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

Hua Li1 · Huajie Liu1 · Xinxin Pei1 · Hongyu Chen1 · Xiao Li1 · Jiarui Wang2 · Chenyang Wang2

Received: 1 February 2021 / Accepted: 28 March 2021 / Published online: 13 April 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

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, 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, etc.). 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.

Keywords Histone acetyltransferases · Histone deacetylases · Gene expression · Drought stress · Wheat

Abbreviations

HATs Histone acetyltransferases
HDACs Histone deacetylases
GNAT GCN5-related N-terminal acetyltransferase
p300/CBP CREB-binding protein
MYST MOZ, Ybf2/Sas3, Sas2, and Tip60
TAFII250 TATA-binding protein-associated factors
RPD3/HDA1 Reduced Potassium Dependency 3/Histone DeAcetylase 1
SIR2 Silent Information Regulator 2
HD2 Histone Deacetylase 2
pI The theoretical isoelectric point
ABRE Abscisic acid-responsive element
STRE Stress-responsive element
ARE Essential for anaerobic induction
CCGTC-box Meristem-specific activation
G-Box Light responsiveness
MYB MYB-related binding sites
TGA-element Auxin-responsive element
PSII Photosystem II
LHCb Light-harvesting chlorophyll a/b-binding proteins

Handling Editor: Rhonda Peavy.

Hua Li and Huajie Liu have contributed equally to this study.

Hua Li lihua@henau.edu.cn
Chenyang Wang xmzxwang@163.com
Jiarui Wang xmzxwang@163.com

1 College of Life Science, Henan Agricultural University, Zhengzhou, China
2 College of Agronomy/State Key Laboratory of Wheat and Maize Crop Science, Henan Agricultural University, Zhengzhou, China

Introduction

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 (Kouzarides 2007). Among these, acetylation is one kind of well-characterized histone post-translational modifications. 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 (Hassig and Schreiber 1997). Histone acetylation can weaken the interaction between histones and negatively charged DNA, change the surface structure of nucleosomes, and stretch chromatin (Hollender and Liu 2008; Allis and Jenuwein 2002). 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) (Ekwall 2005). 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 (Wang et al. 2014; Zhao et al. 2019; Jiang et al. 2020; Ueda and Seki 2020). For example, Arabidopsis GCN5 mutants lose their apical advantage and show abnormal anther development (Cohen et al. 2009). In addition, GCN5 mutations or down-regulation increase plant sensitivity to salt, drought, heat, and disease stresses (Hu et al. 2015; Li et al. 2019; Zheng et al. 2019). The MYST family members HAM1 and HAM2 are involved in the regulation of flowering time, and their inhibition leads to early flowering in Arabidopsis (Xiao et al. 2013). The histone deacetylase HDA9 prevents seedling traits and germination (van Zanten et al. 2014) and negatively regulates salt and drought stress responsiveness in Arabidopsis (Zheng et al. 2016). HDA19 is involved in drought, heat, and salt responses (Chen and Wu 2010; Mehdi et al. 2016; Ueda et al. 2017; 2018). The overexpression of HDA705 in rice decreases ABA and salt stress resistance during seed germination and enhances osmotic stress resistance during the seedling stage (Zhao et al. 2016).

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 (Pandey et al. 2002), rice (Oryza sativa) (Liu et al. 2012), Brachypodium distachyon (Tan et al. 2019), litchi (Litchi chinensis Sonn) (Peng et al. 2017), Vitis vinifera (Aquea et al. 2010), Citrus sinensis (Xu et al. 2015), and cotton (Gossypium raimondii/arboreum) (Imran et al. 2019), and HDACs have been identified in Arabidopsis (Hollender and Liu 2008), rice (Hu et al. 2009), soybean (Liew et al. 2013), Vitis vinifera (Aquea et al. 2010), litchi (Peng et al. 2017), tomato (Zhao et al. 2015), and cotton (Imran et al. 2020). 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.

Materials and Methods

Identification of HATs and HDACs

Wheat HAT and HDAC proteins were identified following previously described methods in studies of Arabidopsis (Pandey et al. 2002) and rice (Liu et al. 2012). 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) (Gasteiger et al. 2005). The subcellular localization of each HAT and HDAC protein was predicted using WoLF PSORT (https://www.genscript.com/wolf-psort.html) (Horton et al. 2007).
Phylogenetic and Domain Analysis

To investigate evolutionary relationships, a multiple sequence alignment was generated for HAT and HDAC 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 (Tamura et al. 2013). Domain annotation was performed using Pfam, and TBtools was used for visualization (Chen et al. 2020).

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 (Chen et al. 2020).

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

BN207 (Bainong NO.207) is the main wheat cultivar in the Huang-Huai region of China, and its drought resistance is higher than that of its parents BN64 (Bainong NO.64) and ZM16 (Zhoumai NO.16); therefore, BN207 and its parents BN64 and ZM16 plants were used for gene expression analyses under drought stress. Seeds were sown into 27 x 26 cm pots filled with 12 kg loam soil, which was collected from the top 20 cm soil layer at the experimental site, and then the pots were buried in the field, keeping the soil surface inside the pot at the same level as that of the field. The plants were thinned to 12 seedlings per pot at the three-leaf stage. 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. The soil moisture was monitored daily with a soil moisture analyzer (RIME-PICO32 TDR, Germany), and water was added in time to maintain the required soil moisture. Flag leaves were collected for a gene expression analysis after 10 days of drought treatment. 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/) (Ramírez-González et al. 2018). A heatmap was generated using HemI (http://hem.biocuckoo.org/down.php).

Quantitative Real-Time PCR Analysis

Total RNA was isolated using the Omniplant RNA Kit (CWBio, Beijing, China) following the manufacturer’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™ Universal SYBR qPCR Master Mix (Vazyme) and the ABI StepOne-Plus Real-Time PCR System. The expression levels were calculated using the $2^{-\Delta \Delta CT}$ method. TaACT-1 was used as a housekeeping gene for normalization (Garrido et al. 2020). 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.

Results

Identification of HAT and HDAC 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 isoelectric 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
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         |
|               | TaHAF1A   | TraesCS7A02G515000.1  | 1782            | chlo         |
|               | TaHAF1B   | TraesCS7B02G431700.2  | 1762            | nucl         |
|               | 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         |
nucleus as well as in the chloroplasts, cytoplasm, mitochondria, cytoskeleton, etc. (Table S1).

**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

| Subfamily | Gene name | Gene ID | Protein length | Localization |
|-----------|-----------|---------|----------------|--------------|
| 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           |
| 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           |
| TaHDA19C  | 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           |
| TaHDA21B  | TraesCS5D02G302400.1 | 495     | cyto           |
| TaHDA22B  | TraesCS3B02G318000.1 | 380     | cysk           |
| TaHDA22C  | TraesCS3D02G422300.1 | 327     | cyto           |
| TaHDA22D  | TraesCS3D02G445700.4 | 309     | nucl           |
| TaHDT1A   | TraesCS1A02G445400.2 | 311     | nucl           |
| TaHDT2A   | TraesCS3A02G415200.1 | 403     | nucl           |
| TaHDT2B   | TraesCS3B02G450300.1 | 383     | nucl           |
| TaHDT2D   | TraesCS3D02G410300.2 | 364     | nucl           |
| TaHDT3A   | TraesCS3B02G450400.1 | 378     | nucl           |
| TaHDT3B   | TraesCS3B02G410400.1 | 432     | nucl           |
| TaHDT4A   | TraesCS3A02G158900.1 | 432     | nucl           |
| TaHDT4B   | TraesCS5B02G156700.1 | 433     | nucl           |
| TaHDT4C   | TraesCS5D02G164000.1 | 433     | nucl           |
| TaSRT1A   | TraesCS2A02G077800.1 | 440     | cyto, nucl     |
| TaSRT1B   | TraesCS2B02G092700.1 | 465     | cyto, nucl     |
| TaSRT1D   | TraesCS2D02G075800.1 | 678     | nucl, E.R      |
| TaSRT2A   | TraesCS5A02G114700.3 | 414     | cyto           |
| TaSRT2D   | TraesCS5D02G124700.1 | 396     | chlo           |
| TaSRT2U   | TraesCSU02G136000.1  | 396     | chlo           |
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 (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 (Figure S2; 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 TaHDASD/B/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 TaHDAC subfamilies.

**Expression Analysis of TaHATs and TaHDACs in Different Tissues**

The RNA-seq data for different tissues were obtained from expVIP. All TaHATs and TaHDACs were differentially expressed in the leaf, root, spike, and grain (Fig. 4). 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. 4a). 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, TaHDT1A, TaHDT2A, and TaHDT3B/D were almost undetectable in the four tissues (Fig. 4b). 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. 4b). 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

Fig. 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
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. 5). This indicated that these genes may not be related to the regulation of drought stress. All other genes were up-regulated or down-regulated at least one variety under drought stress. It is worth noting that TaHAG2, TaHAG3, and TaHAC2 in the TaHAT family (Fig. 5A) and TaHDA2, TaHDA18, TaHDT1, and TaHDT2 in the TaHDAC family (Fig. 5B) showed a significant response to drought stress only in BN207; in particular, TaHAG2, TaHAG3, and TaHAC2 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 TaHAT and TaHDAC Proteins

HATs and HDACs are considered to mediate plant growth and development and the response to environmental stresses (Zhao et al. 2019). 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 (Pandey et al. 2002; Hu et al. 2009; Liu et al. 2012). 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 (Bienz 2006; Gamsjaeger et al. 2007). Bromodomains bind to acetylated lysine residues (Marmorstein and Berger 2001; Servet et al. 2010). 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). Although previous studies have shown that homologs of AtHAG5 and AtHAC12 are lacking in Brachypodium distachyon (Tan et al. 2019), the homologous genes of AtHAC2 and AtHAF2 (also missing from B. distachyon) were detected in wheat (Fig. 1) (Tan et al. 2019). 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) (Hu et al. 2009). 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 (Wu et al. 2011; Nallamilli et al. 2014; Xue et al. 2018). 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 (Xue et al. 2018). Similarly, a large number of acetylated proteins are related to photosynthesis in Arabidopsis (Wu et al. 2011), 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 (Li et al. 2019). 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 (Zheng et al. 2016). The overexpression of the Populus histone deacetylase gene 84KHD1A results in reduced tolerance to salt, drought, and ABA (Yu et al. 2018). The overexpression of the Brachypodium histone deacetylase BdHD1 results in hypersensitivity to ABA and enhanced drought resistance (Song et al. 2019). The overexpression of the cotton histone deacetylase GhHDT4D in Arabidopsis increases plant tolerance to drought (Zhang et al. 2020). 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. 5). 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 (Figure S2). 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 *TaHDACs* 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.
Fig. 5 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. Values are the means ± standard deviation (SD) (n = 3). Bars denoted by the different letters indicate statistical significance at p < 0.05

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00344-021-10364-9.

Acknowledgments This research was supported by the National Key Research and Development Program of China (Grant No. 2017YFD0300408) and Natural Science Foundation of Henan Province (Grant No. 212300410352).

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 revise the manuscript. All the authors read and approved the final manuscript.

Declarations

Conflict of interest The authors declare that they have no competing financial interests.

References

Allis CD, Jenuwein T (2016) The molecular hallmarks of epigenetic control. Nat Rev Genet 17:487–500

Aquea F, Timmermann T, Arce-Johnson P (2010) Analysis of histone acetyltransferase and deacetylase families of Vitis vinifera. Plant Physiol Bioch 48:194–199

Bienz M (2006) The PHD finger, a nuclear protein-interaction domain. Trends Biochem Sci 31:35–40

Chen LT, Wu K (2010) Role of histone deacetylases HDA6 and HDA19 in ABA and abiotic stress response. Plant Signal Behav 5:1318–1320

Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, HeY XR (2020) TBtools—an integrative toolkit developed for interactive analyses of big biological data. Mol Plant 13:1194–1202

Cohen R, Schocken J, Kaldis A, Vilachanasios KE, Hark AT, McCain ER (2009) The histone acetyltransferase GCN5 affects the inflorescence meristem and stamen development in Arabidopsis. Planta 230:1207–1221

Ekwall K (2005) Genome-wide analysis of HDAC function. Trends Genet 21:608–615

Gamsjaeger R, Liew CK, Loughlin FE, Crossley M, Mackay JP (2007) Sticky fingers: zinc-fingers as protein-recognition motifs. Trends Biochem Sci 32:63–70

Garrido J, Aguilar S, Prieto P (2020) Identification and validation of reference genes for RT-qPCR normalization in wheat meiosis. Sci Rep 10:1–12

Gasteiger E, Hoogland C, Gattiker A, Wilkins M, Appel R, Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed) The proteomics protocols handbook. Humana Press, Totowa, pp 571–607

Hassig CA, Schreiber SL (1997) Nuclear histone acetylases and deacetylases and transcriptional regulation: HATs off to HDACs. Curr Opin Chem Biol 1:300–308

Hollender C, Liu Z (2008) Histone deacetylases genes in Arabidopsis development. J Integr Plant Biol 50:875–885

Horton P, Park K, Obayashi T, Fujita N, Harada H, Adams-Collier C, Nakai K (2007) WoLF PSORT: protein localization predictor. Nucleic Acids Res 35(suppl_2):585–587

Hu Y, Qin F, Huang L, Sun Q, Li C, Zhao Y, Zhou DX (2009) Rice histone deacetylases genes display specific expression patterns and developmental functions. Biochem Bioph Res Co 388:266–271

Hu Z, Song N, Zheng M, Liu X, Liu Z, Xing J et al (2015) Histone acetyltransferase GCN 5 is essential for heat stress-responsive gene activation and thermotolerance in Arabidopsis. Plant J 84:1178–1191

Imran M, Shaﬁq S, Farooq MA, Naeem MK, Widemann E, Bakhsh A et al (2019) 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 20:5311

Imran M, Shaﬁq S, Naeem MK, Widemann E, Munir MZ, Jensen KB, Wang R (2020) Histone Deacetylase (HDAC) gene family in Allotetraploid cotton and its diploid progenitors silico identiﬁcation molecular characterization, and gene expression analysis under multiple abiotic stresses DNA damage and phytohormone treatments. Int J Mol Sci 21(1):321

Jiang J, Ding AB, Liu F, Zhong X (2020) Linking signaling pathways to histone acetylation dynamics in plants. J Exp Bot 71:5179–5190

Kouzarides T (2007) Chromatin modiﬁcations and their function. Cell 128:693–705

Li S, Lin Y, Wang P, Zhang B, Li M, Chen S et al (2019) The AREB1 transcription factor inﬂuences histone acetylation to regulate drought responses and tolerance in Populus trichocarpa. Plant Cell 31:663–686

Liew LC, Singh MB, Bhalla PL (2013) An RNA-seq transcriptome analysis of histone modiﬁers and RNA silencing genes in soybean during ﬂoral initiation process. PLoS ONE 8:e77502

Liu X, Luo M, Zhang W, Zhao J, Zhang J, Wu K et al (2012) Histone acetyltransferases in rice (Oryza sativa L.): phylogenetic analysis subcellular localization and expression. BMC Plant Biol 12(1):145

Ma X, Zhang B, Liu C, Tong B, Guan T, Xia D (2017) Expression of a populus histone deacetylase gene 84KHD903 in tobacco enhances drought tolerance. Plant Sci 265:1–11

Marmorstein R, Berger SL (2001) Structure and function of bromodomains in chromatin-regulating complexes. Gen 272:1–9

Mehdii S, Derkacheva M, Ramström M, Krallemann L, Bergquist J, Hennig L (2016) The WD40 domain protein MSII functions in a histone deacetylase complex to ﬁne-tune abscisic acid signaling. Plant Cell 28:42–54

Nallamilli BR, Edelmann MJ, Zhong X, Tan F, Mujahid H, Zhang J, Nanduri B, Peng Z (2014) Global analysis of lysine acetylation suggests the involvement of protein acetylation in diverse biological processes in rice (Oryza sativa). PLoS ONE 9:e89283

Onufriev AV, Schiessel H (2019) The nucleosome: from structure to function through physics. Curr Opin Struc Biol 56:119–130

Pandey R, MuÈller A, Napoli CA, Selinger DA, Pikaard CS, Richards JE et al (2002) Analysis of histone acetyltransferase and histone deacetylase families of Vitis vinifera. Curr Opin Genet Dev 12(1):145

Pandey R, MuÈller A, Napoli CA, Selinger DA, Pikaard CS, Richards JE et al (2002) Analysis of histone acetyltransferase and histone deacetylase families of Vitis vinifera. Curr Opin Genet Dev 12(1):145
Ramírez-González RH, Borrill P, Lang D, Harrington SA, Brinton J, Venturini L et al (2018) The transcriptional landscape of polyploid wheat. Science 361:662

Servet C, Silva N, Zhou D (2010) Histone acetyltransferase AtGCN5/HAG1 is a versatile regulator of developmental and inducible gene expression in Arabidopsis. Mol Plant 3:670–677

Shahbazian M, Grunstein M (2007) Functions of site-specific histone acetylation and deacetylation. Annu Rev Biochem 76:75–100

Song J, Henry HA, Tian L (2019) Brachypodium histone deacetylase BdHD1 positively regulates ABA and drought stress responses. Plant Sci 283:355–365

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729

Tan S, Gao L, Li T, Chen L (2019) Phyllogenetic and expression analysis of histone acetyltransferases in Brachypodium distachyon. Genomics 111:1966–1976

Ueda M, Seki M (2020) Histone modifications form epigenetic regulatory networks to regulate abiotic stress response. Plant Physiol 182:15–26

Wang Z, Cao H, Chen F, Liu Y (2014) The roles of histone acetylation in seed performance and plant development. Plant Physiol Bioch 84:125–133

Wu X, Oh MH, Schwarz EM, Larue CT, Sivaguru M, Imai BS et al (2011) Lysine acetylation is a widespread protein modification for diverse proteins in Arabidopsis. Plant Physiol 155:1769–1778

Zheng Y, Ding Y, Sun X, Xie S, Wang D, Liu X, Su L, Lei Pan W, Zhou DX (2016) Histone deacetylase HDA9 negatively regulates salt and drought stress responsiveness in Arabidopsis. J Exp Bot 67:1703–1713

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.