Comparative Genome-wide Analysis and Expression Profiling of Histone Acetyltransferases and Histone Deacetylases Involved in the Response to Drought in Wheat

Hua Li (lihuahnnd@163.com)
Henan Agricultural University
Huajie Liu
Henan Agricultural University
Xinxin Pei
Henan Agricultural University
Hongyu Chen
Henan Agricultural University
Xiao Li
Henan Agricultural University
Jiarui Wang
Henan Agricultural University
Chenyang Wang
Henan Agricultural University

Research Article

Keywords: histone acetyltransferases, histone deacetylases, gene expression, drought stress, wheat

Posted Date: December 15th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-123885/v1

Version of Record: A version of this preprint was published at Journal of Plant Growth Regulation on April 13th, 2021. See the published version at https://doi.org/10.1007/s00344-021-10364-9.

Abstract

Background: Histone acetyltransferases (*HATs*) and histone deacetylases (*HDACs*) contribute to plant growth, development, and stress responses. A number of *HAT* and *HDAC* genes have been identified in several plants. However, wheat *HATs* and *HDACs* have not been comprehensively characterized. In this study, we identified *TaHATs* and *TaHDACs* in the wheat genome using bioinformatics tools.

Result: In total, 30 *TaHAT* genes and 53 *TaHDAC* genes were detected in the wheat genome. As described in other plants, *TaHATs* were classified into four subfamilies (i.e., GNAT, p300/CBP, MYST, and TAFII250) and *TaHDACs* were divided into three subfamilies (i.e., RPD3/HDA1, HD2, and SIR2). Phylogenetic and conserved domain analyses showed that *TaHATs* and *TaHDACs* are highly similar to those in *Arabidopsis* and rice; however, divergence and expansion from *Arabidopsis* and rice were also observed. We detected many stress-related cis-regulatory elements in the promoter regions of these genes (i.e., ABRE, STRE, MYB et al.). Further, based on a comparative expression analyses of three varieties with different degrees of drought resistance under drought stress, we found that *TaHAG2*, *TaHAG3*, *TaHAC2*, *TaHDA18*, *TaHDT1*, and *TaHDT2* are likely regulate drought stress in wheat.

Conclusions: In this study, *TaHATs* and *TaHDACs* from the wheat genome were identified. Three *TaHATs* and three *TaHDACs* were very likely to regulate drought stress based on a promoter analysis and gene expression analysis. These results provide a foundation for further research on the regulation of acetylation in wheat and its role in the response to drought stress.

Background

Plant growth, development, and stress responses depend on precise and accurate gene regulatory systems. Epigenetic modification is an important regulatory mechanism underlying gene expression and includes the regulation of transcription via DNA methylation or histone modifications, without changing the DNA sequence. Histone modifications mainly occur on lysine, arginine, and serine residues at the N-terminus of histone H3 or H4; the modification types include methylation, acetylation, phosphorylation, ubiquitination, and glycosylation [1]. Among these, acetylation is one kind of well-characterized histone post-translational modification. Histone acetylation mainly occurs on the lysine residues at the N-terminus of histones. It is completed by two dynamic and reversible processes, histone acetylation and deacetylation [2]. Histone acetylation can weaken the interaction between histones and negatively charged DNA, change the surface structure of nucleosomes, and stretch chromatin [3–5], which is conducive to the formation of binding sites for transcriptional regulators. Therefore, histone acetylation tends to induce gene activation. Conversely, histone deacetylation often causes chromatin condensation and inhibits gene transcription [6].

The processes of histone acetylation and deacetylation are catalyzed by histone acetyltransferases (HATs) and histone deacetylases (HDACs), respectively. HATs and HDACs regulate the dynamic balance of histone acetylation levels. In plants, HATs are classified into four distinct subfamilies according to their

structures: GNAT (GCN5-related N-terminal acetyltransferase), p300/CBP (CREB-binding protein), MYST (MOZ, Ybf2/Sas3, Sas2, and Tip60), and TAFII250 (TATA binding protein-associated factors) [7]. HDACs can be divided into three subfamilies: RPD3/HDA1 (Reduced Potassium Dependency 3/Histone DeAcetylase 1), SIR2 (Silent Information Regulator 2), and HD2 (Histone Deacetylase 2) [8]. The RPD3/HDA1 and SIR2 subfamilies are homologous to yeast HDACs and the HD2 subfamily is plant-specific. Accumulating studies have shown that both HATs and HDACs are involved in the regulation of plant growth and development as well as the response to stress [9–12]. For example, *Arabidopsis GCN5* mutants lose their apical advantage and show abnormal anther development [13]. In addition, *GCN5* mutations or downregulation increase plant sensitivity to salt, drought, heat, and disease stresses [14–16]. The MYST family members *HAM1* and *HAM2* are involved in the regulation of flowering time, and their inhibition leads to early flowering in *Arabidopsis* [17]. The histone deacetylase *HDA9* prevents seedling traits and germination [18] and negatively regulates salt and drought stress responsiveness in *Arabidopsis* [19]. *HDA19* is involved in drought, heat, and salt responses [20–23]. The overexpression of *HDA705* in rice decreases ABA and salt stress resistance during seed germination and enhances osmotic stress resistance during the seedling stage [24].

There is accumulating evidence for the important roles of HATs and HDACs in plant development and responses to environmental stresses. Therefore, to better understand the changes in histone acetylation, it is necessary to comprehensive identify and analyze HATs and HDACs in plant genomes. Genome-wide analyses of HATs have been performed in *Arabidopsis* [7], rice (*Oryza sativa*) [25], *Brachypodium distachyon* [26], litchi (*Litchi chinensis* Sonn) [27], *Vitis vinifera* [28], *Citrus sinensis* [29], and cotton (*Gossypium raimondi/ arboretum*) [30], and HDACs have been identified in *Arabidopsis* [31], rice [32], soybean [33], *Vitis vinifera* [28], litchi [27], tomato [34], and cotton [35]. However, little is known about HATs and HDACs in wheat (*Triticum aestivum*).

In this study, we first identified genes encoding HATs and HDACs in the wheat genome and then comprehensively analyzed these loci with respect to phylogenetic relationships, gene domain and structure, chromosomal location, and cis-regulatory elements in their promoters. Furthermore, the gene expression patterns of *HATs* and *HDACs* in different tissues (i.e., leaf, root, spike, and grain) and under drought stress were analyzed. This study could provide a foundation for further research on the regulation of histone acetylation in wheat and closely related monocots.

Results

Identification of HATs and HDACs protein in wheat

To identify HATs and HDACs in the genomes of wheat, a systematic blast search was performed using *Arabidopsis* and rice sequences as queries. Pfam and InterProScan databases were used to further verify the candidate HATs and HDACs based on structural domains. In total, 30 HATs and 53 HDACs were identified in the wheat genome (Table 1, Table S1). The polypeptide lengths of HATs and HDACs were 438–1796 and 309–693 amino acids, respectively, the predicted molecular weights were 50.14–201.47

and 33.16–74.54 kDa, and the theoretical isoelectronic point (pI) values were 4.71–8.9 and 4.6–9.42. In addition, the intron–exon organization of these *HATs* and *HDACs* were analyzed. The numbers of conserved coding regions ranged from 9 to 21 in *HATs* and from 1 to 17 in *HDACs*. With respect to subcellular localization, most HATs were detected in the nucleus, but HDACs were located in the nucleus as well as in the chloroplasts, cytoplasm, mitochondria, cytoskeleton, etc.

Table 1
Histone acetyltransferases (HATs) and Histone deacetylases (HDACs) identified in wheat

Subfamily	Gene name	Gene ID	Protein length	Localization
HATs family				
GNAT	TaHAG1A	TraesCS1A02G138200.2	507	nucl, chlo
	TaHAG1D	TraesCS1D02G134200.1	507	nucl, chlo
	TaHAG1U	TraesCSU02G003200.1	507	nucl, chlo
	TaHAG2A	TraesCS5A02G197700.1	463	cyto
	TaHAG2B	TraesCS5B02G186000.1	463	cyto
	TaHAG2D	TraesCS5D02G193200.1	463	cyto
	TaHAG3A	TraesCS2A02G320900.1	569	cyto
	TaHAG3B	TraesCS2B02G361800.1	569	cyto
	TaHAG3D	TraesCS2D02G341600.1	569	cyto
MYST	TaHAG4A	TraesCS2A02G159700.1	438	nucl, cyto
	TaHAG4B	TraesCS2B02G185300.1	482	nucl, mito
	TaHAG4D	TraesCS2D02G166900.1	438	nucl, cyto
CBP	TaHAC1A	TraesCS3A02G524800.1	1286	nucl
	TaHAC1B	TraesCS3B02G592100.1	1288	nucl
	TaHAC1D	TraesCS3D02G530000.2	1286	nucl
	TaHAC2A	TraesCS2A02G039500.3	1186	nucl
	TaHAC2B	TraesCS2B02G052300.4	1186	nucl
	TaHAC2D	TraesCS2D02G038100.1	1185	nucl
	TaHAC4A	TraesCS7A02G414500.1	1518	nucl
	TaHAC4B	TraesCS7B02G314400.2	1512	nucl, pero
	TaHAC4D	TraesCS7D02G407600.2	1518	nucl, pero
	TaHAC5A	TraesCS6A02G107300.1	1726	nucl
	TaHAC5B	TraesCS6B02G135800.1	1726	nucl
	TaHAC5D	TraesCS6D02G095400.1	1728	nucl
TAFII250	TaHAF1A	TraesCS7A02G515000.1	1782	chlo
	TaHAF1B	TraesCS7B02G431700.2	1762	nucl

Subfamily	Gene name	Gene ID	Protein length	Localization
	TaHAF1D	TraesCS7D02G505400.1	1762	nucl
	TaHAF2A	TraesCS7A02G514800.1	1796	nucl
	TaHAF2B	TraesCS7B02G431500.1	1796	nucl
	TaHAF2D	TraesCS7D02G505200.1	1796	nucl
HDACs family				
RPD3/HDA1	TaHDA2A	TraesCS7A02G362600.1	355	chlo, cyto
	TaHDA2B	TraesCS7B02G266000.1	353	chlo
	TaHDA2D	TraesCS7D02G360500.1	353	chlo
	TaHDA5A	TraesCS1A02G317100.1	397	chlo
	TaHDA5B	TraesCS1B02G329500.1	390	cyto, chlo
	TaHDA5D	TraesCS1D02G317100.1	394	chlo
	TaHDA6A	TraesCS6A02G181100.1	458	cysk
	TaHDA6B	TraesCS6B02G210200.1	458	cysk, cyto
	TaHDA6D	TraesCS6D02G168400.1	458	cysk
	TaHDA7A	TraesCS6A02G184100.2	519	nucl
	TaHDA7B	TraesCS6B02G212600.3	520	nucl
	TaHDA7D	TraesCS6D02G171000.1	574	chlo, nucl
	TaHDA8A	TraesCS1A02G275300.1	391	chlo
	TaHDA8B	TraesCS1B02G284500.1	393	chlo
	TaHDA8D	TraesCS1D02G274900.1	366	chlo
	TaHDA9A	TraesCS2A02G293200.1	430	cyto
	TaHDA9B	TraesCS2B02G309700.1	430	cyto
	TaHDA9D	TraesCS2D02G291000.1	430	cyto
	TaHDA14A	TraesCS5A02G119300.2	444	chlo
	TaHDA14B	TraesCS5B02G121300.1	453	chlo
	TaHDA14D	TraesCS5D02G126600.1	444	chlo
	TaHDA15A	TraesCS5A02G065300.1	614	nucl
	TaHDA15B	TraesCS5B02G072100.1	614	nucl

Subfamily	Gene name	Gene ID	Protein length	Localization
	TaHDA15D	TraesCS5D02G076100.1	612	nucl
	TaHDA18A	TraesCS2A02G177100.1	693	chlo, nucl
	TaHDA18B	TraesCS2B02G204100.1	693	chlo, nucl
	TaHDA18D	TraesCS2D02G185200.1	693	nucl, chlo
	TaHDA19A	TraesCS7A02G365600.3	523	nucl
	TaHDA19B	TraesCS7B02G261800.1	519	nucl
	TaHDA19D	TraesCS7D02G356800.1	519	nucl
	TaHDA20A	TraesCS4A02G213200.1	471	mito, nucl
	TaHDA20B	TraesCS4B02G102600.1	471	mito, nucl
	TaHDA20D	TraesCS4D02G100000.1	471	cyto, nucl
	TaHDA21A	TraesCS5A02G295000.1	484	cyto
	TaHDA21D	TraesCS5D02G302400.1	495	cyto
	TaHDA22B	TraesCS3B02G318000.1	380	cysk
	TaHDA22D	TraesCS3D02G422300.1	327	cyto
HD2	TaHDT1A	TraesCS1A02G445700.4	309	nucl
	TaHDT1D	TraesCS1D02G454400.2	311	nucl
	TaHDT2A	TraesCS3A02G415200.1	403	nucl
	TaHDT2B	TraesCS3B02G450300.1	383	nucl
	TaHDT2D	TraesCS3D02G410300.2	364	nucl
	TaHDT3B	TraesCS3B02G450400.1	378	nucl
	TaHDT3D	TraesCS3D02G410400.1	432	nucl
	TaHDT4A	TraesCS5A02G158900.1	432	nucl
	TaHDT4B	TraesCS5B02G156700.1	433	nucl
	TaHDT4D	TraesCS5D02G164000.1	433	nucl
SIR2	TaSRT1A	TraesCS2A02G077800.1	440	cyto, nucl
	TaSRT1B	TraesCS2B02G092700.1	465	cyto, nucl
	TaSRT1D	TraesCS2D02G075800.1	678	nucl, E.R
	TaSRT2A	TraesCS5A02G114700.3	414	cyto

Subfamily	Gene name	Gene ID	Protein length	Localization
	TaSRT2D	TraesCS5D02G124700.1	396	chlo
	TaSRT2U	TraesCSU02G136000.1	396	chlo

Phylogenetic and conserved domain analyses of HATs and HDACs in wheat

To reveal the evolutionary relationships among HATs and HDACs in wheat, a phylogenetic tree was constructed using MEGA 6.0 based on the amino acid sequences (Table S1) for the newly identified HAT and HDAC proteins in wheat and previously identified HATs from *Arabidopsis thaliana* and rice. Similar to *Arabidopsis* and rice, wheat HATs could be grouped into four distinct subfamilies: 12 HATs belonged to the CBP subfamily, 9 HATs belonged to the GNAT subfamily, 3 HATs belonged to the MYST subfamily, and 6 HATs belonged to the TAFII250 subfamily (Fig. 1, Table 1). The 53 HDACs in wheat could be classified into three subfamilies, RPD3/HDA1, HD2, and SIR2, with 37, 10, and 6 loci, respectively (Fig. 2, Table 1). *HATs* and *HDACs* in wheat were named based on the nomenclature suggestions for *Arabidopsis*; each gene was assigned a two-letter code corresponding to *T. aestivum* (Ta), followed by family designation and number, followed by A, B, or D (according to the subgenome in wheat).

In an analysis of domain architectures, all TaHAT subfamilies in wheat had conserved domains; for example, the CBP subfamily of wheat TaHATs contained the HAT-KAT11 domain, the GNAT subfamily of TaHATs contained the Hat1_N or Acetyltransferase domain, the MYST subfamily contained MOZ_SAS, zf-MYST, and Tudor-knot domains, and the TAFII250 subfamily contained DUF3591 and Bromodomain. In addition to these highly conserved domains, the CBP subfamily also contained the PHD, ZZ, and zf-TAZ domains, the GNAT subfamily had the Radical_SAM domain and Bromodomain, and the TBP-binding and ubiquitin domain was found in the TAFII250 subfamily (Fig. 1). For HDACs of wheat, RPD3/HDA1, HD2, and SIR2 subfamilies had the conserved domains Hist-deacetyl, NPL, and SIR2, respectively. In addition, TaHDT2 and TaHDT4 in the HD2 subfamily contained zf-C2H2_6 and FKBP_C domains, respectively. TaSRT2A in SIR2 subfamilies contained the Fibrillarin domain (Fig. 2). In general, wheat HATs and HDACs had similar domain organizations to those of their counterparts in *Arabidopsis* and rice.

Genomic localization of TaHATs and TaHDACs

The newly identified wheat *HATs* and *HDACs* were mapped to chromosomes. Both *TaHATs* and *TaHDACs* were unevenly distributed along the chromosomes (Fig. 3, Figure S1). In particular, there were no *TaHAT* genes on chromosomes 4A/B/D and 1B, three *TaHATs* were located on chromosomes 2A/B/D and 7 A/B/D, and the remaining chromosomes had only a single *TaHAT* gene. However, *TaHDACs* were distributed across all chromosomes, with the greatest number on chromosomes 5A/D (five *HDACs*) and

only one *TaHDAC* gene on chromosomes 3A, 4A/B/D, and Un. In terms of all *TaHATs* and *TaHDACs*, most were found on chromosomes 2A/B/D and 5 A/D, and the fewer were found on chromosomes 4A/B/D.

Putative cis-regulatory elements in the promoter regions of TaHATs and TaHDACs

To gain more insight into the putative functions of *TaHATs* and *TaHDACs*, the promoter region (1500 bp upstream of the transcription start site) was scanned using the PlantCARE database. Many putative cis-regulatory elements were detected in the promoters of both *TaHATs* and *TaHDACs* (Fig. 4; Table S2; Table S3), such as ABRE (abscisic acid-responsive element), STRE (stress-responsive element), ARE (essential for anaerobic induction), CCGTCC-box (meristem-specific activation), G-Box (light responsiveness), MYB (MYB-related binding sites), and TGA-element (auxin-responsive element). Most *TaHATs* and *TaHDACs* had ABRE (25 *HAT* genes and 43 *HDAC* genes) or STRE (27 *HAT* genes and 46 *HDAC* genes) elements. However, only *TaHDACs* had the GARE-motif (gibberellin-responsive element), indicating that the transmission and regulation of GA may be more closely related to histone deacetylation. Moreover, in *TaHATs*, the number of ABRE elements was higher in all genes in the GNAT subfamily as well as *TaHAC4A/B/D* and *TaHAC5A/B/D* in the CBP subfamily than in the TAFII250 subfamily, which contained few or no ABRE elements. Genes with a large number of ABRE elements in *HDACs* were *TaHDA5A/B/D*, *TaHDA8A/B/D*, and *TaHDA9B/D* in the RPD3/HDA1 subfamily, *TaHDT2A/B* in the HD2 subfamily, and *TaSRT1D* and *TaSRT2U* in the SIR2 subfamily. These genes may mediate the ABA signaling pathway. The TGA-element was only detected in GNAT subfamily genes, such as *TaHAG1A/B/U* and *TaHAG2A/B/D*, and was not observed in genes of the CBP and MYST subfamilies (except *TaHAC1D*). In general, the distribution of cis-acting elements was more similar in *TaHAT* subfamilies than in *TaHDACs* subfamilies.

Expression analysis of TaHATs and TaHDACs in different tissues

The RNA-seq data for different tissues were obtained from expVIP [36] (Fig. 5). All *TaHATs* and *TaHDACs* were differentially expressed in the leaf, root, spike, and grain. *TaHAG4A/B/D* was expressed in all four tissues and showed the highest expression levels among *TaHATs*. *TaHAC1A/B/D* and *TaHAF1B/D* had very low or no expression in these four tissues. The expression levels of *TaHAG2A/B/D* in the leaf and *TaHAC4B/D*, *TaHAC5A/B*, and *TaHAF2A/B* in the grain were nearly undetectable (Fig. 5A). *TaHDT1D* was expressed in all four tissues and showed relatively higher expression levels than those of other *TaHDACs*, while *TaHDA20A/B/D*, *TaHDA21A/D*, *TaHDA22D*, T*aHDT1A*, *TaHDT2A*, and *TaHDT3B/D* were almost undetectable in the four tissues (Fig. 5B). Additionally, the expression levels differed among A, B, and D genomes, e.g., *TaHDA19A* was not expressed in the leaf, root, or spike, while *TaHDA19B/D* were highly expressed. *TaHDA22B* was highly expressed in all four tissues, unlike *TaHDA22D* (Fig. 5B). These results suggested that the A, B, and D genomes may jointly contribute to their functional roles.

Expression analysis of TaHATs and TaHDACs in response to drought stress

The identification of putative cis-regulatory elements in the promoter regions suggested that *TaHATs* and *TaHDACs* contribute to the response to abiotic stresses. In the main wheat-producing area, plants often encounter drought, leading to reduced yields. Therefore, we focused on the expression of *TaHATs* and *TaHDACs* under drought stress. Drought resistance is significantly higher in the wheat variety BN207 than in its parents BN64 and ZM16 (Table S4). Therefore, we used these varieties to identify *TaHATs* and *TaHDACs* that may contribute to the response to drought stress by comparative expression analyses.

Similar to the tissue expression results, *TaHDA5*, *TaHDA20*, *TaHDA21*, and *TaHDT3* were not detected in the leaf by qRT-PCR. For the remaining *TaHATs* and *TaHDACs*, regardless of conditions (i.e., normal or drought), a number of *TaHATs* and *TaHDACs* were significantly differentially expressed among the three varieties. However, the expression levels of *TaHAG4* and *TaHAF1* in the *TaHAT* family and *TaHDA7*, *TaHDA15* and *TaSRT2* in the *TaHDACs* family were not affected or were slightly affected by drought stress in all three varieties (Fig. 6). This indicated that these genes may not be related to the regulation of drought stress. All other genes were up-regulated or down-regulated in at least one variety under drought stress. It is worth noting that *TaHAG2*, *TaHAG3*, and *TaHAC2* in the *TaHAT* family (Fig. 6A) and *TaHDA2*, *TaHDA18*, *TaHDT1*, and *TaHDT2* in the *TaHDAC* family (Fig. 6B) showed a significant response to drought stress only in BN207; in particular, *TaHAG2*, *TaHAG3*, *TaHAC2*, and *TaHDT1* were up-regulated under drought stress, while *TaHDA2*, *TaHDA18*, and *TaHDT2* were down-regulated. Further, the expression of these genes (except *TaHDA2*) in BN207 under drought stress was significantly different from levels in its parents BN64 and ZM16. Therefore, combined with the observation that BN207 had higher drought resistance than that of its parents BN64 and ZM16, our results suggested that these six genes were likely to mediate drought stress in wheat.

Discussion

Characterization of TaHATs and TaHDACs proteins

HATs and *HDACs* could mediate plant growth and development and the response to environmental stresses [10]. A number of *HATs* and *HDACs* have been identified in several plants. However, little is known about these enzymes in wheat. In this study, 30 TaHATs and 53 TaHDACs were identified in the wheat genome using bioinformatics tools (Table 1, Table S1). Wheat is hexaploid; accordingly, the numbers of TaHATs and TaHDACs are approximately three times those of *Arabidopsis* and rice [7, 25, 32]. All subfamilies of TaHATs and TaHDACs contained subfamily-specific domains, as in other plants. Further, TaHATs and TaHDACs also contain other conserved domains, such as zf-TAZ, ZZ, PHD, and Bromodomains. The zf-TAZ, ZZ, and PHD domains are thought to be involved in protein recognition and protein–protein interactions [37–38]. Bromodomains bind to acetylated lysine residues [39–40]. In general, the protein domains of wheat TaHATs and TaHDACs are highly similar to those of homologues in

Arabidopsis and rice, suggesting that all these genes in wheat have similar functions to those described in other plant species.

Evolution and function divergence of TaHATs and TaHDACs

All *HATs* (except *AtHAG5* and *AtHAC12*) in *Arabidopsis* correspond to homologous genes on the A, B, and D genomes of wheat (Fig. 1). Previous studies have also shown that homologues of *AtHAG5* and *AtHAC12* are lacking in *Brachypodium distachyon* [26]. Homologous genes of *AtHAC2* and *AtHAF2* were detected in wheat (Fig. 1) but not in *B. distachyon* [26]. Similar results were found in the *HADC* family; for example, ten genes in the HD2 subfamily were found in wheat, corresponding to four HD2 subfamily genes in *Arabidopsis* (*AtHD1*, *AtHD2*, *AtHD3*, and *AtHD4*) (Fig. 2), but only two HD2 subfamily genes exist in rice (*OsHD701* and *OsHD702*) [32]. In addition, *TaHDA21A/D* and *TaHDA22B/D* in the RPD3/HDA1 subfamily have the conserved Hist-deacetyl domain, but they have low homology with all RPD3/HDA1 subfamily members in *Arabidopsis*. These results indicated that compared with the dicotyledonous *Arabidopsis* and the monocotyledonous rice, wheat *TaHAT* and *TaHDAC* families have diverged and expanded.

A subcellular localization analysis showed that a majority of *TaHATs* (77%) are located in the nucleus. However, only 38% of *TaHDAC* family members were located in the nucleus; the others were mainly located in the chloroplast (30%) and cytoplasm (21%) (Table 1). These results indicated that *TaHATs* and *TaHDACs* may have vital functions other than histone acetylation. In particular, the *TaHDAC* family may have broader functions than those of the *TaHAT* family based on the predominant localization in the chloroplast and cytoplasm (Table 1). In fact, lysine-acetylated proteins are located in various cellular compartments in *Arabidopsis* and rice [41–43]. In rice, 45% of acetylated proteins are distributed in the chloroplast, followed by the cytoplasm (27%), and only 13% of acetylated proteins are located in the nucleus [43]. Similarly, a large number of acetylated proteins are related to photosynthesis in *Arabidopsis* [41], e.g., photosystem II (PSII) subunits, light-harvesting chlorophyll a/b-binding proteins (LHCb), Rubisco large and small subunits, and chloroplastic ATP synthase (b-subunit).

Expression level of TaHATs and TaHDACs under drought stress

Numerous studies have confirmed that *HATs* and *HDACs* are involved in the regulation of plant abiotic stresses. A number of histone acetylation modification genes that regulate drought resistance have recently been identified in various plants. For instance, the histone acetylase GCN5 in *Populus trichocarpa* can improve the drought resistance of plants by enhancing H3K9ac and enriching RNA polymerase II at *PtrNAC* genes [15]. *AtHDA9* can regulate the acetylation level in the promoters of stress-responsive genes and thus negatively regulate plant sensitivity to salt and drought stresses in *Arabidopsis* [19]. The over-expression of the *Populus* histone deacetylase gene *84KHDA903* in tobacco enhances drought tolerance [44]. The silencing of the tomato histone deacetylase gene *SlHDA5* results in reduced tolerance to salt, drought, and ABA [45]. The overexpression of the *Brachypodium* histone deacetylase *BdHD1* results in

hypersensitivity to ABA and enhanced drought resistance [46]. The overexpression of the cotton histone deacetylase *GhHDT4D* in *Arabidopsis* increases plant tolerance to drought [47]. In this study, we analyzed the expression changes of all *TaHATs* and *TaHDACs* under drought stress in different drought-resistant wheat varieties. Different *TaHATs* and *TaHDACs* had different expression profiles in the three varieties (Fig. 6). However, under drought stress, *TaHAG2*, *TaHAG3*, *TaHAC2*, and *TaHDT1* levels were significantly higher in the drought-resistant variety BN207 than in BN64 and ZM16. Conversely, *TaHDA18* and *TaHDT2* were significantly down-regulated by drought in BN207. BN64 and ZM16 are the parents of BN207; however, BN207 shows higher drought resistance (Table S4), implying that the six genes described above are likely to participate in the regulation of drought stress. Based on a promoter analysis, these six genes contain a large number of STRE or ABRE elements (Fig. 4). *TaHAG2*, *TaHAG3*, *TaHDA18*, and *TaHDT2* had multiple ABA elements, suggesting that the regulation of drought stress by these four genes may depend on the ABA signaling pathway. *TaHAC2* and *TaHDT1* contained no or few ABRE elements, indicating that the responses of these loci to drought may be ABA-independent.

Conclusions

HATs and *HDACs* exert important functions related to plant growth, development, and stress responses. In this study, *TaHATs* and *TaHDACs* in the wheat genome were identified, and the expression patterns of these genes in various tissues and under drought stress were analyzed. In total, 30 *TaHATs* and 53 *TaHDAC*s were found in wheat and were classified into four and three subfamilies, respectively. Further, based on expression analyses, we found that *TaHAG2*, *TaHAG3*, *TaHAC2*, *TaHDA18*, *TaHDT1*, and *TaHDT2* are closely related to drought stress. In summary, this study provides a foundation for further studies of *TaHATs* and *TaHDACs* in wheat and their regulatory roles in the response to drought stress.

Methods

Identification of HATs and HDACs

Wheat HAT and HDAC proteins were identified following previously described methods in studies of *Arabidopsis* [7] and rice [25]. The HAT and HDAC protein sequences from *Arabidopsis* and rice were downloaded from Phytozome (http://phytozome.jgi.doe.gov/pz/portal.html) and used as queries for searches against the *T. aestivum* L. genomes using BLASTP in Ensembl (http://plants.ensembl.org/index.html). Furthermore, the predicted HATs and HDACs were confirmed by searching the protein domains in Pfam (http://pfam.xfam.org/search) and InterPro database (http://www.ebi. ac.uk/interpro/search/sequence-search).

The physicochemical properties of predicted HAT and HDAC proteins were predicted using ExPASy (http://cn.expasy.org/tools). The subcellular localization of each HAT and HDAC protein was predicted using WoLF PSORT (https://www.genscript.com/wolf-psort.html).

Phylogenetic and domain analysis

To investigate evolutionary relationships, a multiple sequence alignment was generated for HATs and HDACs proteins in *Triticum aestivum*, *Oryza sativa,* and *Arabidopsis thaliana* using ClustalW with default parameters. An un-rooted phylogenetic tree was constructed based on the neighbor-joining method with 1000 bootstrap replicates using MEGA6 [48]. Domain annotation was performed using Pfam, and TBtools was used for visualization [49].

Chromosomal distribution and localization

The positional information of *TaHATs* and *TaHDACs* was downloaded from Ensembl database, and then their chromosomal distribution was mapped using TBtools [49].

Promoter cis-element analysis

The 1.5 kb upstream sequences of the transcription start site were downloaded from IWGSC (http://www.wheatgenome.org/), and then the PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) was used to analyze cis-regulatory elements.

Plant materials and stress treatments

The wheat cultivar BN207 and its parents BN64 and ZM16 plants were used for gene expression analyses. Seeds were sown into 27 × 26 cm pots filled with cultivated soil, with 12 plants per pot. Growth conditions and management were the same as those for field cultivation. Drought treatment was performed at the booting stage of wheat, using 75% soil water as a control and 55% soil water as a drought treatment. Flag leaves were collected for a gene expression analysis after 10 days of drought. After 15 days of treatment, plants were rehydrated and the yield was measured at the mature period.

Tissue expression analysis

The expression levels (TPM values) of *TaHATs* and *TaHDACs* in the leaf, root, spike, and grain were downloaded from expVIP (http://www.wheat-expression.com/) [36]. A heatmap was generated using HemI (http://hemi.biocuckoo.org/down.php).

Quantitative real-time PCR analysis

Total RNA was isolated using the Omniplant RNA Kit (CWBIO, Beijing, China) following the manufacture's protocols. Then, 2 μg of total RNA was used to synthesize cDNA with cDNA Synthesis Supermix (Vazyme, Nanjing, China). qRT-PCR was performed using AceQ^{TM} Universal SYBR qPCR Master Mix (Vazyme) and the ABI StepOnePlus Real-Time PCR System. The expression levels were calculated using the $2^{-\Delta Ct}$ method. *TaACT-1* was used as a housekeeping gene for normalization [50]. The primers used in this study are listed in Table S5.

Statistical analysis

The qRT-PCR results are reported as the means of three independent experiments. Differences among treatments were analyzed by one-way ANOVA and Duncan's multiple range tests, setting $p < 0.05$ as the threshold for significance.

Declarations

Acknowledgments:

Not applicable.

Author Contributions:

HL designed the experiments and wrote the manuscript. HL and HJL carried out the bioinformatics analysis. XXP and HYC performed the biological experiment. XL and JRW prepared the plant sample. CYW helped to revise the manuscript. All authors read and approved the final manuscript.

Funding:

This research was supported by the National Key Research and Development Program of China (2017YFD0300408).

Availability of data and materials:

All data generated or analyzed during this study are included in this published article and its Additional files. The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate:

Not applicable.

Consent for publication:

Not applicable.

Competing interests:

The authors declare that they have no competing interests.

References

1. Kouzarides T. Chromatin modifications and their function. Cell. 2007; 128:693–705.
2. Hassig CA, Schreiber SL. Nuclear histone acetylases and deacetylases and transcriptional regulation: HATs off to HDACs. Curr Opin Chem Biol. 1997; 1:300–308.
3. Hollender C, Liu Z. Histone deacetylase genes in *Arabidopsis*J Integr Plant Biol. 2008; 50:875–885.
4. Allis CD, Jenuwein T. The molecular hallmarks of epigenetic control. Nat Rev Genet. 2016; 17:487–500.
5. Onufriev AV, Schiessel H. The nucleosome: from structure to function through physics. Curr Opin Struc Biol. 2019; 56:119–130.
6. Shahbazian M, Grunstein M. Functions of site-specific histone acetylation and deacetylation. Annu Rev Biochem. 2007; 76:75–100.
7. Pandey R, MuÈller A, Napoli CA, Selinger DA, Pikaard CS, Richards EJ, et al. Analysis of histone acetyltransferase and histone deacetylase families of *Arabidopsis thaliana* suggests functional diversification of chromatin modification among multicellular eukaryotes. Nucleic acids Res. 2002; 30:5036–5055.
8. Ekwall K. Genome-wide analysis of HDAC function. Trends Genet. 2005; 21:608–615.
9. Wang Z, Cao H, Chen F, Liu Y. The roles of histone acetylation in seed performance and plant development. Plant Physiol Bioch. 2014; 84:125-133.
10. Zhao T, Zhan Z, Jiang D. Histone modifications and their regulatory roles in plant development and environmental memory. J Genet Genomics. 2019; 46:467–476.
11. Jiang J, Ding AB, Liu F, Zhong X. Linking signaling pathways to histone acetylation dynamics in plants. J Exp Bot. 2020; 71: 5179–5190.
12. Ueda M, Seki M. Histone modifications form epigenetic regulatory networks to regulate abiotic stress response. Plant Physiol. 2020; 182: 15–26.
13. Cohen R, Schocken J, Kaldis A, Vlachonasios KE, Hark AT, McCain ER. The histone acetyltransferase GCN5 affects the inflorescence meristem and stamen development in *Arabidopsis*. Planta. 2009; 230:1207–1221.
14. Hu Z, Song N, Zheng M, Liu X, Liu Z. Xing J, et al. Histone acetyltransferase GCN 5 is essential for heat stress-responsive gene activation and thermotolerance in Arabidopsis. Plant J. 2015; 84:1178–

1191.

15. Li S, Lin Y, Wang P, Zhang B, Li M, Chen S, et al. The AREB1 transcription factor influences histone acetylation to regulate drought responses and tolerance in *Populus trichocarpa*. Plant Cell. 2019; 31:663–686.
16. Zheng M, Liu X, Lin J, Liu X, Wang Z, Xin M, et al. Histone acetyltransferase GCN 5 contributes to cell wall integrity and salt stress tolerance by altering the expression of cellulose synthesis genes. Plant J. 2019; 97:587–602.
17. Xiao J, Zhang H, Xing L, Xu S, Liu H, Chong K, Xu Y. Requirement of histone acetyltransferases HAM1 and HAM2 for epigenetic modification of FLC in regulating flowering in *Arabidopsis*. J Plant Physiol. 2013; 170:444–451.
18. van Zanten M, Zöll C, Wang Z, Philipp C, Carles A, Li Y, et al. HISTONE DEACETYLASE 9 represses seedling traits in Arabidopsis thaliana dry seeds. Plant J. 2014; 80:475–488.
19. Zheng Y, Ding Y, Sun X, Xie S, Wang D, Liu X, Su L, Lei Pan W, Zhou DX. Histone deacetylase HDA9 negatively regulates salt and drought stress responsiveness in *Arabidopsis*. J Exp Bot. 2016; 67:1703–1713.
20. Chen LT, Wu K. Role of histone deacetylases HDA6 and HDA19 in ABA and abiotic stress response. Plant Signal Behav. 2010; 5: 1318–
21. Mehdi S, Derkacheva M, Ramström M, Kralemann L, Bergquist J, Hennig L. The WD40 domain protein MSI1 functions in a histone deacetylase complex to fine-tune abscisic acid signaling. Plant Cell. 2016; 28:42–54.
22. Ueda M, Matsui A, Tanaka M, Nakamura T, Abe T, Sako K, et al. The distinct roles of class I and II RPD3-like histone deacetylases in salinity stress response. Plant Physiol. 2017; 175:1760–1773.
23. Ueda M, Matsui A, Nakamura T, Abe T, Sunaoshi Y, Shimada H, Seki M. Versatility of HDA19-deficiency in increasing the tolerance of *Arabidopsis* to different environmental stresses. Plant Signal Behav. 2018; 13:e1475808.
24. Zhao J, Li M, Gu D, Liu X, Zhang J, Wu K, Zhang X, Teixeira da Silva J, Duan J. Involvement of rice histone deacetylase HDA705 in seed germination and in response to ABA and abiotic stresses. Biochem Bioph Res Co. 2016; 470:439–444.
25. Liu X, Luo M, Zhang W, Zhao J, Zhang J, Wu K, et al. Histone acetyltransferases in rice (*Oryza sativa*): phylogenetic analysis, subcellular localization and expression. BMC Plant Biol. 2012; 12:145.
26. Tan S, Gao L, Li T, Chen L. Phylogenetic and expression analysis of histone acetyltransferases in *Brachypodium distachyon*. Genomics. 2019; 111:1966–1976.
27. Peng M, Ying P, Liu X, Li C, Xia R, Li J, Zhao M. Genome-wide identification of histone modifiers and their expression patterns during fruit abscission in Litchi. Front Plant Sci. 2017; 8:639.
28. Aquea F, Timmermann T, Arce-Johnson P. Analysis of histone acetyltransferase and deacetylase families of *Vitis vinifera*. Plant Physiol Bioch. 2010; 48:194–199.

29. Xu J, Xu H, Liu Y, Wang X, Xu Q, Deng X. Genome-wide identification of sweet orange (*Citrus sinensis*) histone modification gene families and their expression analysis during the fruit development and fruit-blue mold infection process. Front Plant Sci. 2015; 6:607.
30. Imran M, Shafiq S, Farooq MA, Naeem MK, Widemann E, Bakhsh A, et al. Comparative genome-wide analysis and expression profiling of histone acetyltransferase (HAT) gene family in response to hormonal applications, metal and abiotic stresses in cotton. Int J Mol Sci. 2019; 20:5311.
31. Hollender C, Liu Z. Histone deacetylase genes in *Arabidopsis*J Integr Plant Biol. 2008; 50:875–885.
32. Hu Y, Qin F, Huang L, Sun Q, Li C, Zhao Y, Zhou DX. Rice histone deacetylase genes display specific expression patterns and developmental functions. Biochem Bioph Res Co. 2009; 388:266–271.
33. Liew LC, Singh MB, Bhalla PL. An RNA-seq transcriptome analysis of histone modifiers and RNA silencing genes in soybean during floral initiation process. PLoS One. 2013; 8:e77502.
34. Zhao L, Lu J, Zhang J, Wu PY, Yang S, Wu K. Identification and characterization of histone deacetylases in tomato (*Solanum lycopersicum*). Front Plant Sci. 2015; 5:760.
35. Imran M, Shafiq S, Naeem MK, Widemann E, Munir MZ, Jensen KB, Wang R. Histone Deacetylase (HDAC) Gene Family in Allotetraploid Cotton and Its Diploid Progenitors: In Silico Identification, Molecular Characterization, and Gene Expression Analysis under Multiple Abiotic Stresses, DNA Damage and Phytohormone Treatments. Int J Mol Sci. 2020; 21:321.
36. Ramírez-González RH, Borrill P, Lang D, Harrington SA, Brinton J, Venturini L, et al. The transcriptional landscape of polyploid wheat. Science. 2018; 361:
37. Bienz M. The PHD finger, a nuclear protein-interaction domain. Trends Biochem. Sci. 2006; 31:35–40.
38. Gamsjaeger R, Liew CK, Loughlin FE, Crossley M, Mackay JP. Sticky fingers: Zinc-fingers as protein-recognition motifs. Trends Biochem Sci. 2007; 32:63–70.
39. Marmorstein R, Berger SL. Structure and function of bromodomains in chromatin-regulating complexes. Gene. 2001; 272:1–9.
40. Servet C, e Silva NC, Zhou DX. Histone acetyltransferase AtGCN5/HAG1 is a versatile regulator of developmental and inducible gene expression in *Arabidopsis*. Mol Plant. 2010; 3:670–677.
41. Wu X, Oh MH, Schwarz EM, Larue CT, Sivaguru M, Imai BS, et al. Lysine acetylation is a widespread protein modification for diverse proteins in *Arabidopsis*. Plant Physiol. 2011; 155:1769–
42. Nallamilli BR, Edelmann MJ, Zhong X, Tan F, Mujahid H, Zhang J, Nanduri B, Peng Z. Global analysis of lysine acetylation suggests the involvement of protein acetylation in diverse biological processes in rice (*Oryza sativa*). PLoS One. 2014; 9:e89283.
43. Xue C, Liu S, Chen C, Zhu J, Yang X, Zhou Y, et al. Global proteome analysis links lysine acetylation to diverse functions in *Oryza sativa*. Proteomics. 2018; 18: 1700036.
44. Ma X, Zhang B, Liu C, Tong B, Guan T, Xia D. Expression of a populus histone deacetylase gene 84KHDA903 in tobacco enhances drought tolerance. Plant Sci. 2017; 265:1–
45. Yu X, Gao Q, Chen G, Guo JE, Guo X, Tang B, Hu Z. SlHDA5, a tomato histone deacetylase gene, is involved in responding to salt, drought, and ABA. Plant Mol Biol Rep. 2018; 36:36–

46. Song J, Henry HA, Tian L. Brachypodium histone deacetylase BdHD1 positively regulates ABA and drought stress responses. Plant Sci. 2019; 283:355–
47. Zhang JB, He SP, Luo JW, Wang XP, Li DD, Li XB. A histone deacetylase, GhHDT4D, is positively involved in cotton response to drought stress. Plant Mol Biol. 2020; 104: 67–
48. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 2013; 30:2725–2729.
49. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, HeY, Xia R. TBtools-an integrative toolkit developed for interactive analyses of big biological data. Mol Plant. 2020; 13:1194–1202.
50. Garrido J, Aguilar M, Prieto P. Identification and validation of reference genes for RT-qPCR normalization in wheat meiosis. Sci Rep. 2020; 10:1-12.

Figures

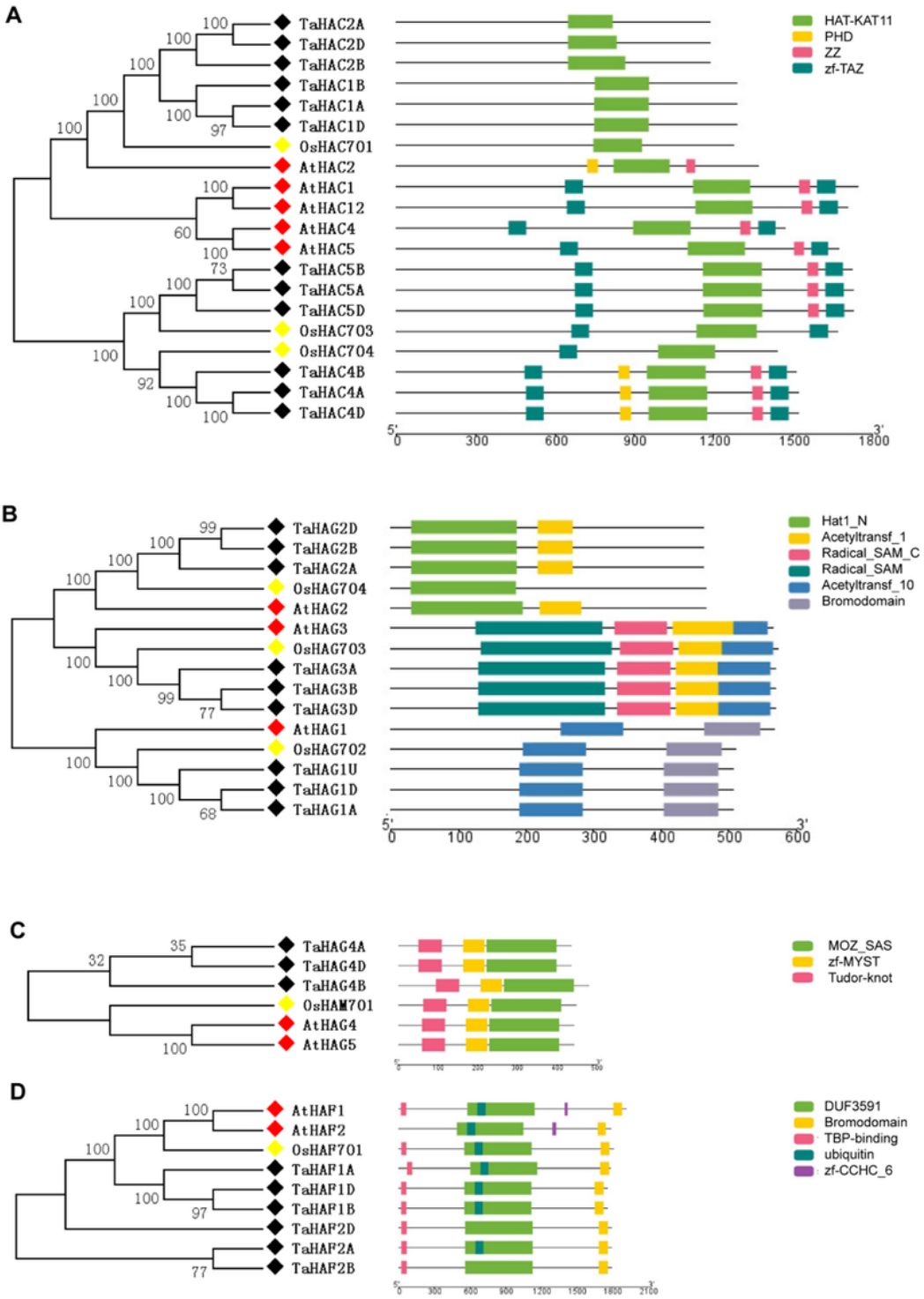

Figure 1

Molecular phylogenetic analysis and domain structures of HATs in Triticum aestivum, Arabidopsis thaliana and Oryza sativa. (A) CBP subfamily. (B) GNAT subfamily. (C) MYST subfamily and (D) TAFII250 subfamily. The HATs from Triticum aestivum, Oryza sativa, and Arabidopsis thaliana are marked with black, yellow, and red diamonds, respectively.

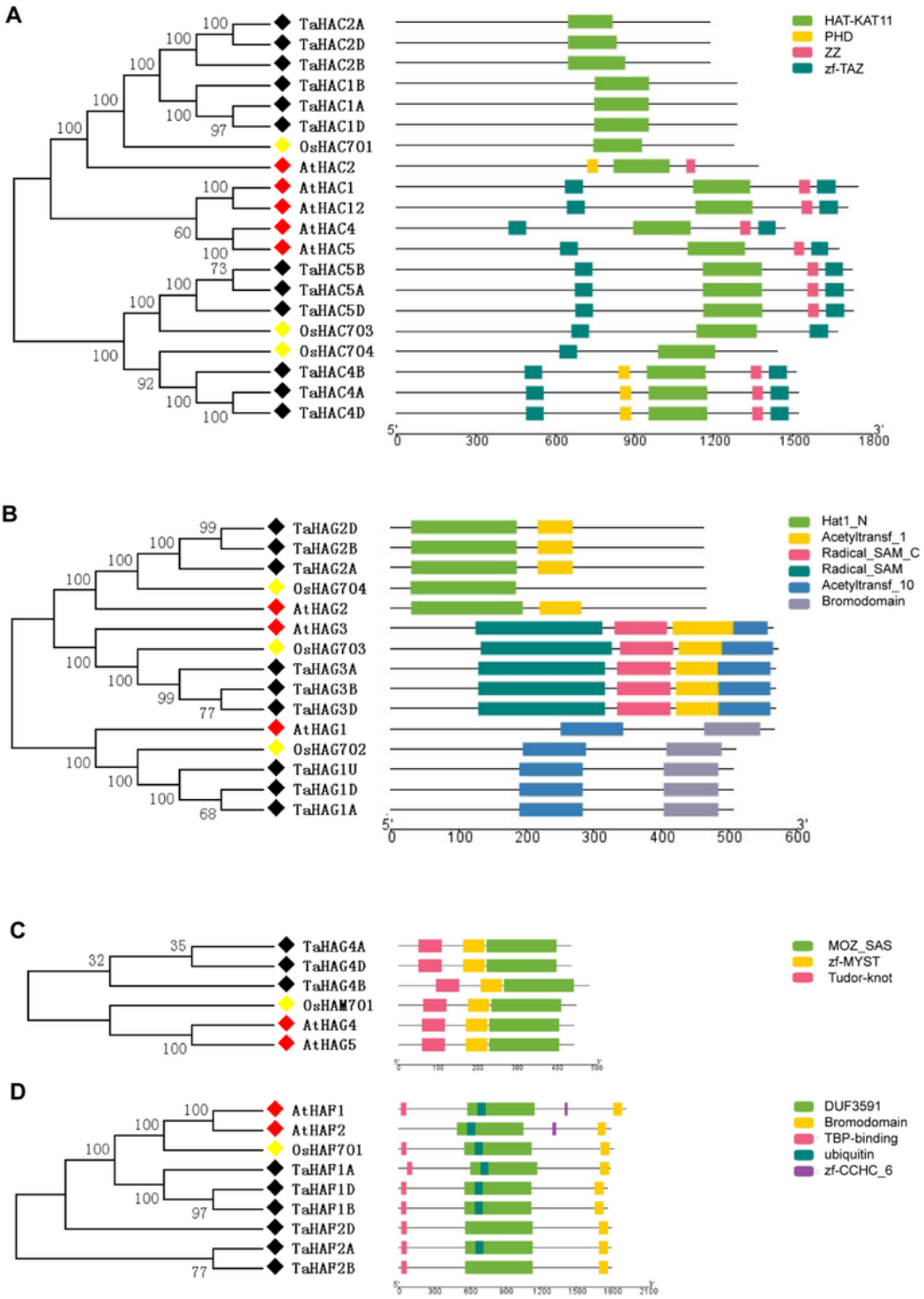

Figure 1

Molecular phylogenetic analysis and domain structures of HATs in Triticum aestivum, Arabidopsis thaliana and Oryza sativa. (A) CBP subfamily. (B) GNAT subfamily. (C) MYST subfamily and (D) TAFII250 subfamily. The HATs from Triticum aestivum, Oryza sativa, and Arabidopsis thaliana are marked with black, yellow, and red diamonds, respectively.

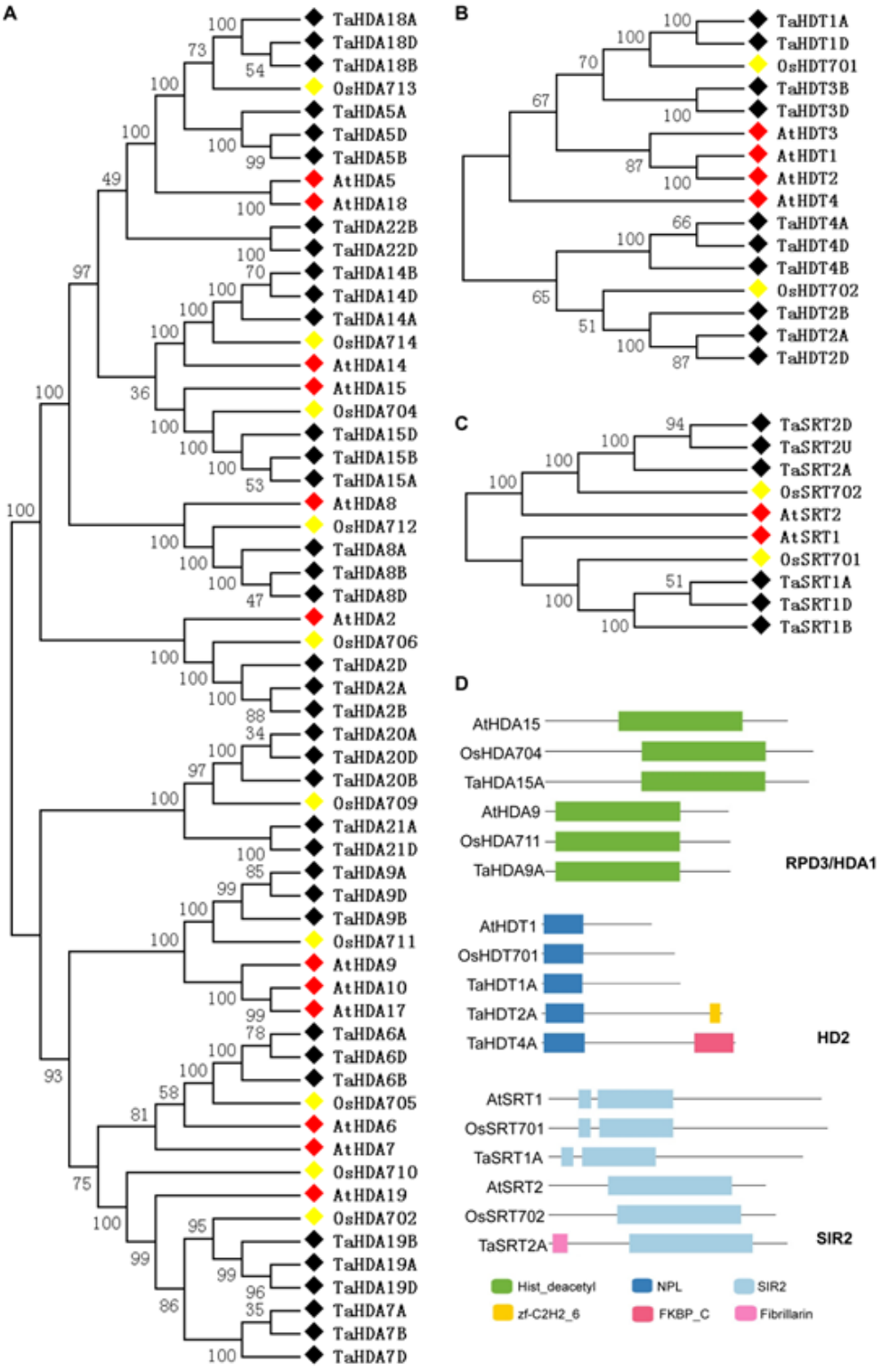

Figure 2

Molecular phylogenetic analysis and domain structures of HDACs in Triticum aestivum, Arabidopsis thaliana, and Oryza sativa. (A) RPD3/HDA1 subfamily. (B) HD2 subfamily. (C) SIR2 subfamily and (D) Typical domains of HDAC genes in each subfamily. The HDACs from Triticum aestivum, Oryza sativa, and Arabidopsis thaliana are marked with black, yellow, and red diamonds, respectively.

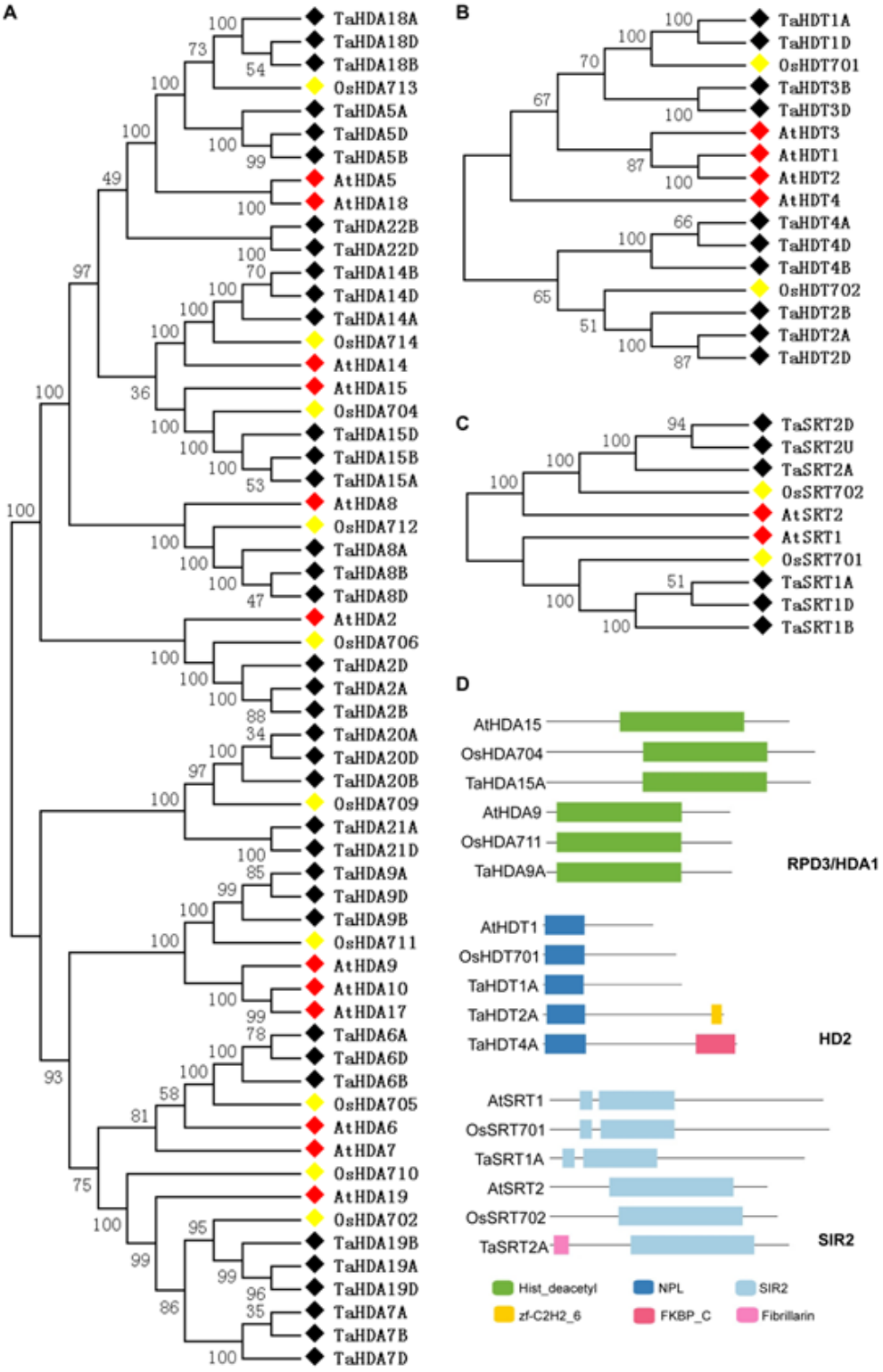

Figure 2

Molecular phylogenetic analysis and domain structures of HDACs in Triticum aestivum, Arabidopsis thaliana, and Oryza sativa. (A) RPD3/HDA1 subfamily. (B) HD2 subfamily. (C) SIR2 subfamily and (D) Typical domains of HDAC genes in each subfamily. The HDACs from Triticum aestivum, Oryza sativa, and Arabidopsis thaliana are marked with black, yellow, and red diamonds, respectively.

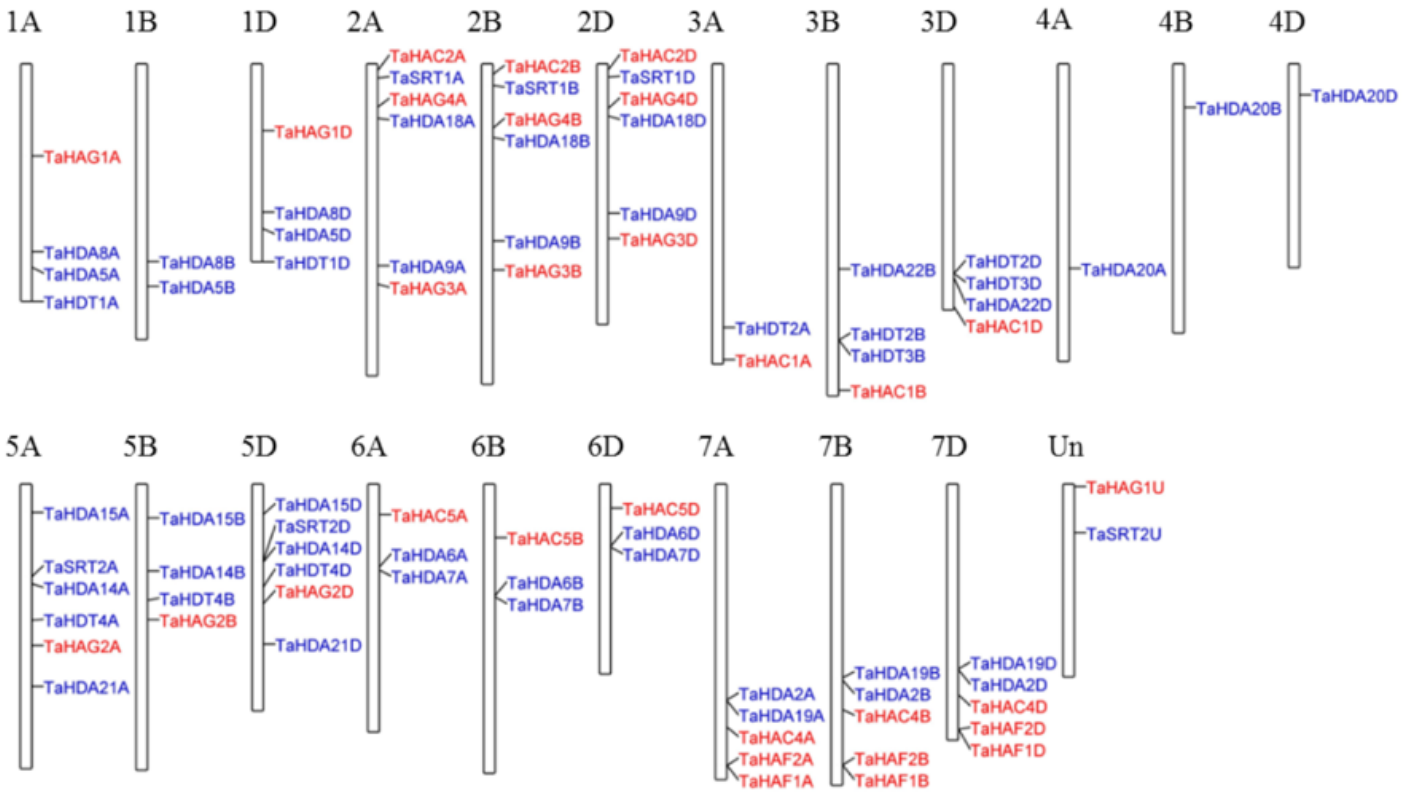

Figure 3

Chromosomes locations of TaHATs and TaHDACs genes on the wheat genome. TaHATs and TaHDACs are marked with red and blue, respectively.

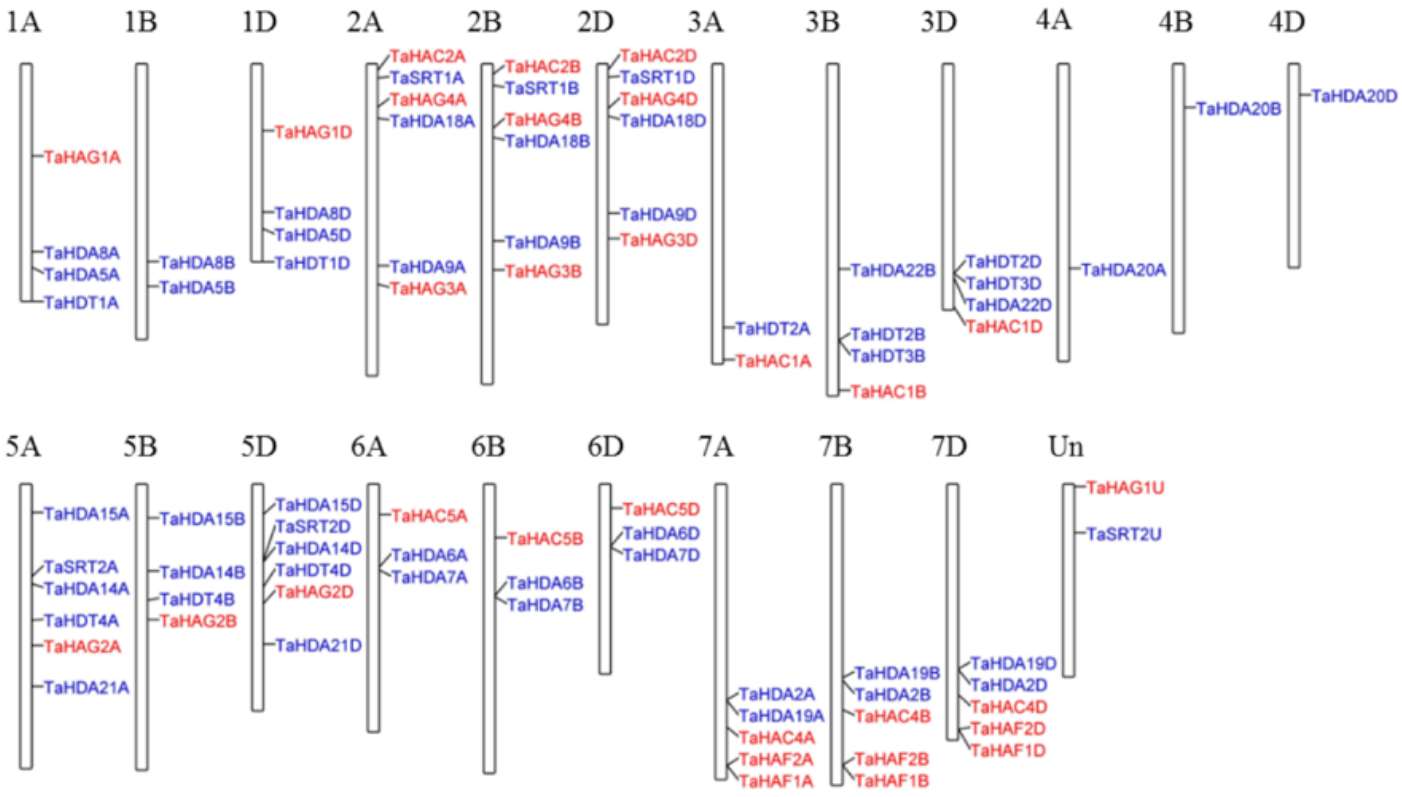

Figure 3

Chromosomes locations of TaHATs and TaHDACs genes on the wheat genome. TaHATs and TaHDACs are marked with red and blue, respectively.

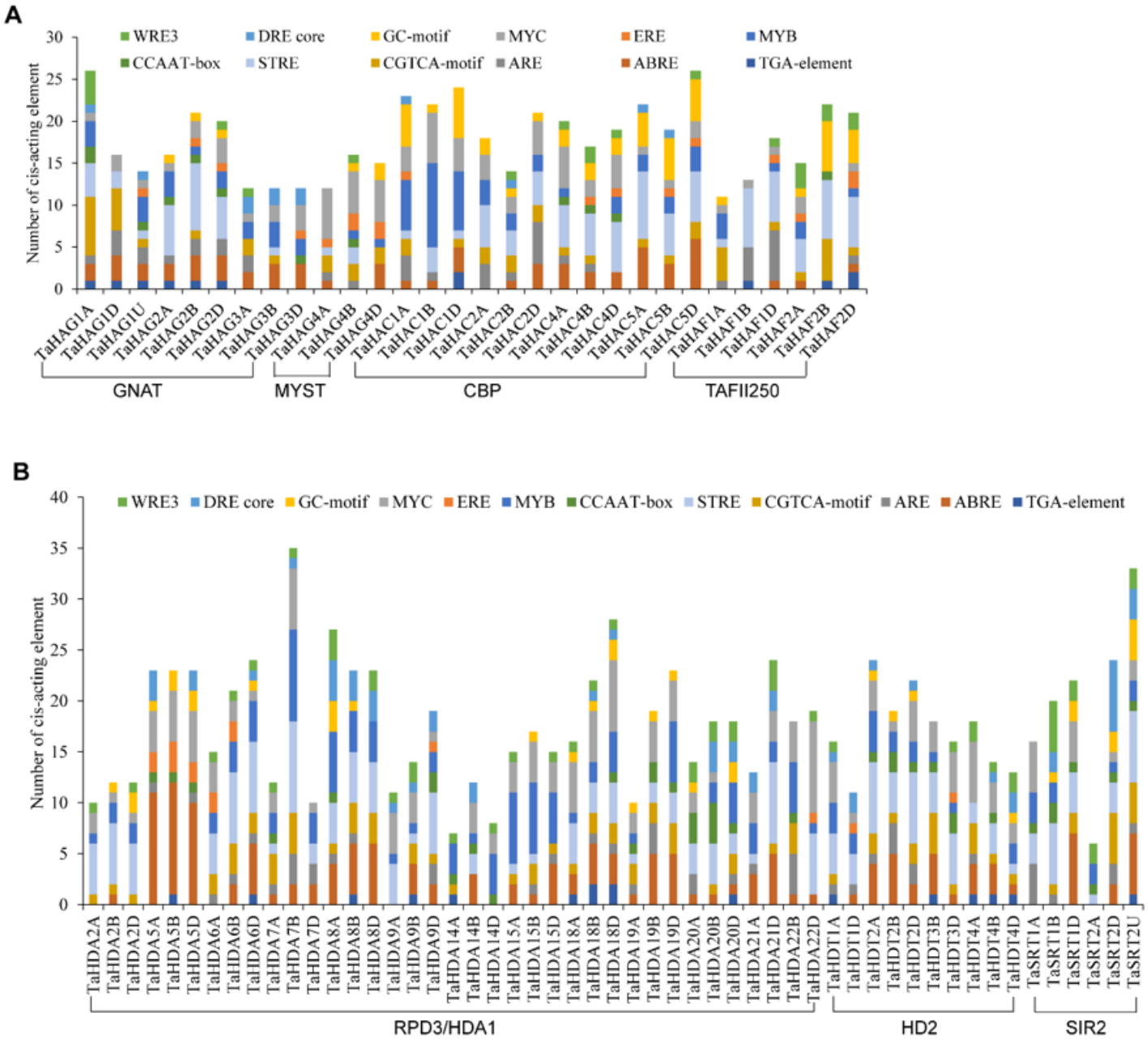

Figure 4

The number of putative cis-regulatory elements in the HATs (A) and HDACs (B) promoters.

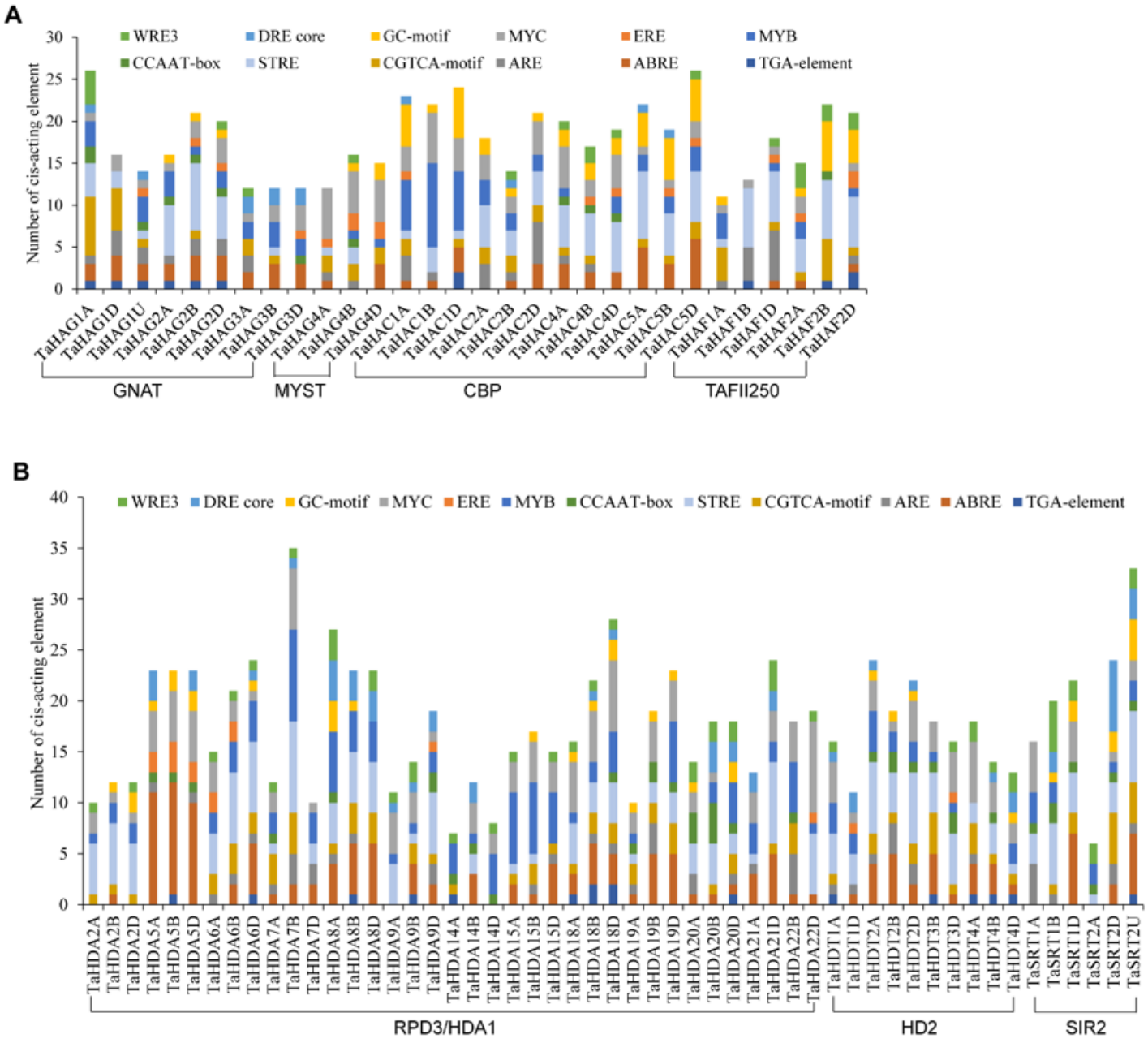

Figure 4

The number of putative cis-regulatory elements in the HATs (A) and HDACs (B) promoters.

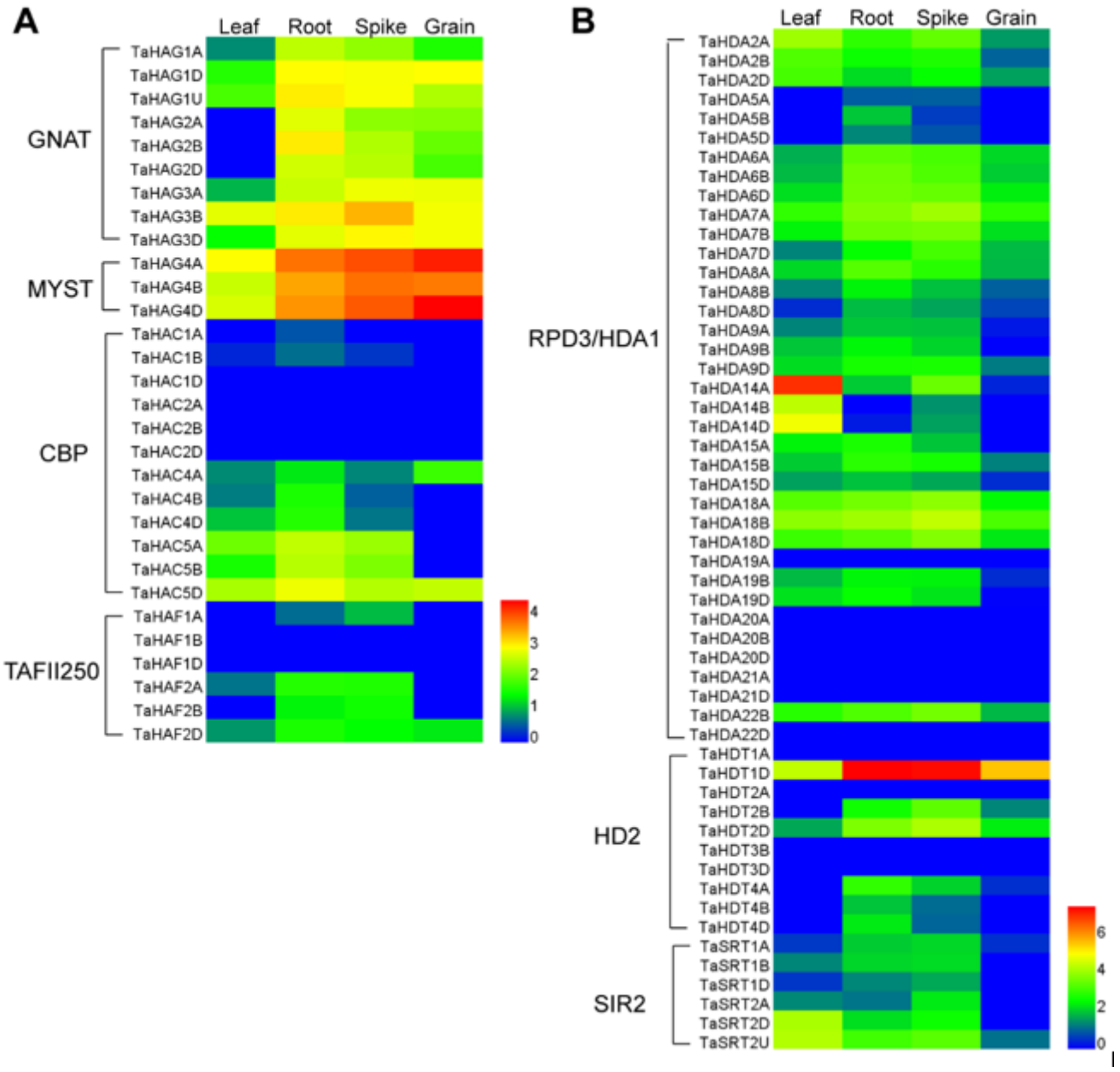

Figure 5

Gene expression pattern of wheat HATs (A) and HDACs (B) in different tissues. The expression levels (TPM values) of TaHATs and TaHDACs in the leaf, root, spike, and grain were downloaded from expVIP.

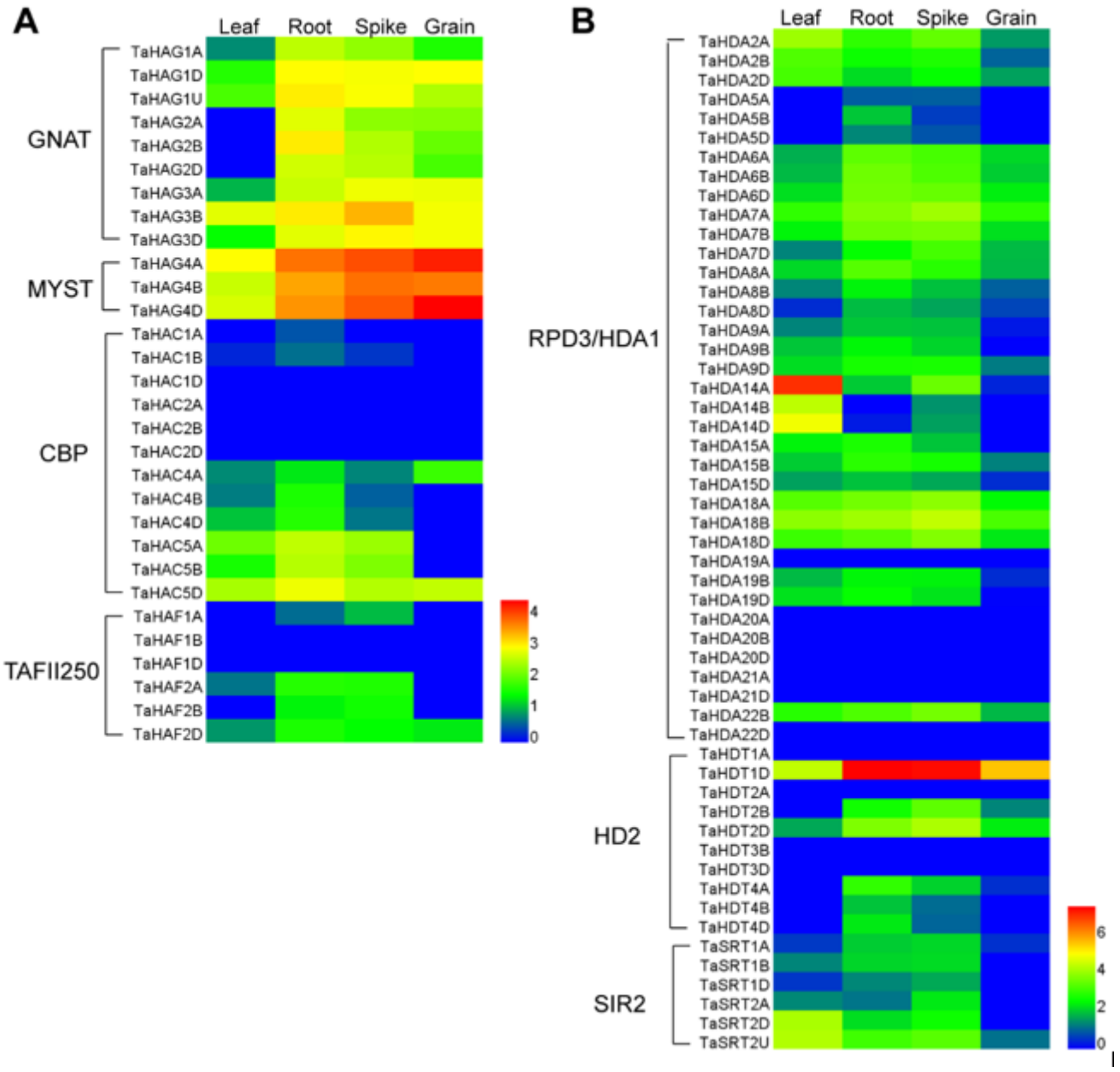

Figure 5

Gene expression pattern of wheat HATs (A) and HDACs (B) in different tissues. The expression levels (TPM values) of TaHATs and TaHDACs in the leaf, root, spike, and grain were downloaded from expVIP.

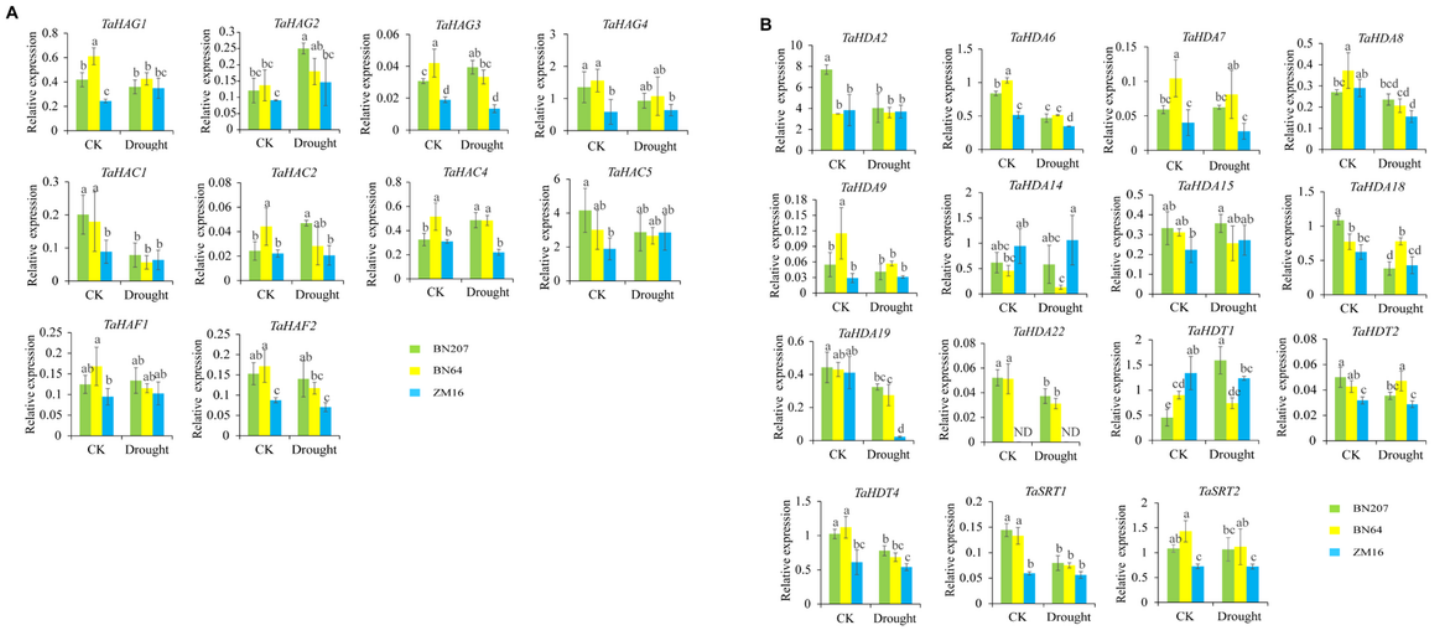

Figure 6

qRT-PCR analysis of the expression of TaHATs and TaHDACs gene under drought stress. At the booting stage of wheat, control soil water content to 55% as drought treatment. After 10 days later, flag leaves were collected for the gene expression analysis.

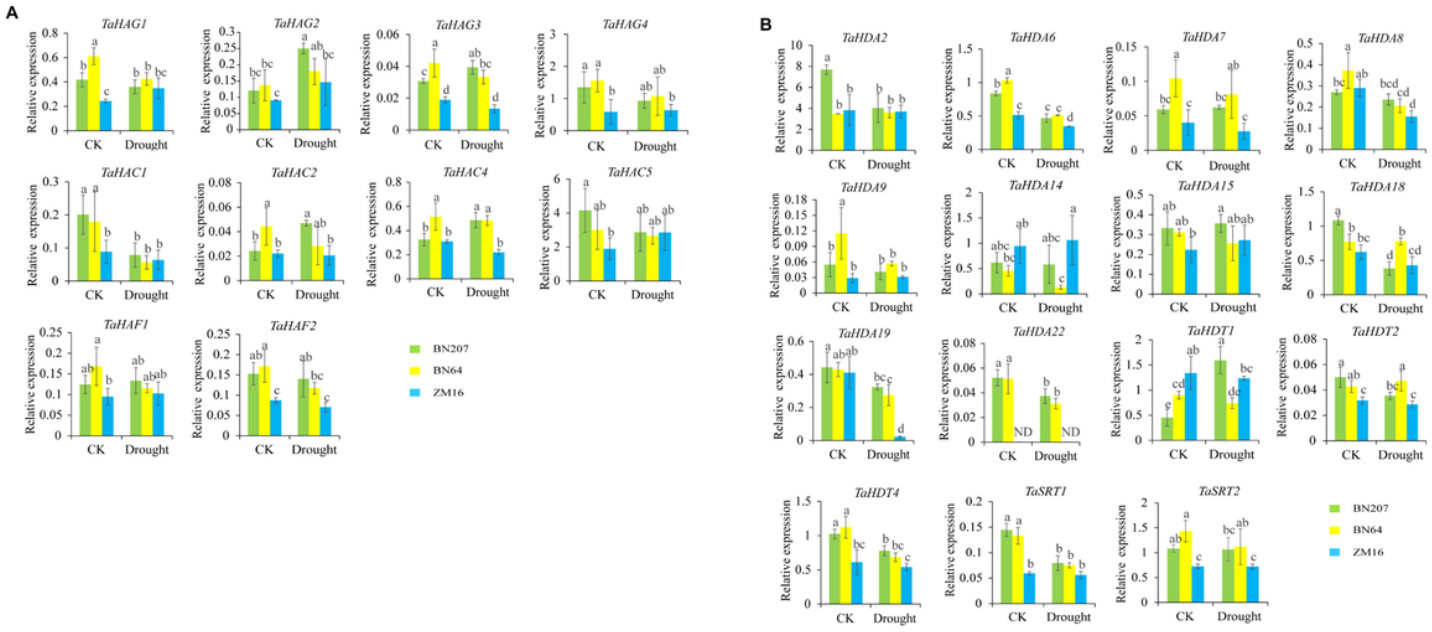

Figure 6

qRT-PCR analysis of the expression of TaHATs and TaHDACs gene under drought stress. At the booting stage of wheat, control soil water content to 55% as drought treatment. After 10 days later, flag leaves were collected for the gene expression analysis.

Supplementary Files

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

- FigureS1.docx
- FigureS1.docx
- TableS1.xlsx
- TableS1.xlsx
- TableS2.xlsx
- TableS2.xlsx
- TableS3.xlsx
- TableS3.xlsx
- TableS4.docx
- TableS4.docx
- TableS5.xlsx
- TableS5.xlsx