/D* genes. Clade β includes *Arabidopsis ATHB8* [17] and *ATHB15/CNA* [19], which are the orthologs of wheat *TaHDZ24-3A/B/D*, and clade γ contains PHB and PHV [55, 56], which are the orthologs of wheat *TaHDZ27-5A/B/D* and *TaHDZ28-5A/B/D*. The HD-Zip IV subfamily also consists of six subclades, designated α, β, γ, δ, ε, and ζ (Fig. 2). Clade α contains *Arabidopsis ANL2*, a regulator of anthocyanin accumulation in the leaf sub-epidermal layer and of cell identity in the root [57]; *ANL2* is orthologous to the wheat *TaHDZ36-6A/B/D*, *TaHDZ37-2A/B/D*, and *TaHDZ38-5A/B/D* genes. Clade β includes *Arabidopsis GL2* [22], which is the ortholog of wheat *TaHDZ27-5A/B/D* and *TaHDZ32-3A/B/D*. Clade γ contains trichome formation genes, *HDG4*, *HDG5*, and *HDG8-12* [58], which are the orthologs of wheat *TaHDZ33-6A/B/D*, *TaHDZ34-7A/B/D*, and *TaHDZ35-1A/B/D*. Clade ε is composed of *AtML1* and *PDF2* that are responsible for shoot epidermal cell differentiation [21], and are the orthologs of the wheat *TaHDZ39-7A/B/D* genes. These results will help us to further understand the function of wheat *HD-Zip* genes, especially those that are orthologous with *Arabidopsis HD-Zip* genes.

To better understand the roles of the wheat *HD-Zip* genes during the life cycle of wheat, we performed an expression analysis of publicly-available RNA-seq data in 10 organs/tissues at different developmental stages. Genes in the HD-Zip family have been reported to be involved in the development of different organs, and expression *HD-Zip* genes varies widely in different organs (Fig. 4); for example, HD-Zip I subfamily genes may be involved in flower and leaf development, as it has been shown that they regulate cotyledon, spike, and leaf development [8, 10, 59, 60]. Our results also showed that the HD-Zip I genes are mainly expressed in seedling leaves, flag leaves, and young spikes (Additional file 4: Figure S3). For HD-Zip II genes, most showed higher expression levels in leaves and young spikes in our study (Additional file 5: Figure S4), although HD-Zip II genes have been previously shown to take part in carpel margin, flower development [15, 61], and leaf polarity [62]. We found that HD-Zip III genes were mainly expressed in seedling stems and leaves (Additional file 6: Figure S5), and these genes may have potential functions in organ polarity, vascular development, and meristem function as suggested in previous reports [54, 63].

HD-Zip IV genes in wheat may function in trichome, anther, and grain development as suggested in previous studies [64, 65], because most of them show higher expression levels in seedling stems, young spikes, and during grain development (Additional file 7: Figure S6). These results suggest the *TaHDZ* genes may play a variety of roles in wheat development.

The responsiveness of HD-Zip genes to stress strongly suggests that they play roles in developmental adaptation to changing environmental conditions. As in other plants, the expression of a subset of the HD-Zip family I and II genes is either induced or repressed by drought stress (Fig. 5; Fig. 6). Our qRT-PCR analyses further revealed that a noval HD-Zip I gene, *TaHDZ5* is highly expressed in seedling leaves and is induced by drought stress (Additional file 10: Figure S8). To investigate the role of *TaHDZ5* in the abiotic stress response, a homologous gene *TaHDZ5-6A* was transformed into *Arabidopsis*, and its expression was confirmed by qRT-PCR (Fig. 7A). The transgenic plants showed significantly improved drought resistance compared with WT plants (Fig. 7B, C). Adverse environmental factors often cause physiological changes in plants, and physiological indices can be used to evaluate the abiotic stress resistance of crops. Our analysis of the transpiration rate of detached leaves showed that the rate of water loss in the transgenic plants was lower than in WT plants (Fig. 7F). Consistently, we found that the stomata closed faster in transgenic plants than in WT plants under drought conditions (Fig. 7D, E). In addition, the proline content was higher in the transgenic plants compared to WT plants under drought conditions (Fig. 7G). We also found that constitutive expression of *TaHDZ5-6A* in *Arabidopsis* significantly increased the transcription of many stress-responsive genes, including *RD29A*, *RD29B*, *RAB18*, *DREB2A*, *NCED3*, and *RD17* (Fig. 8). These data provide strong evidence that *TaHDZ5-6A* can enhance drought tolerance in the transgenic arabidopsis plants. Previous studies have reported that overexpression of TF genes may cause the growth retardation in transgenic plants [66-68], restricting the applicability of target genes in transgenic breeding. The growth and morphological features of transgenic plants expressing *TaHDZ5-6A* were closely monitored, and no obvious adverse effects were observed (Fig. 7A), indicating the potential for using *TaHDZ5-6A* in plant breeding.

Conclusions

In conclusion, we performed a comprehensive analysis of the genome organization, evolutionary relationships, and expression profiles of the HD-Zip gene family members in wheat and also functionally characterized *TaHDZ5-6A* by showing that it confers drought tolerance in transgenic *Arabidopsis*. The present study has built a foundation and provides an essential framework for the further functional characterization of wheat *HD-Zip* genes in various physiological processes, including their role and the underlying molecular mechanism in the regulation of drought tolerance in wheat.

Methods

Plant materials and drought stress treatments

Wheat (*Triticum aestivum*) variety Chinese spring (CS) was identified and obtained from the Prof. Zhensheng Kang's Lab (Northwest A&F University, China) and was used to analysis the expression of *TaHDZ* genes, this wheat variety can also obtained from Chinese Crop Germplasm Resources Information System (http://www.cgris.net/ zhongzhidinggou/index.php). After surface-sterilized with 75% ethanol and washed with deionized water, the seeds were germinated on wet filter paper at 25 ^{o}C for 3 days. The germinated seeds were placed in a nutrient solution (0.1 mM KCl, 0.75 mM K_2SO_4, 0.65 mM $MgSO_4$, 0.25 mM KH_2PO_4, 1.0 mM $MnSO_4$, 1.0 mM $ZnSO_4$, 0.1 mM EDTA-Fe, 2.0 mM $Ca(NO_3)_2$, 0.005 mM $(NH_4)_6Mo_7O_{24}$, 0.1 mM $CuSO_4$) for hydroponic cultivation with a 16/8 h light/dark cycle at 16 ^{o}C in a growth chamber. For drought treatment, the three-leaf stage seedlings were placed on a clean bench and subjected to dehydration (25 ^{o}C, relative humidity of 40-60%). Leaves and roots from three seedlings were collected after 0, 1, 3, 6, 12 and 24 hr for drought treatment.

To investigate the tissue-specific expression patterns of *TaHDZ5-6A* in wheat, field grown wheat cv. Chinese spring were used. Wheat plants were grown during the growing season at the experimental station of the Northwest A & F University, Yangling, Shanxi, China (longitude 108°E, latitude 34°15'N) from 2016 to 2017. Ten tissue/organ samples including root, stem, leaf of wheat seedling at five-leaf stage, young spike at early booting stage, spike at heading stage, flag leaf at heading stage, and the grain of 5, 10, 15 and 20 DPA. Each sample was collected from at least five individual plants for two repeats. All collected samples were immediately frozen into liquid nitrogen and stored at -80 ^{o}C for RNA extraction.

Identification and annotation of *HD-Zip* genes in wheat

The HD-Zip domain (PF00046) was downloaded from Pfam (http://pfam.xfam.org/) and then used for identification of the *HD-Zip* genes from the Chinese Spring IWGSC RefSeq v1.0 reference genome assembly (https://wheat-urgi.versailles.inra.fr/) using HMMER3.1 [69](*E*-value < 1e-10). After remove all redundant sequences using CD-hit program, the rest of protein sequences were further subjected to identify the HD and LZ domains using the Simple Modular Architecture Research Tool (SMART; http://smart.embl-heidelberg.de/smart/set_mode.cgi?NORMAL=1). We further filtered these genes through phylogenetic analysis along with previously identified HD-Zip proteins from *Arabidopsis thaliana*, *Vitis vinifera*, *Populus trichocarpa*, *Brachypodium distachyon*, *Oryza sativa*, *Zea mays*, and *Physcomitrella patens* [37]. Phylogenetic analysis was also implemented to categorize different HD-Zip subfamilies. Homeologous genes from each of the three wheat subgenomes (A, B, and D genomes) were named *TaHDZX_ZA*, *TaHDZX_ZB*, or *TaHDZX_ZD*, where X denotes the gene number and Z the wheat chromosome where it is located. The theoretical pI (isoelectric point) and Mw (molecular weight) of each putative wheat HD-Zip protein was calculated using compute pI/Mw tool online (http://web.expasy.org/compute_pi/).

Phylogenetic analysis and conserved protein motif/domain identification

Multiple sequence alignments were generated using the ClustalW program with the default settings [70]. To investigate the evolutionary relationship among HD-Zip proteins, an unrooted phylogenetic tree was

obtained by neighbor-joining (NJ) method using MEGA6.0 software based on the full-length of HD-Zip protein sequences [71]. The bootstrap probability of each branch was estimated with 10,000 replications to obtain confidence support.

The gene structure imformation of *TaHDZ* genes were got from the Chinese Spring IWGSC RefSeq v1.0 reference genome, and analysed using the Gene Structure Display Server 2.0 (GSDS; http://gsds.cbi.pku.edu.cn/). The conserved motifs of TaHDZs were identified using SALAD database (http://salad.dna.affrc.go.jp/salad/ en/).

Gene expression analysis by RNA-seq data

To study the expression of *TaHDZ* genes in different tissues, RNA-seq data from ten tissues, including root, stem, leaf of wheat seedling at five-leaf stage, young spike at early booting stage, spike at heading stage, flag leaf at heading stage, and the grain of 5, 10, 15 and 20 DPA were collected from database (http://genedenovoweb.ticp.net: 81/Wheat_GDR1246/index.php?m=index&f=index). For further analysis the expression of *TaHDZ* genes in response to drought stress, we harvested the leaves and roots from three-week-old wheat seedlings subjicted to drought treatment at 0, 1, 3, 6, and 12 hr to conduct the RNA-seq analysis. Wheat plantation and sampling was mentioned above. TopHat and Cufflinks were used to analyze the genes' expression based on the RNA-seq data [72,73]. The FPKM value (fragments per kilobase of transcript per million fragments mapped) was calculated for each *TaHDZ* gene, the log10-transformed (FPKM +1) values of the 113 *TaHDZ* genes were used for heat map generation.

RNA Extraction and Quantitative Real-Time PCR

Total RNA was isolated and purified using Total RNA Rapid Extraction Kit for Polysaccharides Polyphenol Plant (BioTeke) according to the manufacturer's directions. The purified RNA was treated with RNase-free DNase I (TaKaRa, China) to remove any genomic DNA contamination. First-strand cDNA was synthesized from 1 μg of total RNA using Recombinant M-MLV reverse transcriptase (Promega, USA). Quantitative real time-PCR (qRT-PCR) was performed in optical 96-well plates using an ABI7300 Thermo-cycler (Applied Biosystems, USA). Reactions were carried out in 10 μl volume, containing 1 μl diluted cDNA, 200 nM gene-specific primers, and 5 μl SYBR Premix Ex Taq II (TaKaRa) with the following conditions: 10 min at 95 ^{o}C, followed by 40 cycles of 15 s at 95 ^{o}C and 30 s at 60 ^{o}C. The specificity of the amplicon for each primer pair was verified by melting curve analysis. The wheat *Actin* (Gene ID: 542814) and *Arabidopsis Actin2* (AT3G18780) were used as the internal controls for the expression analysis of *TaHDZ5* in wheat and stress-responsive genes in *Arabidopsis*, respectively. The relative gene expression levels were calculated according to the $2^{-\Delta\Delta Ct}$ method [74], with the variation in expression being estimated from three biological replicates. The primer pairs used for qRT-PCR analysis are listed in Additional file 12: Table S4.

TaHDZ5-6A isolation and *Arabidopsis* transformation

The full-length opening reading frame of *TaHDZ5-6A* was amplified from wheat cDNA with gene-specific primers (forward: 5'-ATGGAGCCCGGCCGGCTCAT-3'; reverse: 5' -CTAGTTCCACATCCAGTAGCTGATC-3'), and cloned into the pGreen vector [75] driven by cauliflower mosaic virus (CaMV) 35S promoter. The recombinant vector (*35S::TaHDZ5-6A*) was introduced into *Agrobacterium tumefaciens* and transformed into *Arabidopsis* (*Arabidopsis thaliana*; ecotype Columbia) using the floral dip method. For selection of transformants, T_1 seeds were plated on MS medium containing 2% sucrose and 50 mg/mL kanamycin. Homozygous T_3 plants were used for phenotypic analysis.

Drought tolerance assay

For the drought tolerance assays, seven-day-old *35S::TaHDZ5-6A* plants germinated on MS medium were transferred into pots containing 230 g 2:1 mixture of Jiffy mix and vermiculite. Thirty two-day-old plants growing under favorable water conditions (22 ^{o}C, relative humidity 60%, and 16/8 h light/dark photoperiod) were exposed to drought stress. Water was withheld from the plants for 14 days. Watering was then resumed to allow the plants to recover. Six days later, the number of surviving plants was recorded. At least 64 plants of each line were compared with wild-type (WT) plants in each test, and statistical data were based on data obtained from three independent experiments. The student's *t*-test was used to assess the difference between wild-type and transgenic plants.

Water loss measurement

Water loss rates were measured using eight plants each of wild-type and *35S::TaHDZ5-6A* transgenic plants. Three-week-old soil-grown plants were detached from roots and weighed immediately (fresh weight, FW), the plants were then left on the laboratory bench (22-24 ^{o}C, relative humidity 40-45%) and weighed at designated time intervals (desiccated weights). The proportions of fresh weight loss were calculated relative to the initial weights. Plants were finally oven-dried for 24 h at 80 ^{o}C to a constant dry weight (DW). Water loss was represented as the percentage of initial fresh weight at each time point. Three replicates were performed for each line. The student's *t*-test was used to assess the difference between wild-type and transgenic plants.

Stomatal aperture analysis

Stomatal apertures were measured as described previously [76]. Leaves of similar size and age were sampled from WT and *35S::TaHDZ5-6A* plants that had been subjected to drought for 10 d. Rosette leaves were floated in solutions containing 30 mM KCl, 10 mM Mes-Tris, pH 6.15, and exposed to light for 3h. A light microscope (Olympus ix71, Tokyo, Japan) was used to examine the stomata on epidermal strips obtained from rosette leaves. The width and length of stomatal pores, as determined by the software Image J (http://rsbweb.nih.gov/ij), were used to calculate stomatal apertures (ratio of width to length).

Proline content measurement

Arabidopsis leaves were harvested at designated time-points during drought treatment and used to measure the free proline content. To maximize the sample uniformity at each time-point, leaves of the same size and location were detached from *35S::TaHDZ5-6A* and WT plants. The free proline content was determined as described previously [77]. Samples (~0.1 g) were homogenized in 3% sulfosalicylic acid and boiled for 10min. After the reaction between proline and acid ninhydrin, the absorbance of sample solutions was measured at 520nm with a UV-Vis spectrophotometer (NanoDrop 2000c, Thermo Scientific, Wilmington, DE, USA).

RNA-seq analysis of transgenic *Arabidopsis* plants

For *Arabidopsis* RNA-seq analysis, pooled tissues from ten three-week-old *Arabidopsis* seedlings were collected from *35S::TaHDZ5-6A* and wild-type plants under normal condition. The total RNA was isolated using Total RNA Rapid Extraction Kit for Polysaccharides Polyphenol Plant (BioTeke) according to the manufacturer's directions. The preliminary quantitative of the concentration and purity of the total RNA were implemented using NanoDrop 2000 spectrophotometer (Thermo) and RNase free agarose gel electrophoresis. To remove the residual DNA, the extracted RNA was treated with RNase-free DNase I (New England Biolabs) for 30 min at 37°C. Libraries from the resulting total RNA were prepared using the TruSeq paired-end mRNA-Seq kit and followed by multiplex adapter ligation, and 125 base paired-end sequencing on the Illumina Hiseq-2500 platform. Differential gene expression was determined using Tuxedo RNA-seq analysis pipeline [72,73]. Enrichment analysis of gene ontology of biological pathways (GOBPs) was performed using the DAVID software program [78] to compute *P*-values that indicate the significance of each GOBP being represented by the genes. GOBPs with $P < 0.01$ were identified as enriched processes.

Abbreviations

ABA: Abscisic acid; BLAST: Basic local alignment search tool; CS: Chinese spring; DW: Dry weight; FDR: False Discovery Rate; FPKM: Fragments per kilobase of transcript per million fragments mapped; FW: Fresh weight; GOBP: Gene ontology of biological pathway; HD: Homeobox domain; HD-SAD: HD-START-associated domain; LZ: Leucine Zipper motif; Mw: Molecular weight; NJ: Neighbor-joining; OCL4: OUTER CELL LAYER4; pI: Isoelectric point; PRE: PopREVOLUTA; qRT-PCR: Quantitative reverse transcription polymerase chain reaction; START: Steroidogenic acute regulatory protein-related lipid transfer; TF: transcription factor; WT: Wild-type

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

The relevant data sets supporting the results of this article are included within the article and its additional files.

Competing interests

The authors declare that they have no conflict of interest.

Funding

This work was supported by grants from the National Natural Science Foundation of China (grant no. 31701418), Science Foundation of Shaanxi (grant no. 2018JQ3067), and Talent Fund of Northwest A&F University (grant no. Z111021602). The funders had no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Authors' contributions

HM and ZK conceived and initiated the research; HM designed the experiments; ZW, FL, FM, and XC carried out the experiments. SL, NC, and HM analyzed the data and wrote the manuscript. All authors have read and approved the final manuscript.

Acknowledgements

We thank reviewers for checking our manuscript and the editors for editing the paper. We would like to thank the members of the Bioinformatics Center of Northwest A & F University for their useful input. We also thank the Research Core Facility at the State Key Laboratory of Crop Stress Biology for Arid Areas, NWAFU for support in this work.

Authors' Information

State Key Laboratory of Crop Stress Biology for Arid Areas, College of Plant Protection, Northwest A&F University, Yangling 712100, Shaanxi, China.

References

1. Nakashima K, Ito Y, Yamaguchi-Shinozaki K. Transcriptional regulatory networks in response to abiotic stresses in *Arabidopsis* and grasses. *Plant Physiol*. 2009;149(1):88-95.
2. Ariel FD, Manavella PA, Dezar CA, Chan RL. The true story of the HD-Zip family. *Trends Plant Sci*. 2007;12(9):419-426.
3. Zhao Y, Zhou Y, Jiang H, Li X, Gan D, Peng X, Zhu S, Cheng B. Systematic analysis of sequences and expression patterns of drought-responsive members of the HD-Zip gene family in maize. *PLoS One*.

2011;6(12):e28488.

4. Jain M, Tyagi AK, Khurana JP. Genome-wide identification, classification, evolutionary expansion and expression analyses of homeobox genes in rice. *FEBS J*. 2008;275(11):2845-2861.
5. Li Z, Zhang C, Guo Y, Niu W, Wang Y, Xu Y. Evolution and expression analysis reveal the potential role of the HD-Zip gene family in regulation of embryo abortion in grapes (*Vitis vinifera* L.). *BMC Genomics*. 2017;18:744.
6. Mukherjee K, Bürglin TR. MEKHLA, a novel domain with similarity to PAS domains, is fused to plant homeodomain-leucine zipper III proteins. *Plant Physiol*. 2006;140(4):1142-1150.
7. Valdés AE, Overnäs E, Johansson H, Rada-Iglesias A, Engström P. The homeodomain-leucine zipper (HD-Zip) class I transcription factors ATHB7 and ATHB12 modulate abscisic acid signalling by regulating protein phosphatase 2C and abscisic acid receptor gene activities. *Plant Mol Biol*. 2012;80(4-5):405-418.
8. Capella M, Ribone PA, Arce AL, Chan RL. *Arabidopsis thaliana* HomeoBox 1 (AtHB1), a Homedomain-Leucine Zipper I (HD-Zip I) transcription factor, is regulated by PHYTOCHROME-INTERACTING FACTOR 1 to promote hypocotyl elongation. *New Phytol*. 2015;207(3):669-682.
9. Aoyama T, Dong CH, Wu Y, Carabelli M, Sessa G, Ruberti I, Morelli G, Chua NH. Ectopic expression of the *Arabidopsis* transcriptional activator *Athb-1* alters leaf cell fate in tobacco. *Plant Cell*. 1995;7(11):1773-1785.
10. Kovalchuk N, Chew W, Sornaraj P, Borisjuk N, Yang N, Singh R, Bazanova N, Shavrukov Y, Guendel A, Munz E, Borisjuk L, Langridge P, Hrmova M, Lopato S. The homeodomain transcription factor TaHDZipI-2 from wheat regulates frost tolerance, flowering time and spike development in transgenic barley. *New Phytol*. 2016;211(2):671-687.
11. Harris JC, Sornaraj P, Taylor M, Bazanova N, Baumann U, Lovell B, Langridge P, Lopato S, Hrmova M. Molecular interactions of the *γ*-clade homeodomain- leucine zipper class I transcription factors during the wheat response to water deficit. *Plant Mol Biol*. 2016;90(4-5):435-452.
12. Agalou A, Purwantomo S, Overnäs E, Johannesson H, Zhu X, Estiati A, de Kam RJ, Engström P, Slamet-Loedin IH, Zhu Z, Wang M, Xiong L, Meijer AH, Ouwerkerk PB. A genome-wide survey of *HD-Zip* genes in rice and analysis of drought-responsive family members. *Plant Mol Biol*. 2008;66(1-2):87-103.
13. Carabelli M, Morelli G, Whitelam G, Ruberti I. Twilight-zone and canopy shade induction of the *Athb-2* homeobox gene in green plants. *Proc Natl Acad Sci U S A*. 1996;93(8):3530-3535.
14. Steindler C, Matteucci A, Sessa G, Weimar T, Ohgishi M, Aoyama T, Morelli G, Ruberti I. Shade avoidance responses are mediated by the ATHB-2 HD-zip protein, a negative regulator of gene expression. *Development*. 1999;126(19): 4235-4245.
15. Dezar CA, Giacomelli JI, Manavella PA, Ré DA, Alves-Ferreira M, Baldwin IT, Bonaventure G, Chan RL. HAHB10, a sunflower HD-Zip II transcription factor, participates in the induction of flowering and in the control of phytohormone-mediated responses to biotic stress. *J Exp Bot*. 2011;62(3):1061- 1076.

16. Robischon M, Du J, Miura E, Groover A. The *Populus* class III HD-ZIP, *popREVOLUTA*, influences cambium initiation and patterning of woody stems. *Plant Physiol*. 2011;155(3):1214-1225.
17. Baima S, Possenti M, Matteucci A, Wisman E, Altamura MM, Ruberti I, Morelli G. The *Arabidopsis* ATHB-8 HD-zip protein acts as a differentiation-promoting transcription factor of the vascular meristems. *Plant Physiol*. 2001;126(2):643- 655.
18. Kim J, Jung JH, Reyes JL, Kim YS, Kim SY, Chung KS, Kim JA, Lee M, Lee Y, Narry Kim V, Chua NH, Park CM. MicroRNA-directed cleavage of *ATHB15* mRNA regulates vascular development in *Arabidopsis* inflorescence stems. *Plant J*. 2005;42(1):84-94.
19. Landau U, Asis L, Eshed Williams L. The ERECTA, CLAVATA and class III HD-ZIP pathways display synergistic interactions in regulating floral meristem activities. *PLoS One*. 2015;10(5):e0125408.
20. Hawker NP, Bowman JL. Roles for Class III *HD-Zip* and *KANADI* genes in *Arabidopsis* root development. *Plant Physiol*. 2004;135(4):2261-2270.
21. Kamata N, Okada H, Komeda Y, Takahashi T. Mutations in epidermis-specific HD-ZIP IV genes affect floral organ identity in *Arabidopsis thaliana*. *Plant J*. 2013;75(3):430-440.
22. Di Cristina M, Sessa G, Dolan L, Linstead P, Baima S, Ruberti I, Morelli G. The *Arabidopsis* Athb-10 (GLABRA2) is an HD-Zip protein required for regulation of root hair development. *Plant J*. 1996;10(3):393-402.
23. Vernoud V, Laigle G, Rozier F, Meeley RB, Perez P, Rogowsky PM. The HD-ZIP IV transcription factor OCL4 is necessary for trichome patterning and anther development in maize. *Plant J*. 2009;59:883-894.
24. Yu LH, Wu SJ, Peng YS, Liu RN, Chen X, Zhao P, Xu P, Zhu JB, Jiao GL, Pei Y, Xiang CB. *Arabidopsis EDT1/HDG11* improves drought and salt tolerance in cotton and poplar and increases cotton yield in the field. *Plant Biotechnol J*. 2016;14(1):72-84.
25. Jia J, Zhao S, Kong X, Li Y, Zhao G, He W, Appels R, Pfeifer M, Tao Y, Zhang X, Jing R, Zhang C, Ma Y, Gao L, Gao C, Spannagl M, Mayer KF, Li D, Pan S, Zheng F, Hu Q, Xia X, Li J, Liang Q, Chen J, Wicker T, Gou C, Kuang H, He G, Luo Y, Keller B, Xia Q, Lu P, Wang J, Zou H, Zhang R, Xu J, Gao J, Middleton C, Quan Z, Liu G, Wang J; International Wheat Genome Sequencing Consortium, Yang H, Liu X, He Z, Mao L, Wang J. *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. *Nature*. 2013;496(7443):91-95.
26. Ling HQ, Zhao S, Liu D, Wang J, Sun H, Zhang C, Fan H, Li D, Dong L, Tao Y, Gao C, Wu H, Li Y, Cui Y, Guo X, Zheng S, Wang B, Yu K, Liang Q, Yang W, Lou X, Chen J, Feng M, Jian J, Zhang X, Luo G, Jiang Y, Liu J, Wang Z, Sha Y, Zhang B, Wu H, Tang D, Shen Q, Xue P, Zou S, Wang X, Liu X, Wang F, Yang Y, An X, Dong Z, Zhang K, Zhang X, Luo MC, Dvorak J, Tong Y, Wang J, Yang H, Li Z, Wang D, Zhang A, Wang J. Draft genome of the wheat A-genome progenitor *Triticum urartu*. *Nature*. 2013;496(7443):87-90.
27. Feldman M, Levy AA. Allopolyploidy-a shaping force in the evolution of wheat genomes. *Cytogenet Genome Res*. 2005;109(1-3):250-258.

28. Marcussen T, Sandve SR, Heier L, Spannagl M, Pfeifer M; International Wheat Genome Sequencing Consortium, Jakobsen KS, Wulff BB, Steuernagel B, Mayer KF, Olsen OA. Ancient hybridizations among the ancestral genomes of bread wheat. *Science*. 2014;345(6194):1250092.
29. Brenchley R, Spannagl M, Pfeifer M, Barker GL, D'Amore R, Allen AM, McKenzie N, Kramer M, Kerhornou A, Bolser D, Kay S, Waite D, Trick M, Bancroft I, Gu Y, Huo N, Luo MC, Sehgal S, Gill B, Kianian S, Anderson O, Kersey P, Dvorak J, McCombie WR, Hall A, Mayer KF, Edwards KJ, Bevan MW, Hall N. Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature*. 2012;491(7426):705-710.
30. Choulet F, Alberti A, Theil S, Glover N, Barbe V, Daron J, Pingault L, Sourdille P, Couloux A, Paux E, Leroy P, Mangenot S, Guilhot N, Le Gouis J, Balfourier F, Alaux M, Jamilloux V, Poulain J, Durand C, Bellec A, Gaspin C, Safar J, Dolezel J, Rogers J, Vandepoele K, Aury JM, Mayer K, Berges H, Quesneville H, Wincker P, Feuillet C. Structural and functional partitioning of bread wheat chromosome 3B. *Science*. 2014;345(6194):1249721.
31. Luo MC, Gu YQ, Puiu D, Wang H, Twardziok SO, Deal KR, Huo N, Zhu T, Wang L, Wang Y, McGuire PE, Liu S, Long H, Ramasamy RK, Rodriguez JC, Van SL, Yuan L, Wang Z, Xia Z, Xiao L, Anderson OD, Ouyang S, Liang Y, Zimin AV, Pertea G, Qi P, Bennetzen JL, Dai X, Dawson MW, Müller HG, Kugler K, Rivarola-Duarte L, Spannagl M, Mayer KFX, Lu FH, Bevan MW, Leroy P, Li P, You FM, Sun Q, Liu Z, Lyons E, Wicker T, Salzberg SL, Devos KM, Dvořák J. Genome sequence of the progenitor of the wheat D genome *Aegilops tauschii*. *Nature*. 2017;551(7681):498-502.
32. Zhao G, Zou C, Li K, Wang K, Li T, Gao L, Zhang X, Wang H, Yang Z, Liu X, Jiang W, Mao L, Kong X, Jiao Y, Jia J. The *Aegilops tauschii* genome reveals multiple impacts of transposons. *Nat Plants*. 2017;3(12):946-955.
33. Ling HQ, Ma B, Shi X, Liu H, Dong L, Sun H, Cao Y, Gao Q, Zheng S, Li Y, Yu Y, Du H, Qi M, Li Y, Lu H, Yu H, Cui Y, Wang N, Chen C, Wu H, Zhao Y, Zhang J, Li Y, Zhou W, Zhang B, Hu W, van Eijk MJT, Tang J, Witsenboer HMA, Zhao S, Li Z, Zhang A, Wang D, Liang C. Genome sequence of the progenitor of wheat A subgenome *Triticum urartu*. *Nature*. 2018;557(7705): 424-428.
34. Yang Y, Luang S, Harris J, Riboni M, Li Y, Bazanova N, Hrmova M, Haefele S, Kovalchuk N, Lopato S. Overexpression of the class I homeodomain transcription factor *TaHDZipI-5* increases drought and frost tolerance in transgenic wheat. *Plant Biotechnol J*. 2018;16(6):1227-1240.
35. Yue H, Shu D, Wang M, Xing G, Zhan H, Du X, Song W, Nie X. Genome-wide identification and expression analysis of the HD-Zip gene family in wheat (*Triticum aestivum* L.). *Genes (Basel)*. 2018;9(2).
36. Henriksson E, Olsson AS, Johannesson H, Johansson H, Hanson J, Engstrom P, Soderman E. Homeodomain leucine zipper class I genes in *Arabidopsis*. Expression patterns and phylogenetic relationships. *Plant Physiol*. 2005; 139(1): 509-518.
37. Hu R, Chi X, Chai G, Kong Y, He G, Wang X, Shi D, Zhang D, Zhou G. Genome-wide identification, evolutionary expansion, and expression profile of homeodomain-leucine zipper gene family in poplar (*Populus trichocarpa*). *PLoS One*. 2012;7(2):e31149.

38. Xu G, Guo C, Shan H, Kong H. Divergence of duplicate genes in exon-intron structure. *Proc Natl Acad Sci U S A*. 2012;109(4):1187-1192.
39. Sakuma Y, Maruyama K, Osakabe Y, Qin F, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. Functional analysis of an *Arabidopsis* transcription factor, DREB2A, involved in drought-responsive gene expression. *Plant Cell*. 2006;18(5):1292-1309.
40. Nakashima K, Fujita Y, Katsura K, Maruyama K, Narusaka Y, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. Transcriptional regulation of ABI3- and ABA-responsive genes including *RD29B* and *RD29A* in seeds, germinating embryos, and seedlings of *Arabidopsis*. *Plant Mol Biol*. 2006;60(1):51-68.
41. Fujita M, Fujita Y, Maruyama K, Seki M, Hiratsu K, Ohme-Takagi M, Tran LS, Yamaguchi-Shinozaki K, Shinozaki K. A dehydration-induced NAC protein, RD26, is involved in a novel ABA-dependent stress-signaling pathway. *Plant J*. 2004;39(6):863-876.
42. Nylander M, Svensson J, Palva ET, Welin BV. Stress-induced accumulation and tissue-specific localization of dehydrins in *Arabidopsis thaliana*. *Plant Mol Biol*. 2001;45(3):263-279.
43. Yoshida T, Nishimura N, Kitahata N, Kuromori T, Ito T, Asami T, Shinozaki K, Hirayama T. *ABA-hypersensitive germination 3* encodes a protein phosphatase 2C (AtPP2CA) that strongly regulates abscisic acid signaling during germination among *Arabidopsis* protein phosphatase 2Cs. *Plant Physiol*. 2006;140(1):115- 126.
44. Tran LS, Nakashima K, Sakuma Y, Simpson SD, Fujita Y, Maruyama K, Fujita M, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. Isolation and functional analysis of *Arabidopsis* stress-inducible NAC transcription factors that bind to a drought-responsive *cis*-element in the *early responsive to dehydration stress 1* promoter. *Plant Cell*. 2004;16(9):2481-2498.
45. Iuchi S, Kobayashi M, Taji T, Naramoto M, Seki M, Kato T, Tabata S, Kakubari Y, Yamaguchi-Shinozaki K, Shinozaki K. Regulation of drought tolerance by gene manipulation of 9-*cis*-epoxycarotenoid dioxygenase: a key enzyme in abscisic acid biosynthesis in *Arabidopsis*. *Plant J*. 2001;27:325-333.
46. Aubert Y, Vile D, Pervent M, Aldon D, Ranty B, Simonneau T, Vavasseur A, Galaud JP. RD20, a stress-inducible caleosin, participates in stomatal control, transpiration and drought tolerance in *Arabidopsis thaliana*. *Plant Cell Physiol*. 2010;51(12):1975-1987.
47. He Y, Ahmad D, Zhang X, Zhang Y, Wu L, Jiang P, Ma H. Genome-wide analysis of family-1 UDP glycosyltransferases (UGT) and identification of *UGT* genes for FHB resistance in wheat (*Triticum aestivum* L.). *BMC Plant Biol*. 2018;18(1):67.
48. Harris JC, Hrmova M, Lopato S, Langridge P. Modulation of plant growth by HD-Zip class I and II transcription factors in response to environmental stimuli. *New Phytol*. 2011;190(4):823-837.
49. Cabello JV, Arce AL, Chan RL. The homologous HD-Zip I transcription factors HaHB1 and AtHB13 confer cold tolerance via the induction of pathogenesis- related and glucanase proteins. *Plant J*. 2012;69(1):141-153.
50. Cabello JV, Chan RL. The homologous homeodomain-leucine zipper transcription factors HaHB1 and AtHB13 confer tolerance to drought and salinity stresses via the induction of proteins that stabilize

membranes. *Plant Biotechnol J*. 2012; 10(7):815-825.

51. Johannesson H, Wang Y, Hanson J, Engström P. The *Arabidopsis thaliana* homeobox gene *ATHB5* is a potential regulator of abscisic acid responsiveness in developing seedlings. *Plant Mol Biol*. 2003;51(5):719-729.
52. Lechner E, Leonhardt N, Eisler H, Parmentier Y, Alioua M, Jacquet H, Leung J, Genschik P. MATH/BTB CRL3 receptors target the homeodomain-leucine zipper ATHB6 to modulate abscisic acid signaling. *Dev Cell*. 2011;21(6):1116- 1128.
53. González-Grandío E, Pajoro A, Franco-Zorrilla JM, Tarancón C, Immink RG, Cubas P. Abscisic acid signaling is controlled by a BRANCHED1/HD-ZIP I cascade in *Arabidopsis* axillary buds. *Proc Natl Acad Sci U S A*. 2017;114(2): E245-E254.
54. Prigge MJ, Otsuga D, Alonso JM, Ecker JR, Drews GN, Clark SE. Class III homeodomain-leucine zipper gene family members have overlapping, antagonistic, and distinct roles in *Arabidopsis* development. *Plant Cell*. 2005; 17(1):61-76.
55. Lee C, Clark SE. A WUSCHEL-independent stem cell specification pathway is repressed by PHB, PHV and CNA in *Arabidopsis*. *PLoS One*. 2015;10(5):e0126006.
56. Yamada T, Sasaki Y, Hashimoto K, Nakajima K, Gasser CS. CORONA, PHABULOSA and PHAVOLUTA collaborate with BELL1 to confine *WUSCHEL* expression to the nucellus in *Arabidopsis* ovules. *Development*. 2016;143(3):422-426.
57. Kubo H, Peeters AJ, Aarts MG, Pereira A, Koornneef M. *ANTHOCYANINLESS2*, a homeobox gene affecting anthocyanin distribution and root development in *Arabidopsis*. *Plant Cell*. 1999;11(7):1217-1226.
58. Nakamura M, Katsumata H, Abe M, Yabe N, Komeda Y, Yamamoto KT, Takahashi T. Characterization of the class IV homeodomain-Leucine Zipper gene family in *Arabidopsis*. *Plant Physiol*. 2006;141(4):1363-1375.
59. Gao S, Fang J, Xu F, Wang W, Chu C. Rice HOX12 regulates panicle exsertion by directly modulating the expression of *ELONGATED UPPERMOST INTERNODE1*. *Plant Cell*. 2016;28(3):680-695.
60. Hur YS, Um JH, Kim S, Kim K, Park HJ, Lim JS, Kim WY, Jun SE, Yoon EK, Lim J. *Arabidopsis thaliana* homeobox 12 (ATHB12), a homeodomain-leucine zipper protein, regulates leaf growth by promoting cell expansion and endoreduplication. *New Phytol*. 2015;205(1):316-328.
61. Reymond MC, Brunoud G, Chauvet A, Martinez-Garcia JF, Martin-MagnietteML, Moneger F, Scutt CP. A light-regulated genetic module was recruited to carpel development in *Arabidopsis* following a structural change to SPATULA. *Plant cell*. 2012;24(7):2812-2825.
62. Bou-Torrent J, Salla-Martret M, Brandt R, Musielak T, Palauqui JC, Martinez-Garcia JF, Wenkel S. ATHB4 and HAT3, two class II HD-ZIP transcription factors, control leaf development in *Arabidopsis*. *Plant Signal Behav*. 2012;7(11):1382-1387.
63. Izhaki A, Bowman JL. KANADI and class III HD-Zip gene families regulate embryo patterning and modulate auxin flow during embryogenesis in *Arabidopsis*. *Plant Cell*. 2007;19(2):495-508.

64. Vernoud V, Laigle G, Rozier F, Meeley RB, Perez P, Rogowsky PM. The HD-ZIP IV transcription factor OCL4 is necessary for trichome patterning and anther development in maize. *Plant J*. 2009;59(6):883-894.
65. Ogawa E, Yamada Y, Sezaki N, Kosaka S, Kondo H, Kamata N, Abe M, Komeda Y, Takahashi T. ATML1 and PDF2 play a redundant and essential role in *Arabidopsis* embryo development. *Plant Cell Physiol*. 2015;56(6):1183-1192.
66. Kim S, Kang JY, Cho DI, Park JH, Kim SY. ABF2, an ABRE-binding bZIP factor, is an essential component of glucose signaling and its overexpression affects multiple stress tolerance. *Plant J*. 2004;40(1):75-87.
67. Maruyama K, Sakuma Y, Kasuga M, Ito Y, Seki M, Goda H, Shimada Y, Yoshida S, Shinozaki K, Yamaguchi-Shinozaki K. Identification of cold-inducible downstream genes of the *Arabidopsis* DREB1A/CBF3 transcriptional factor using two microarray systems. *Plant J*. 2004;38(6):982-993.
68. Dai X, Xu Y, Ma Q, Xu W, Wang T, Xue Y, Chong K. Overexpression of an R1R2R3 MYB gene, *OsMYB3R-2*, increases tolerance to freezing, drought, and salt stress in transgenic *Arabidopsis*. *Plant Physiol*. 2007;143(4):1739-1751.
69. Eddy SR. Profile hidden Markov models. *Bioinformatics*. 1998;14(9):755-763.
70. Thompson JD, Gibson TJ, Higgins DG. Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics*. 2002;Chapter 2:Unit 2.3.
71. 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.
72. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc*. 2012;7(3):562- 578.
73. Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat Biotechnol*. 2013;31(1):46-53.
74. Livak KJ, Schmittgen TD. 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.
75. Qin F, Sakuma Y, Tran LS, Maruyama K, Kidokoro S, Fujita Y, Fujita M, Umezawa T, Sawano Y, Miyazono K, Tanokura M, Shinozaki K, Yamaguchi-Shinozaki K. *Arabidopsis* DREB2A-interacting proteins function as RING E3 ligases and negatively regulate plant drought stress-responsive gene expression. *Plant Cell*. 2008;20(6):1693-1707.
76. Pei ZM, Kuchitsu K, Ward JM, Schwarz M, Schroeder JI. Differential abscisic acid regulation of guard cell slow anion channels in *Arabidopsis* wild-type and *abi1* and *abi2* mutants. *Plant Cell*. 1997;9(3):409-423.
77. Bates LS, Waldren RP, Teare ID. Rapid determination of free proline for water-stress studies. *Plant Soil*. 1973;39:205-207.
78. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4(1): 44-57.

Tables

Table 1 Detail information of wheat *HD-Zip* genes.

Name	Protein id	chr	Start	End	Number of amino acids	Molecular weight	Theoretical pI	group	Previous nomenclature[35]
TaHDZ1-4A	TraesCS4A02G405800	chr4A	678872316	678873272	285	31206.22	4.79	I	TaHDZ16-A/Traes_4AL_99A941299
TaHDZ1-4B	TraesCS4B02G305300	chr4B	593422292	593423732	316	35099.79	4.99	I	TaHDZ16-B/Traes_4BL_ECD20BE67
TaHDZ1-4D	TraesCS4D02G303500	chr4D	471961107	471962520	316	34919.58	4.98	I	
TaHDZ2-5A	TraesCS5A02G249000	chr5A	463452047	463453063	270	28917.9	4.77	I	
TaHDZ2-5B	TraesCS5B02G246700	chr5B	428461602	428462604	269	28884.88	4.7	I	TaHDZ22-B/Traes_5BL_5DE02D63E
TaHDZ2-5D	TraesCS5D02G256200	chr5D	362547609	362548803	270	28909.87	4.7	I	
TaHDZ3-4A	TraesCS4A02G016600	chr4A	11296831	11299294	338	35973.87	6.08	I	TaHDZ15-A/Traes_4AS_F04DD4409
TaHDZ3-4B	TraesCS4B02G287600	chr4B	571047644	571049696	331	36494.59	6.17	I	TaHDZ15-B/Traes_4BL_BE3E058A6
TaHDZ3-4D	TraesCS4D02G286400	chr4D	457074633	457076703	330	36411.5	6.17	I	TaHDZ15-D/Traes_4DL_88ABAD6C0
TaHDZ4-5A	TraesCS5A02G199300	chr5A	404177302	404178633	300	32713.28	4.86	I	
TaHDZ4-5B	TraesCS5B02G197700	chr5B	356597007	356598447	299	32566.01	4.86	I	TaHDZ21-B/Traes_5BL_028D02DF6
TaHDZ4-5D	TraesCS5D02G205000	chr5D	310677920	310679395	299	32637.21	4.93	I	
TaHDZ5-6A	TraesCS6B02G321100	chr6B	568257664	568259128	343	37502.12	4.59	I	TaHDZ24-A/Traes_6AL_36AB0312C
TaHDZ5-6D	TraesCS6D02G272000	chr6D	380799345	380800814	374	41012.34	4.82	I	TaHDZ24-D/Traes_6DL_FF4C8C4AB
TaHDZ6-5A	TraesCS5A02G316800	chr5A	527904954	527905823	249	27500.54	4.97	I	
TaHDZ6-5B	TraesCS5B02G317400	chr5B	501464138	501465028	249	27502.62	5.02	I	TaHDZ20-B/Traes_5BL_9C32B27E2
TaHDZ6-5D	TraesCS5D02G323100	chr5D	415272394	415273274	247	27246.32	5.09	I	TaHDZ20-D/Traes_5DL_96F9EED93
TaHDZ7-2A	TraesCS2A02G389400	chr2A	637984504	637985699	265	29697.25	5.12	I	TaHDZ8-A/Traes_2AL_BFB0C6D4C
TaHDZ7-2B	TraesCS2B02G407600	chr2B	578605119	578605996	260	29074.51	5.2	I	TaHDZ8-B/Traes_2BL_B69300543
TaHDZ7-2D	TraesCS2D02G387300	chr2D	492676386	492677218	243	26899.14	5.04	I	
TaHDZ8-6A	TraesCS6A02G240400	chr6A	451659958	451660731	221	24977.72	5.27	I	
TaHDZ8-6B	TraesCS6B02G284300	chr6B	512447547	512448327	226	25584.39	5.09	I	
TaHDZ8-6D	TraesCS6D02G222600	chr6D	314056149	314056946	225	25536.5	5.65	I	TaHDZ26-D/Traes_6DS_D281B7D32
TaHDZ9-4A	TraesCS4A02G040600	chr4A	34026753	34027976	231	25530.54	6.24	I	TaHDZ17-A/Traes_4AS_1EA23DE08
TaHDZ9-4B	TraesCS4B02G261600	chr4B	530519687	530520929	233	25576.61	6.04	I	TaHDZ17-B/Traes_4BL_BE10705D5
TaHDZ9-4D	TraesCS4D02G261600	chr4D	432748587	432749803	234	25708.82	6.24	I	TaHDZ17-D/Traes_4DL_4798D0BBD
TaHDZ10-2B	TraesCS2B02G405700	chr2B	573974813	573975706	221	24267	7.7	I	TaHDZ7-B/Traes_2BL_419CEED79
TaHDZ10-2D	TraesCS2D02G385500	chr2D	490117422	490118593	247	27340.52	7.17	I	
TaHDZ11-2A	TraesCS2A02G188500	chr2A	153452292	153453239	238	25785.75	7.12	I	
TaHDZ11-2B	TraesCS2B02G218800	chr2B	208210912	208211847	238	25858.89	7.12	I	TaHDZ6-B/Traes_2BS_BD0ED621D
TaHDZ11-2D	TraesCS2D02G199200	chr2D	148855464	148856413	238	25755.72	6.76	I	TaHDZ6-D/Traes_2DS_20F748657
TaHDZ12-6A	TraesCS6A02G120300	chr6A	91924600	91925404	203	22045.89	9.64	II	
TaHDZ12-6B	TraesCS6B02G148700	chr6B	149551263	149551965	209	22641.46	9.41	II	
TaHDZ12-6D	TraesCS6D02G110400	chr6D	76042794	76043516	192	20888.56	9.79	II	TaHDZ27-D/Traes_6DS_F00EB2E01
TaHDZ13-6A	TraesCS6A02G120600	chr6A	92054391	92055152	226	24538.57	8.72	II	TaHDZ25-A/Traes_6AS_3E534A2C1
TaHDZ13-6B	TraesCS6B02G149000	chr6B	149773307	149774051	220	24273.43	9.32	II	
TaHDZ13-6D	TraesCS6D02G110600	chr6D	76064358	76065121	225	24759.74	8.38	II	TaHDZ25-D/Traes_6DS_17B737547
TaHDZ14-7A	TraesCS7A02G423800	chr7A	616552387	616553258	219	24421.47	9.25	II	TaHDZ28-A/Traes_7AL_44206BE21
TaHDZ14-7B	TraesCS7B02G326300	chr7B	579825401	579826290	223	24708.84	9.47	II	
TaHDZ14-7D	TraesCS7D02G417700	chr7D	536514689	536515582	222	24526.65	9.61	II	TaHDZ28-D/Traes_7DL_2FE5181AF
TaHDZ15-1A	TraesCS1A02G372900	chr1A	549301452	549303005	305	33195.14	6.27	II	
TaHDZ15-1B	TraesCS1B02G393100	chr1B	626097677	626098967	306	33662.73	6.68	II	TaHDZ3-B/Traes_1BL_BCA60D8B6
TaHDZ15-1D	TraesCS1D02G379700	chr1D	455759848	455761548	304	33117.06	6.67	II	
TaHDZ16-4A	TraesCS4A02G059600	chr4A	56143977	56145391	275	29373.91	6.67	II	
TaHDZ16-4B	TraesCS4B02G235300	chr4B	491464360	491465839	277	29618.16	6.67	II	TaHDZ18-B/Traes_4BL_78DD63002
TaHDZ16-4D	TraesCS4D02G236600	chr4D	398669293	398670678	274	29332.85	6.67	II	

TaHDZ17-3B	TraesCS3B02G000100	chr3B	213438	214466	228	25282.62	8.64	II	TaHDZ10-B/TRAES3BF043500070CFD
TaHDZ17-3D	TraesCS3D02G009700	chr3D	3294033	3295097	226	24874.23	9.24	II	TaHDZ10-D/Traes_3DS_7CCB5ECD2
TaHDZ18-5A	TraesCS5A02G232700	chr5A	448090083	448091497	351	36587.01	7.01	II	
TaHDZ18-5B	TraesCS5B02G231300	chr5B	407816831	407818230	355	37049.33	6.27	II	TaHDZ19-B/Traes_5BL_4A3874701
TaHDZ18-5D	TraesCS5D02G235300	chr5D	344398948	344400554	339	35808.01	6.49	II	
TaHDZ19-3A	TraesCS3A02G231600	chr3A	432374494	432377827	222	24782.22	9.17	II	
TaHDZ19-3B	TraesCS3B02G260800	chr3B	418718384	418721923	222	24704.13	9.42	II	TaHDZ11-B/TRAES3BF026400090CFD
TaHDZ19-3D	TraesCS3D02G221800	chr3D	302265801	302269312	222	24776.24	9.27	II	
TaHDZ20-1A	TraesCS1A02G219200	chr1A	387840646	387841766	329	35072.38	8.71	II	TaHDZ4-A/Traes_1AL_1444D461A
TaHDZ20-1B	TraesCS1B02G232700	chr1B	418105692	418106801	327	34775.13	8.88	II	
TaHDZ20-1D	TraesCS1D02G220900	chr1D	308459300	308460410	326	34691.06	8.89	II	
TaHDZ21-2A	TraesCS2A02G415900	chr2A	672415211	672416239	227	25675.08	8.84	II	TaHDZ9-A/Traes_2AL_EF9549D16
TaHDZ21-2B	TraesCS2B02G434900	chr2B	624891321	624892387	230	26196.48	8.84	II	TaHDZ9-B/Traes_2BL_02479C76A
TaHDZ21-2D	TraesCS2D02G412900	chr2D	527546793	527547829	230	25914.27	8.84	II	TaHDZ9-D/Traes_2DL_67F1183B2
TaHDZ22-4A	TraesCS4A02G382400	chr4A	660653849	660655014	344	36296.63	9.53	II	
TaHDZ23-7A	TraesCS7A02G083800	chr7A	48525735	48526646	266	28231.91	9.16	II	TaHDZ14-A/Traes_4AL_822582A19
TaHDZ23-7D	TraesCS7D02G079000	chr7D	46711457	46712351	269	28558.14	9.26	II	
TaHDZ24-3A	TraesCS3A02G312800	chr3A	552611093	552615178	874	94707.86	6.06	III	
TaHDZ24-3B	TraesCS3B02G159100	chr3B	154220609	154224807	841	91388.76	5.95	III	
TaHDZ24-3D	TraesCS3D02G141500	chr3D	103645450	103649526	845	91803.17	5.92	III	
TaHDZ25-1A	TraesCS1A02G157500	chr1A	279733261	279742276	840	92041.09	5.65	III	TaHDZ1-A/Traes_1AL_0BE456AC0
TaHDZ25-1B	TraesCS1B02G173900	chr1B	311419246	311427777	840	92056.14	5.65	III	TaHDZ1-B/Traes_1BL_43408C9B0
TaHDZ25-1D	TraesCS1D02G155200	chr1D	217636182	217644644	606	66231.68	6.31	III	
TaHDZ26-4B	TraesCS4B02G385200	chr4B	664152394	664159763	839	91509.33	5.55	III	
TaHDZ26-4D	TraesCS4D02G359600	chr4D	506968051	506975377	838	91591.45	5.61	III	
TaHDZ27-5A	TraesCS5A02G375800	chr5A	573645493	573651017	862	93733.6	6.09	III	
TaHDZ27-5B	TraesCS5B02G378000	chr5B	556177511	556182640	862	93776.77	6.09	III	
TaHDZ27-5D	TraesCS5D02G385300	chr5D	454414315	454419575	862	93676.55	6.09	III	
TaHDZ28-5A	TraesCS5A02G043400	chr5A	39845363	39851741	846	91939.04	6.1	III	
TaHDZ28-5B	TraesCS5B02G047200	chr5B	53246630	53252235	879	95339.96	6.61	III	TaHDZ23-B/Traes_5BS_360DD5644
TaHDZ28-5D	TraesCS5D02G052300	chr5D	50483246	50488859	883	95733.49	6.73	III	TaHDZ23-D/Traes_5DS_50846FD0C
TaHDZ29-3A	TraesCS3A02G325800	chr3A	571174922	571179414	683	74860.15	7.92	IV	
TaHDZ29-3B	TraesCS3B02G354900	chr3B	565265645	565270023	683	74621.82	8.17	IV	TaHDZ13-B/TRAES3BF075200070CFD
TaHDZ29-3D	TraesCS3D02G319200	chr3D	433133441	433137721	683	74697.92	7.58	IV	TaHDZ13-D/Traes_3DL_8AAFB7B06
TaHDZ30-4A	TraesCS4A02G231300	chr4A	540600328	540604473	770	83653.94	6.99	IV	
TaHDZ30-4B	TraesCS4B02G084700	chr4B	83052927	83058776	788	85693.54	6.36	IV	
TaHDZ30-4D	TraesCS4D02G082600	chr4D	56025816	56030159	805	87663.69	6.79	IV	
TaHDZ31-5A	TraesCS5A02G330200	chr5A	539504162	539509761	744	80684.05	6.1	IV	
TaHDZ31-5B	TraesCS5B02G330300	chr5B	514678511	514683625	751	81666.17	6.02	IV	
TaHDZ31-5D	TraesCS5D02G335900	chr5D	425465587	425471148	753	82053.73	5.8	IV	
TaHDZ32-3A	TraesCS3A02G305300	chr3A	541243574	541247038	761	82154.35	7.49	IV	
TaHDZ32-3B	TraesCS3B02G331200	chr3B	536388767	536392301	755	81462.58	7.8	IV	TaHDZ12-B/TRAES3BF023000040CFD
TaHDZ32-3D	TraesCS3D02G296500	chr3D	410412719	410416215	778	84493.39	9.14	IV	
TaHDZ33-6A	TraesCS6A02G324500	chr6A	558785948	558788800	685	75094.44	6.33	IV	
TaHDZ33-6B	TraesCS6B02G354900	chr6B	623081402	623084291	697	75242.66	6.48	IV	
TaHDZ33-6D	TraesCS6D02G304300	chr6D	413269725	413272565	685	75009.28	6.25	IV	
TaHDZ34-7A	TraesCS7A02G167900	chr7A	124067724	124070795	725	79209	6.13	IV	
TaHDZ34-7B	TraesCS7B02G072700	chr7B	80658897	80662738	730	79513.39	6.19	IV	

TaHDZ34-7D	TraesCS7D02G168700	chr7D	119565615	119569016	732	79795.75	6.22	IV	
TaHDZ35-1A	TraesCS1A02G193400	chr1A	350395750	350401039	873	94078.96	5.68	IV	
TaHDZ35-1B	TraesCS1B02G208400	chr1B	377965083	377969806	890	96021.18	5.57	IV	
TaHDZ35-1D	TraesCS1D02G197300	chr1D	278105002	278110971	883	95091.14	5.64	IV	TaHDZ2-D/Traes_1DL_9FB53E48A
TaHDZ36-6A	TraesCS6A02G255800	chr6A	474327845	474334088	804	86379.32	5.51	IV	
TaHDZ36-6B	TraesCS6B02G269700	chr6B	485537128	485542736	804	86399.39	5.51	IV	
TaHDZ36-6D	TraesCS6D02G237000	chr6D	335062906	335068647	804	86465.43	5.46	IV	
TaHDZ37-2A	TraesCS2A02G401200	chr2A	654907253	654913197	798	85781.8	5.54	IV	
TaHDZ37-2B	TraesCS2B02G419200	chr2B	600740018	600745678	785	84370.1	5.92	IV	
TaHDZ37-2D	TraesCS2D02G398600	chr2D	511688028	511693305	784	84013.06	5.87	IV	TaHDZ5-D/Traes_2DL_036F2A3FC
TaHDZ38-5A	TraesCS5A02G314400	chr5A	524906817	524907230	849	90375.05	5.7	IV	
TaHDZ38-5B	TraesCS5B02G315100	chr5B	497213945	497219576	849	90378.01	5.7	IV	
TaHDZ38-5D	TraesCS5D02G320600	chr5D	412738548	412744058	849	90401.09	5.7	IV	
TaHDZ39-7A	TraesCS7A02G308400	chr7A	436693693	436698045	796	85517.17	5.99	IV	
TaHDZ39-7B	TraesCS7B02G208600	chr7B	381824085	381829068	798	85647.27	5.99	IV	
TaHDZ39-7D	TraesCS7D02G305200	chr7D	386725334	386730432	796	85592.19	5.99	IV	
TaHDZ40-2A	TraesCS2A02G474000	chr2A	715337511	715343047	777	83524.33	5.59	IV	
TaHDZ40-2B	TraesCS2B02G497500	chr2B	694054407	694060937	775	83454.3	5.65	IV	
TaHDZ40-2D	TraesCS2D02G473700	chr2D	577152396	577158953	776	83438.3	5.59	IV	

Additional Files

Additional file 1: Figure S1 Distribution of *TaHDZ* genes among 21 chromosomes of wheat genome.

Additional file 2: Figure S2 Schematic representation of the conserved motifs in the TaHDZ proteins. Each motif is represented by a colored box. The black lines represent the non-conserved sequences.

Additional file 3: Table S1 FPKM (Fragments Per Kilobase Million) values of the *TaHDZ* genes in ten tissues and under drought stress treatment.

Additional file 4: Figure S3 Hierarchical clustering of the relative expression level of family I *TaHDZ* genes in ten different organs or tissues. The heat map was drawn in Log_{10}-transformed expression values. The red or green colors represent the higher or lower expression level of each transcript in each sample.

Additional file 5: Figure S4 Hierarchical clustering of the relative expression level of family II *TaHDZ* genes in ten different organs or tissues. The heat map was drawn in Log_{10}-transformed expression values. The red or green colors represent the higher or lower expression level of each transcript in each sample.

Additional file 6: Figure S5 Hierarchical clustering of the relative expression level of family III *TaHDZ* genes in ten different organs or tissues. The heat map was drawn in Log_{10}-transformed expression values. The red or green colors represent the higher or lower expression level of each transcript in each sample.

Additional file 7: Figure S6 Hierarchical clustering of the relative expression level of family IV *TaHDZ* genes in ten different organs or tissues. The heat map was drawn in Log_{10}-transformed expression

values. The red or green colors represent the higher or lower expression level of each transcript in each sample.

Additional file 8: Table S2 Potential *cis*-elements within a 2 kb region upstream from the start codon of each *TaHDZ* gene.

Additional file 9: Figure S7 Protein sequence alignment of TaHDZ5-6A and TaHDZ5-6D. Identical amino acids are shaded in black, and similar amino acids are shaded in gray.

Additional file 10: Figure S8 The expression patterns of *TaHDZ5* in wheat. A The expression profiles of *TaHDZ5* in different tissues. R, root of wheat seedling at five-leaf stage; S, stem of wheat seedling at five-leaf stage; L, leaf of wheat seedling at five-leaf stage; FL, flag leaf at heading stage; YS5, young spike at early booting stage; YS15, spike at heading stage; GR5, grain of 5 days post-anthesis; GR15, grain of 15 days post-anthesis. **B** The expression pattern of *TaHDZ5* under drought stress treatment. The error bars indicate standard deviations derived from three independent biological experiments.

Additional file 11: Table S3 Genes up or downregulated in *35S:TaHDZ5-6A* transgenic *Arabidopsis* relative to WT plants under well-watered condition. Genes with an average fold change (FC) > 2.0 and a corrected-*P* < 0.001 are shown. The gene functional description is based on TAIR 10.

Additional file 12: Table S4 Primers used in this research. The name of the primers was based on the gene name and experimental purpose.

Figures

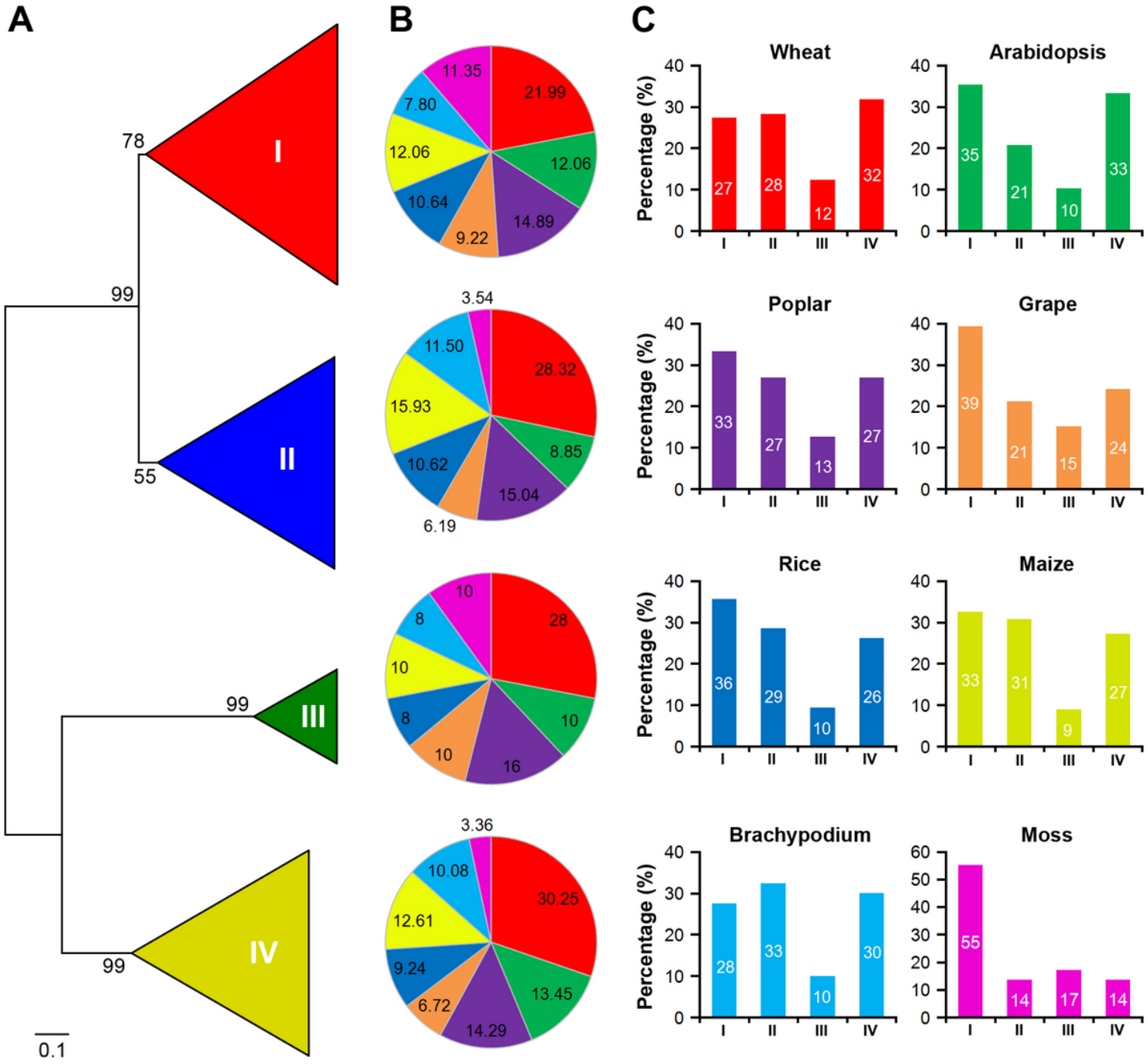

Figure 1

Phylogeny and distribution of HD-Zip proteins from eight plant species. A Phylogenetic tree of HD-Zip proteins from Arabidopsis, Populus, Vitis, wheat, rice, maize, Brachypodium, and moss. Phylogeny was constructed by PhyML using maximum likelihood analysis. Bootstrap support values as percentage, are shown on selected major branches. The bar indicates substitutions per site; B Percentage representation of HD-Zips across the eight plant species within each subfamily; C Percentage representation of distributions for HD-Zips within each plant species.

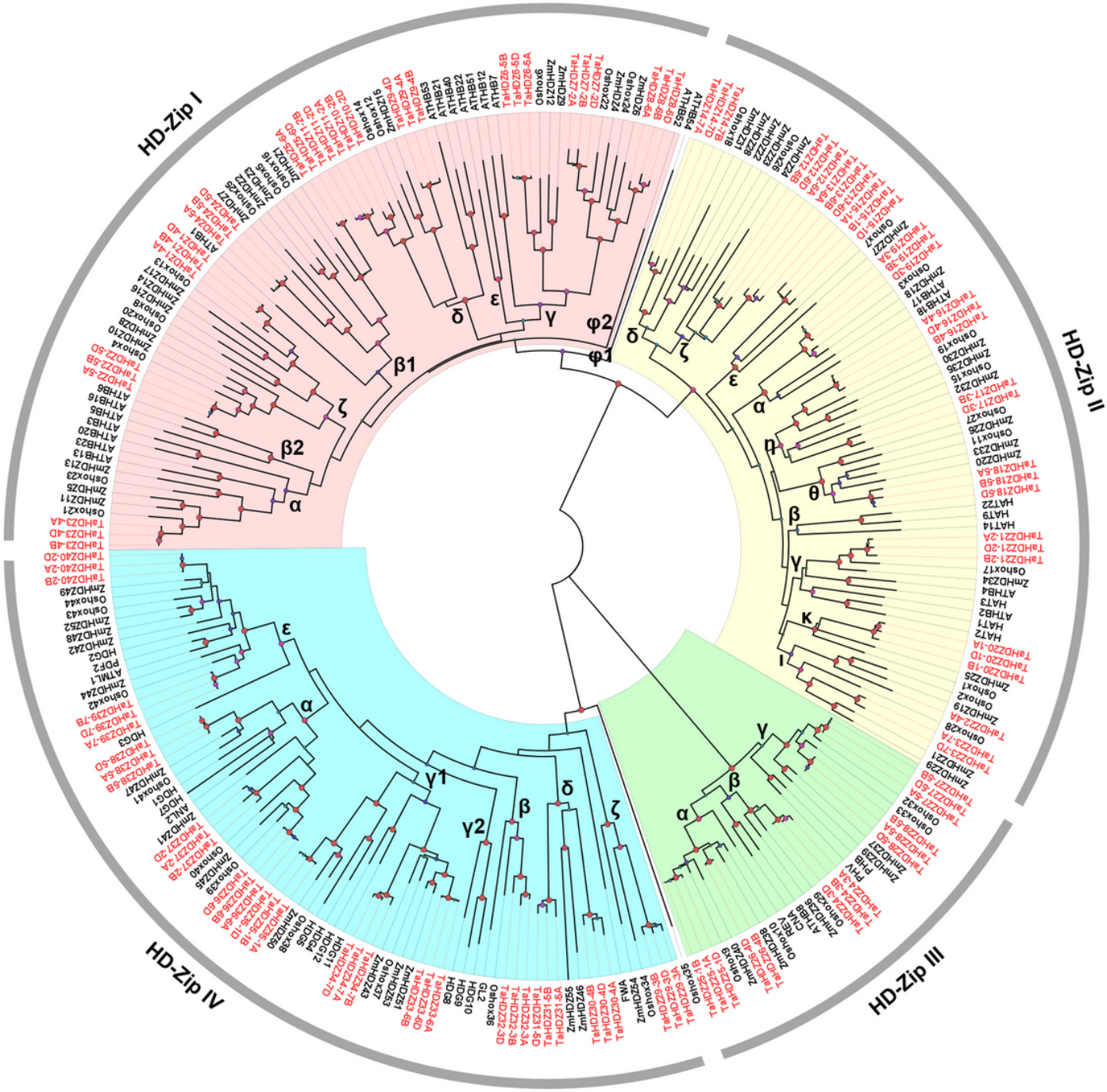

Figure 2

The phylogenetic tree of HD-Zip proteins from wheat, Arabidopsis, maize and rice. Members of the HD-zip genes from wheat are marked in red. Two-letter prefixes for sequence identifiers indicate species of origin. Ta, Triticum aestivum; At, Arabidopsis thaliana; Os, Oryza sativa; Zm, Zea mays. The tree was constructed using the Neighbor-Joining algorithm with 1000 bootstrap based on the full length sequences of HD-Zip proteins. The HD-Zip proteins are grouped into four distinct groups.

Figure 3

Phylogenetic relationships and gene structures of wheat HD-Zip genes. A Phylogenetic tree of 113 full length HD-Zip proteins from wheat were constructed by MEGA 6.0 using the Neighbour-Joining (NJ) method with 1,000 bootstrap values. B Exon/intron structures of wheat HD-Zip genes. Exons and introns are represented by purple boxes and grey lines, respectively. C The distribution of intron numbers between four distinct HD-Zip subfamily of wheat.

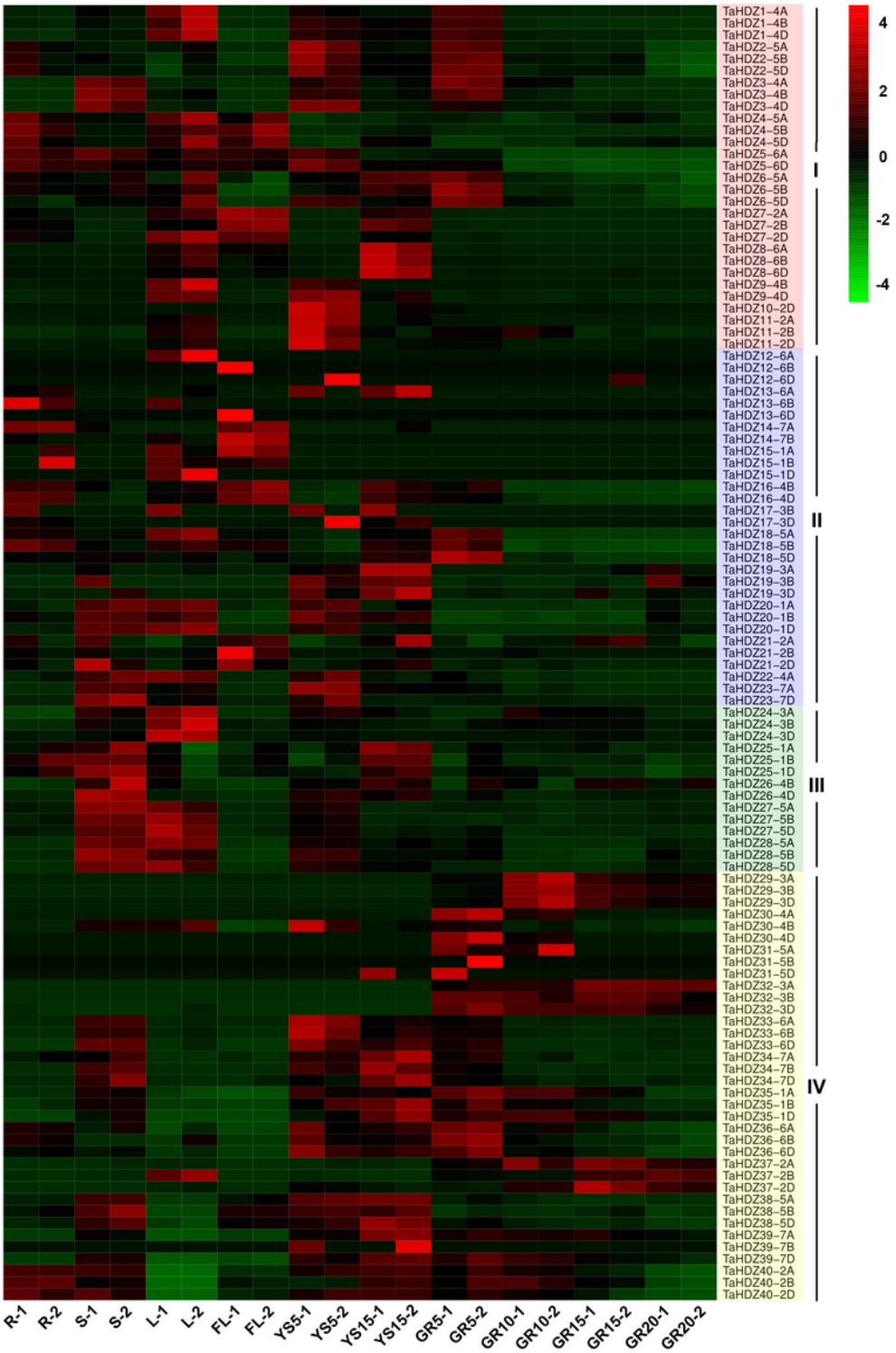

Figure 4

Expression profiles of TaHDZ genes in ten different organs or tissues. The heat map was drawn in Log10-transformed expression values. The red or green colors represent the higher or lower expression level of each transcript in each sample. R, root of wheat seedling at five-leaf stage; S, stem of wheat seedling at five-leaf stage; L, leaf of wheat seedling at five-leaf stage; FL, flag leaf at heading stage; YS5, young spike

at early booting stage; YS15, spike at heading stage; GR5, grain of 5 days post-anthesis; GR10, grain of 10 days post-anthesis; GR15, grain of 15 days post-anthesis; GR20, grain of 20 days post-anthesis.

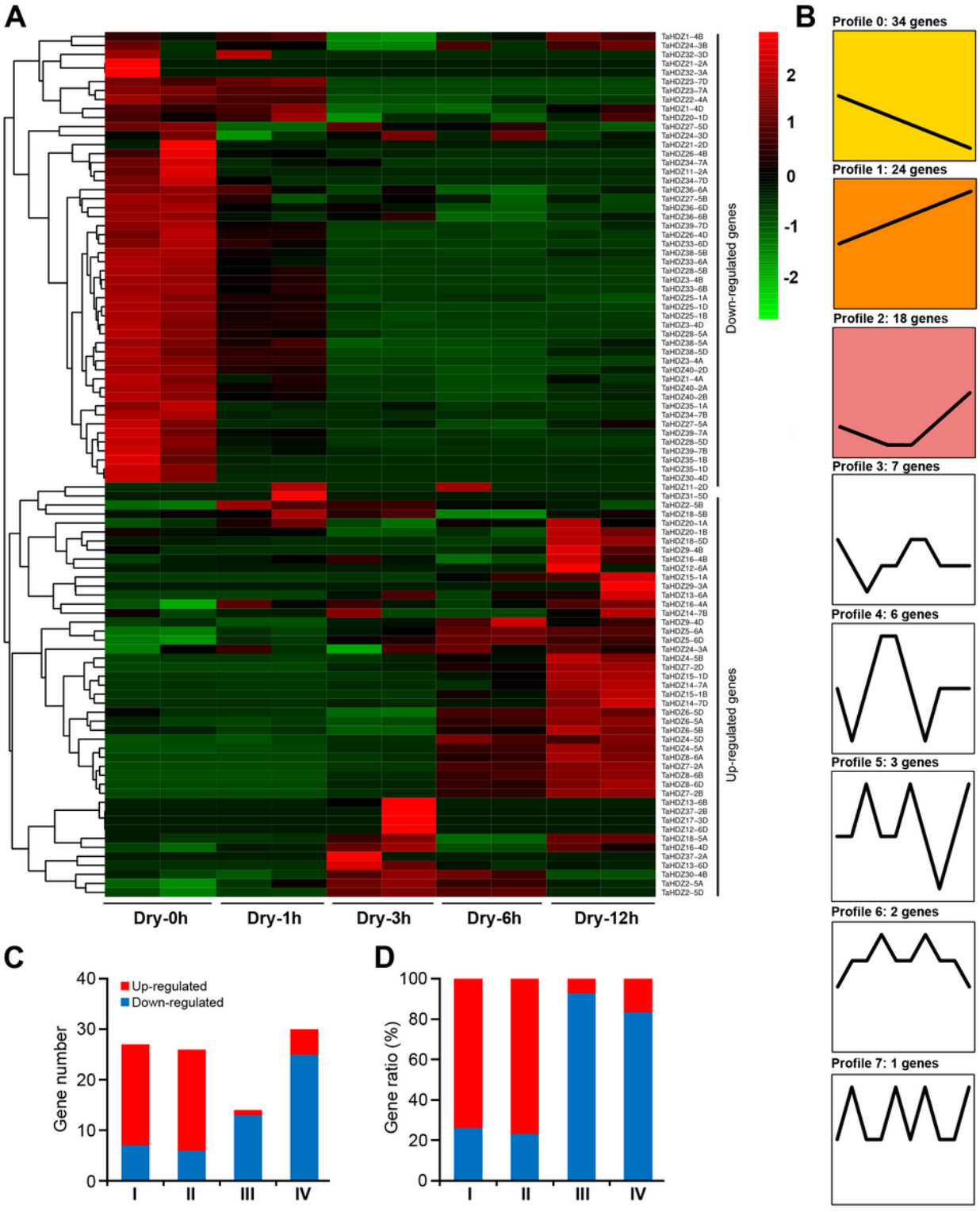

Figure 5

Expression profiles of TaHDZ genes in seedling leaves under drought stress treatment. A hierarchical clustering of the relative expression level of TaHDZ genes under drought stress treatment. The heat map was drawn in Log10-transformed expression values. The red or green colors represent the higher or lower

relative abundance of each transcript in each sample. B Expression patterns of TaHDZ genes under drought stress treatment. C The numbers of up-regulated and down-regulated TaHDZ genes in four HD-Zip subfamilies. D The ratios of up-regulated and down-regulated TaHDZ genes in four HD-Zip subfamilies.

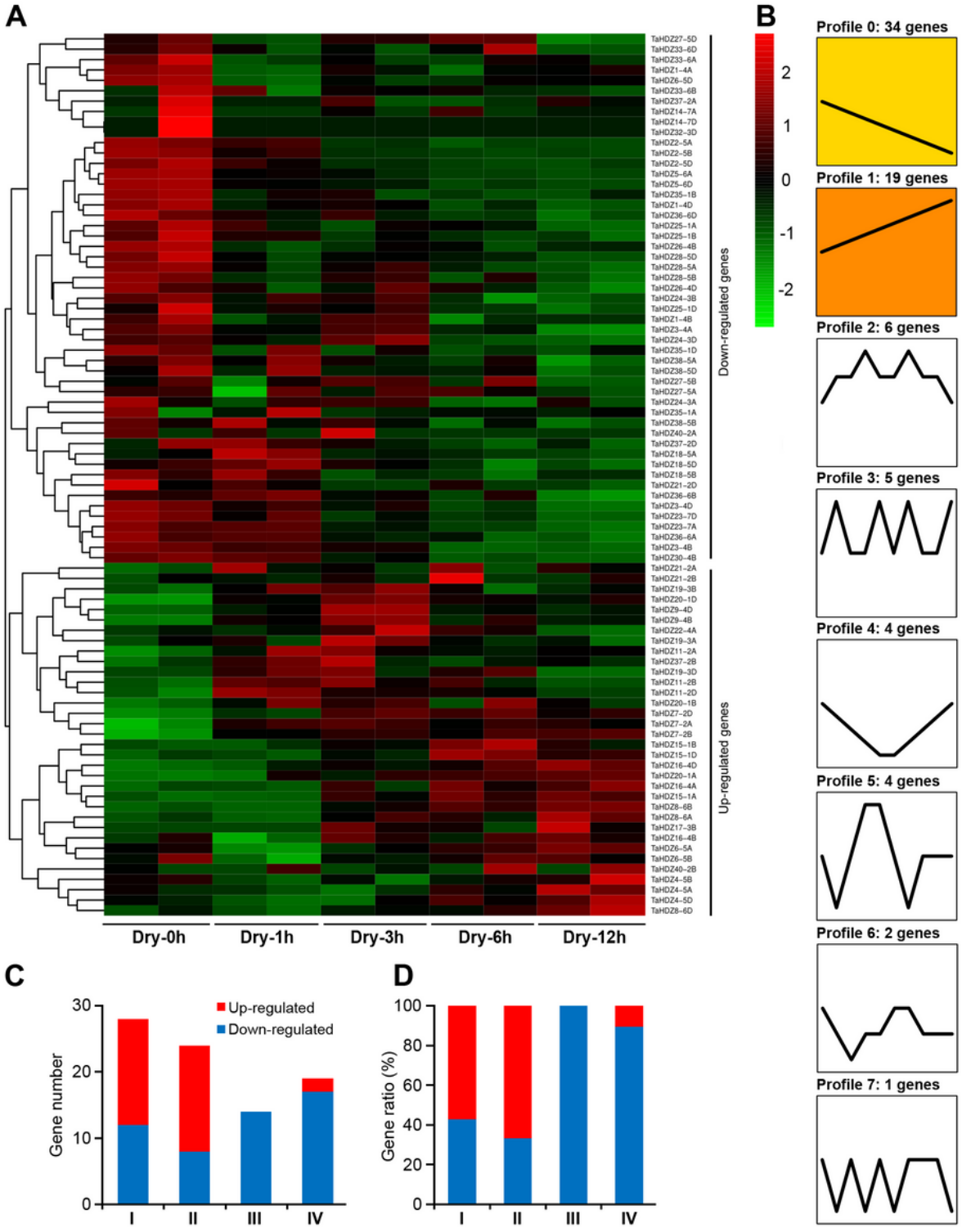

Figure 6

Expression profiles of TaHDZ genes in seedling roots under drought stress treatment. A Hierarchical clustering of the relative expression level of TaHDZ genes under drought stress treatment. The heat map was drawn in Log10-transformed expression values. The red or green colors represent the higher or lower relative abundance of each transcript in each sample. B Expression patterns of TaHDZ genes under drought stress treatment. C The numbers of up-regulated and down-regulated TaHDZ genes in four HD-Zip subfamilies. D The ratios of up-regulated and down-regulated TaHDZ genes in four HD-Zip subfamilies.

Figure 7

Phenotype of the 35S:TaHDZ5-6A transgenic Arabidopsis. A RT-PCR analysis of TaHDZ5-6A transcript levels in the three transgenic lines. B Statistical analysis of survival rates after the drought-stress treatment. The average survival rates and standard errors were calculated based on data obtained from three independent experiments. Significant differences were determined by a t-test. *P < 0.05, **P < 0.01. C Drought tolerance of 35S:TaHDZ5-6A transgenic Arabidopsis. Photographs were taken before and after the drought treatment, and followed by a six-day period of re-watering. D Stomatal aperture of WT and 35S::TaHDZ5-6A transgenic plants under normal and drought conditions. E Statistical analysis of stomatal aperture of WT and 35S::TaHDZ5-6A transgenic plants. Values are mean ratios of width to length. Error bars represent standard errors of three independent experiments (n = 60). Bars, 10 μm. F Water loss from detached rosettes of WT and 35S::TaHDZ5-6A transgenic plants. Water loss was expressed as the percentage of initial fresh weight. Values are means from eight plants for each of three independent experiments. Significant differences were determined by a t-test. *P < 0.05, **P < 0.01. G Free proline content of WT and 35S::TaHDZ5-6A transgenic plants under normal and drought stress treatment.

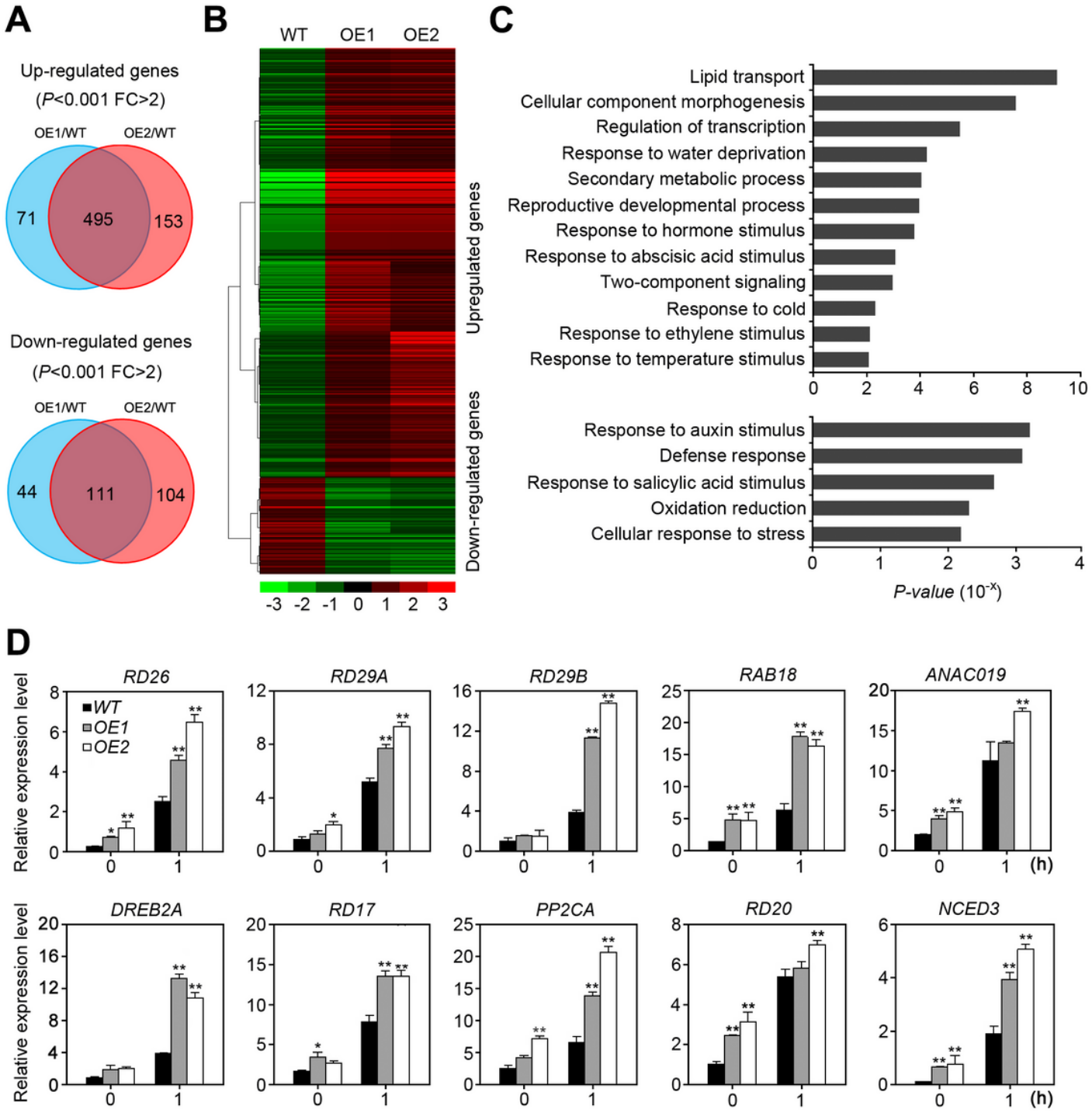

Figure 8

Transcriptomic analyses of the 35S::TaHDZ5-6A transgenic Arabidopsis under normal condition. A venn diagrams of up- or down-regulated genes in transgenic plants relative to WT plants using a significance cutoff of P <0.001, and a fold-change (FC) >2. B Hierarchical clustering of up- or down-regulated genes in 35S::TaHDZ5-6A transgenic Arabidopsis lines relative to WT plants. The indicated scale is the log2 value of the normalized level of gene expression. C Gene ontology of biological pathways (GOBPs) enriched in TaHDZ5-6A transgenic plants based on up or downregulated genes. D qRT-PCR analysis of drought

induced genes in the transgenic and WT plants under normal and drought conditions. The error bars indicate standard deviations derived from three independent biological experiments. Significant differences were determined by a t-test. *P < 0.05, **P < 0.01.

Supplementary Files

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

- Additionalfile7.tif
- Additionalfile8.xlsx
- Additionalfile3.xlsx
- Additionalfile6.tif
- Additionalfile1.tif
- Additionalfile10.tif
- Additionalfile4.tif
- Additionalfile12.xlsx
- Additionalfile11.xlsx
- Additionalfile9.tif
- Additionalfile2.tif
- Additionalfile5.tif