Genome-wide characterization and expression profiling analysis of the xyloglucan endotransglycosylase/hydrolase gene family in Brachypodium distachyon

Hongyan Shen
Shandong Agricultural University
Qiuping Tan
Shandong Agricultural University
Wei Xiao
Shandong Agricultural University
Wenpeng Deng
Shandong Agricultural University
Xiaoyan Yu (yxyxst20040214@163.com)
Shandong Agricultural University
Hongyu Wu (15275380863@163.com)
Shandong Agricultural University

Research article

Keywords: Brachypodium distachyon, Xyloglucan endotransglucosylase/hydrolase, Genome wide analysis, Gene expression

Posted Date: July 29th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-41833/v1

Genome-wide characterization and expression profiling analysis of the xyloglucan endotransglycosylase/hydrolase gene family in *Brachypodium distachyon*

Hongyan Shen[1#], Qiuping Tan[1#], Wei Xiao[1], Wenpeng Deng[1], Xiaoyan Yu[2]* and Hongyu Wu[2]*

[1]State Key Laboratory of Crop Biology, College of Horticulture Science and Engineering, Shandong Agricultural University, 61 Daizong Road, Tai'an 271018, China

[2]State Forestry Administration Key Laboratory of Silviculture in downstream areas of the Yellow River, College of Forestry, Shandong Agricultural University, Tai'an 271018, China

[#] Hongyan Shen and Qiuping Tan contributed equally to this work

***Corresponding author**

Hongyu Wu:

E-mail: 15275380863@163.com

Xiaoyan Yu:

E-mail: yxyxst20040214@163.com

Abstract

Background: Xyloglucan endotransglucosylase/hydrolases (XTHs) are a class of cell wall-associated enzymes involved in the construction and remodeling of cellulose/xyloglucan crosslinks. However, knowledge of this gene family in the model monocot *Brachypodium distachyon* is limited.

Results: A total of 29 *BdXTH* genes were identified from the reference genome, and these were further divided into three main groups (Group I/II, Group III, and the Ancestral Group) through comparative phylogenetic analysis. Gene structure and protein motif analysis indicate that closely clustered *BdXTH* genes are relatively conserved within each group. A highly conserved amino acid domain (DEIDFEFLG) responsible for catalytic activity was identified in all BdXTH proteins. We detected three pairs of segmentally duplicated *BdXTH* genes and five groups of tandemly duplicated *BdXTH* genes, which played vital roles in the expansion of the BdXTH gene family. *Cis*-elements related to hormones, growth, and abiotic stress responses were identified in the promoters of each *BdXTH* gene. Most *BdXTH* genes have distinct expression patterns in different tissues and growth stages. Furthermore, when roots were treated with two abiotic stresses (salinity and drought) and four plant hormones (IAA, auxin; GA3, gibberellin; ABA, abscisic acid and BR, brassinolide), the expression levels of many *BdXTH* genes changed significantly, suggesting possible roles in response to various environmental stimuli and plant hormones.

Conclusion: In this study, we performed genome-wide identification, characterization, and expression pattern analysis of the XTH gene family in *Brachypodium*, which provide valuable information for further elucidation of the biological functions of *BdXTH* genes in the model grass *B. distachyon*.

Key worlds: *Brachypodium distachyon*; Xyloglucan endotransglucosylase/hydrolase; Genome-wide analysis; Gene expression

Background

Xyloglucan endotransglucosylase/hydolases (XTHs), a subfamily of the glycoside hydrolase family GH16, are crucial enzymes that are involved in the regulation of cell wall extension, construction, and degradation in plants [1, 2]. XTH proteins have two significant catalytic activities, and can act either as endotransglucosylases (XET, EC 2.4.1.207) to elongate xyloglucan chains by cleaving the chains and rejoining the reducing ends to other xyloglucan molecules, or as endohydrolases (XEH, EC 3.2.1.151) that cleave xyloglucan chains by rejoining the xyloglucan reducing end to water molecules [3]. XTH proteins are predicted to present several common structural features: a putative signal peptide, a conserved ExDxE motif likely to be the catalytic site for both XET and XEH activities, a potential *N*-glycosylation site necessary for protein stability, and several cysteine residues that stabilize the C-terminal end [4].

The *XTH* gene family is widely distributed in both monocotyledonous and dicotyledonous plants, and the gene numbers vary within individual plant species, with 33 in *Arabidopsis* [5], 29 in rice (*Oryza sativa*) [6], 25 in tomato (*Solanum lycopersicum*) [7], 44 in *Medicago truncatula* [8], 26 in woodland strawberry (*Fragaria vesca*) [9], 56 in tobacco (*Nicotiana tabacum*) [10], 61 in soybean (*Glycine max*) [11], 24 in pineapple (*Ananas comosus*) [12], and 24 in barley (*Hordeum vulgare*) [13]. A phylogenetic analysis revealed that the XTH proteins cluster into three groups (I, II, and III) on the basis of sequence similarity in *A. thaliana* [4, 5]. In rice, the XTH proteins were found to cluster into two major groups named I/II and III, because the boundary between group I and II was no longer apparent[14]. Baumann et al. (2007) used ~130 full-length XTH protein sequences mainly from *Arabidopsis*, rice, black cottonwood (*Populus trichocarpa*), tomato, and hybrid aspen (*Populus tremula* X *Populus tremuloides*) to derive a phylogenetic tree from a structure-based sequence alignment using a maximum likelihood method. This study showed that group III can be further separated into two main clades (group III-A and group III-B), and a small outlying ancestral group that is close to the root [15]. Several studies in *Fragaria vesca* [9], barley [13], and *Ananas comosus* [12] also reported similar classifications of the XTH proteins. Furthermore, XTH proteins in groups I, II, and

III-B have been reported to have significant xyloglucan endotransglucosylase (XET) activity, while proteins in group III-A mainly show xyloglucan endohydrolase (XEH) activity [4, 16, 17].

On the one hand, *XTH* family genes are involved in many physiological responses. For example, *DkXTH1*, *DkXTH4*, and *DkXTH5* in persimmon shown higher expression levels and are associated with fruit firmness. However, the expression of *DkXTH2* and *DkXTH3* reach their maxima concomitant with pronounced fruit softening [18]. In addition, overexpression of *FvXTH9* and *FvXTH6* might promote strawberry fruit ripening by modifying cell wall components [19]. *XTH17* and *XTH24* in *Arabidopsis* are involved in polar cell elongation [20]. The natural variation of *PtoXET16A* in poplar can affect wood properties, and the expression of *PtoXET16A* in *Populus tomentosa* was highest in the root, followed by the phloem, cambium, and developing xylem, suggesting that *PtoXET16A* plays important roles in the development of vascular tissues [21]. On the other hand, *XTH* family genes also play important roles in the response to plant hormones and abiotic stresses [11, 22]. The expression of Arabidopsis *XTH17* was substantially reduced in the presence of aluminum (Al), and the *xth17* and *xth31* mutants were more Al resistant than the wild type [23]. Expression of *DkXTH6* in persimmon (*Diospyros kaki*) was found to be positively up-regulated during ethylene production, as well as by propylene and ABA treatments, although expression was down-regulated by GA3 and cold treatments. However, the mRNA levels of *DkXTH7* were the highest in GA3-treated fruits and cold-treated fruits [24]. In addition, overexpression of persimmon *DkXTH1* enhanced tolerance to salt, ABA, and drought stresses in transgenic *Arabidopsis* plants and delayed fruit softening in transgenic tomatoes [22]. In soybean, the expression levels of many *GmXTH* genes were significantly associated with ethylene and flooding stress, and overexpression of *Arabidopsis XTH31* in soybean can enhance plant flooding tolerance [11].

Brachypodium distachyon is a species of monocot that is used as a model system for genetic and physiological studies in grasses. Plants of *B. distachyon* are small in stature, have a short life cycle, a small genome, modest growth requirements, and

many available mutant resources [25-29]. Although *XTH* genes have been reported to play important roles during plant growth and development, little is known about the functions of the *BdXTH* genes in *Brachypodium*. The available genome sequence of *Brachypodium distachyon* and related studies of the *XTH* gene families from other species will enable the comprehensive characterization of *BdXTH* genes from *B. distachyon* [30, 31]. In this study, we identified 29 *BdXTH* genes in the *B. distachyon* genome based on a bioinformatic analysis. We then performed a comprehensive analysis of the *BdXTH* genes, including their evolutionary relationships, gene structures, duplication events, conserved motifs, and *cis*-regulatory elements. To provide useful information for further functional studies of the *BdXTH* genes in *B. distachyon*, the expression patterns of the *XTH* genes in response to stresses such as plant hormones, salinity, and drought were characterized by qRT-PCR analysis. This study provides novel insights into the functional characteristics of the *BdXTH* genes in *B. distachyon*, and also provides necessary resources for further research into the specific functions and regulatory mechanisms of *BdXTH* genes at different developmental stages.

Results

Identification of the *BdXTH* genes in *Brachypodium distachyon*

The availability of the *B. distachyon* genome makes it possible to identify the *XTH* family genes on a genome-wide level. A total of 29 candidate *BdXTH* genes that are predicted to encode proteins containing both the PF00722 and PF06955 domains were identified. These genes were renamed *BdXTH1* to *BdXTH29* on the basis of their chromosomal positions (Table 1). The lengths of the BdXTH proteins varied from 279 to 372 amino acids with an average length of 306 amino acids. Corresponding with protein length, the MWs ranged from 30.54 kDa to 40.94 kD. Subcellular localization analysis predicted that 27 BdXTH proteins are located in the cell wall, and the other two proteins, BdXTH8 and BdXTH23, are targeted to both the cell wall and the cytoplasm. Owing to the complex amino acid polarities, the PIs ranged from 4.67 to 8.83. Information relating to other parameters of the BdXTH proteins, such as

instability index (II), aliphatic index (AI), and grand average of hydropathicity (GRAVY) are also presented in Table 1.

160 Table 1

Gene Name	Gene ID	Chromosome	Start	End	Length (aa)	MW (kDa)	PI	II	AI	GRAVY	Subcellular Localization
BdXTH1	Bradi1g09690	Bd1	6944672	6946132	284	31.55	5.67	36.99	63.87	-0.27	Cell wall
BdXTH2	Bradi1g09700	Bd1	6956014	6957576	284	31.54	5.67	36.12	63.87	-0.26	Cell wall
BdXTH3	Bradi1g25847	Bd1	20978556	20980916	307	34.16	5.09	39.79	63.03	-0.34	Cell wall
BdXTH4	Bradi1g27867	Bd1	23074430	23076645	302	34.11	8.73	40.95	58.91	-0.46	Cell wall
BdXTH5	Bradi1g33810	Bd1	29475206	29476739	290	32.39	8.78	30.99	58.24	-0.42	Cell wall
BdXTH6	Bradi1g33817	Bd1	29478854	29481513	300	34.28	6.71	43.09	71.2	-0.36	Cell wall
BdXTH7	Bradi1g33827	Bd1	29485795	29487172	284	31.45	5.73	38.54	62.64	-0.34	Cell wall
BdXTH8	Bradi1g33840	Bd1	29494938	29496678	293	32.59	6.22	31.26	69.66	-0.28	Cell wall/Cytoplasm
BdXTH9	Bradi1g44777	Bd1	43043568	43045106	301	33.45	5.05	43.37	68.37	-0.33	Cell wall
BdXTH10	Bradi1g68590	Bd1	67356916	67359656	328	36.46	6.02	52.51	67.65	-0.39	Cell wall
BdXTH11	Bradi1g71937	Bd1	70046381	70048307	284	31.51	6.44	36.9	78.27	-0.12	Cell wall
BdXTH12	Bradi1g77990	Bd1	74331199	74332831	318	35.24	7.00	46.49	61.07	-0.41	Cell wall
BdXTH13	Bradi3g02700	Bd3	1633396	1635150	336	37.66	8.57	51.44	76.10	-0.22	Cell wall
BdXTH14	Bradi3g10290	Bd3	8496465	8498025	288	32.10	4.91	32.13	66.32	-0.33	Cell wall
BdXTH15	Bradi3g18590	Bd3	17168440	17171803	280	31.02	4.91	38.97	59.21	-0.38	Cell wall
BdXTH16	Bradi3g18600	Bd3	17175111	17176766	301	33.71	5.29	28.49	60.66	-0.52	Cell wall
BdXTH17	Bradi3g18607	Bd3	17180627	17182323	291	33.41	4.85	42.45	63.68	-0.51	Cell wall
BdXTH18	Bradi3g18690	Bd3	17286604	17292191	289	32.83	6.89	32.97	71.87	-0.41	Cell wall
BdXTH19	Bradi3g21337	Bd3	20473847	20476548	354	39.32	8.75	48.62	75.06	-0.25	Cell wall
BdXTH20	Bradi3g31767	Bd3	20473847	20476548	289	32.32	5.78	26.27	69.24	-0.26	Cell wall
BdXTH21	Bradi3g34227	Bd3	36405786	36409262	330	35.58	7.83	44.98	79.88	-0.18	Cell wall
BdXTH22	Bradi3g52307	Bd3	53017512	53019765	345	37.56	8.83	33.85	64.09	-0.23	Cell wall
BdXTH23	Bradi4g16990	Bd4	17884860	17887177	294	33.37	6.25	37.64	64.76	-0.49	Cell wall/Cytoplasm
BdXTH24	Bradi4g29707	Bd4	35169644	35172434	338	37.31	6.41	52.96	70.74	-0.29	Cell wall

BdXTH25	Bradi5g20718	Bd5	23655071	23657848	316	34.23	4.67	44.62	44.62	-0.19	Cell wall
BdXTH26	Bradi5g20726	Bd5	23658483	23663566	372	40.94	5.79	43.33	67.96	-0.25	Cell wall
BdXTH27	Bradi5g20734	Bd5	23661776	23664848	295	33.66	4.83	38.07	76.41	-0.29	Cell wall
BdXTH28	Bradi5g20742	Bd5	23668827	23670387	279	30.54	5.19	40.05	67.89	-0.26	Cell wall
BdXTH29	Bradi5g22907	Bd5	25246820	25248701	314	34.27	5.37	31.8	73.34	-0.25	Cell wall

Phylogenetic analysis of BdXTH proteins

To study the evolutionary relationships among XTH proteins in dicots and monocots, a phylogenetic tree was constructed using the full-length candidate XTH protein sequences in *Brachypodium* and other three species, including 29 BdXTHs, 33 AtXTHs from *Arabidopsis*, 30 OsXTHs from rice, and 36 SlXTHs from tomato (Table S1). The results showed that the XTH proteins cluster into three main groups (Group I/II, Group III, and Ancestral Group), and that each phylogenetic group contains XTH proteins from the four species (Fig. 1). Not unexpectedly, proteins from closely related species clustered together; proteins from the monocots (*B. distachyon* and rice) tended to cluster together and proteins from the dicots (tomato and *Arabidopsis*) clustered together. Furthermore, the different groups contained different numbers of XTH proteins. Group I/II contained the largest number of XTH proteins, including 19 BdXTHs, 18 OsXTHs, 22 AtXTHs, and 27 SlXTHs. Group III was further divided into two subgroups (Group III-A and Group III-B), which included two and six BdXTHs, respectively, along with XTH proteins from the other three species included in the analysis. The Ancestral Group contains the smallest number of XTH proteins; two BdXTHs, one OsXTH, four AtXTHs, and two SlXTHs.

Structural and conserved motif analyses of the *BdXTH* genes

Different combinations of exons and introns can lead to diverse gene functions. To gain more knowledge of the structural diversity of the *BdXTH* genes, the structures of the 29 *BdXTH* genes were analyzed using GSDS. The results showed that the number of exons varied from three to five, and that the structures of genes from the same group showed more similarity to one another (Fig. 2b). Most of the *BdXTH* genes (20/29) contained three exons. Eight genes contained four exons and only one gene (*BdXTH19*) in Group III-B contained five exons. In addition, only *BdXTH26* from Group I/II contained two longer introns. Moreover, three genes (*BdXTH15*, *18*, and *27*) from Group I/II had longer UTR sequences compared with the other genes.

A conserved motif analysis of all 29 predicted BdXTH protein sequences from *B. distachyon* conducted using the MEME program predicted 20 motifs (Fig. 3; Fig. S1);

the number of motifs varied from nine to thirteen, and members of the same group usually shared a similar motif composition. The proteins in Group III-A and the Ancestral Group had relatively fewer motifs, nine and ten, respectively, while most of the XTH proteins in the other groups had 12 motifs. Motifs 1, 3, and 4 were found to be highly conserved in all BdXTH proteins. In addition, several conserved motifs were specific to certain groups. For example, motifs 10, 13, and 17-19 were only present in Group I/II proteins, and motifs 15 and 16 were unique to proteins in Group III-B.

Chromosomal location and synteny analysis of the *BdXTH* genes

The chromosomal positions of 29 *BdXTH* genes were located using information derived from the *Brachypodium* genome [30]. The *BdXTH* genes were found to be widely distributed on the chromosomes, but the distribution was not uniform (Fig. 3). None of the *BdXTH* genes were located on Chr. 2. Only two *BdXTH* gens were located on Chr. 4, and five *BdXTH* genes were located on the end of Chr. 5. In addition, most *BdXTH* genes were located on Chrs. 1 and 3, with 12 and 10 genes, respectively.

To identify potential duplication events in the *BdXTH* gene family, a collinearity analysis was performed using MCScanX software. Results revealed that there are three pairs of segmentally duplicated *BdXTH* genes (*BdXTH15/25*, *BdXTH17/27*, and *BdXTH18/27*) and five groups of tandemly duplicated *BdXTH* genes (Fig. 4; *BdXTH1/2*, *BdXTH5/6/7/8*, *BdXTH15/16*, *BdXTH25/26*, and *BdXTH27/28*). *BdXTH15*, *BdXTH25*, and *BdXTH27* were involved in both tandem duplications and segmental duplications. Also, all of the duplicated genes are in Group I/II and account for ~48% (14/29) of all *BdXTH* genes, indicating that tandem duplication and segmental duplication have played important roles in the expansion of the *BdXTH* gene family in *B. distachyon*.

Structure-based sequence alignment

The secondary structures of the BdXTH proteins were predicted by aligning the 29

BdXTH protein sequences with those of two other proteins for which the structures have been experimentally determined (PttXET16-34, PDB id: 1UN1 and TmNXG1, PDB id: 2UWA) [15, 32], using ESPript (http://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi). This analysis showed that the active site (ExDxE) responsible for catalytic activity is highly conserved in all of the BdXTH protein family members (Fig. 4; Fig. S2; Fig. S3). The first glutamate residue (E) is the catalytic nucleophile that initiates the enzymatic reaction, and the second E residue acts as a base to activate the entrant substrate. The *N*-glycosylation site denominated as NxT/S/Y (marked with asterisks) can bind *N*-glycans and is related to protein stability. We found that the *N*-glycosylation site was present in almost all BdXTH proteins except the Ancestral Group members (BdXTH9 and BdXTH11). We also found that the distance between the *N*-glycosylation site and the active site in the Group I/II members was closer than in the Group III members. The BdXTH proteins also contain conserved domains next to the substrate binding site that are called loop 1, loop 2, and loop 3 (underlined in green). Loop 2 in the Group III-A proteins (BdXTH4 and BdXTH12) is longer compared to that in the other groups, which may be the reason why proteins in Group III-A mainly show xyloglucan hydrolase activity.

Cis-element analysis of the _BdXTH_ gene promoter regions

XTH family genes play important roles during plant growth and development, as well as in the response to multiple environmental stresses. *Cis*-elements can regulate gene expression via their interactions with the trans-acting elements such as transcription factors. To understand the functions and regulation network of the *BdXTH* family genes, we analyzed the 2,000 bp of DNA sequence upstream of the promoter regions for 29 *BdXTH* genes on the PlantCARE database [33]. As shown in Fig. 5, the promoter regions include several hormone-related (abscisic acid, MeJA, auxin, salicylic acid, and gibberellin) *cis*-elements (ABRE, CGTCA-motif, TGACG-motif, TGA-element, AuxRR-core, TCA-element, P-box, GARE-motif, TATC-box), and growth regulation *cis*-elements (MBSI, RY-element, CAT-box, HD-Zip 1, O2-site, GCN4_motif, MSA-like, circadian motif I). The promoter regions also contain several

environmental response elements, such as those for anaerobic (ARE), drought (MBS), low temperature (LTR), and anoxic (GC-motif) conditions; light responsive elements (GT1-motif); and defense and stress (TC-rich repeats) *cis*-elements. The distribution of *cis*-elements analyzed in the promoter regions is shown in Fig. S4. *Cis*-acting elements involved in the responses to phytohormones were relatively abundant, especially the abscisic acid-responsive (ABRE) element, and accounted for 27% of the total number of *cis*-elements detected in this study. The *cis*-elements involved in cell cycle regulation (MSA-like), root development (motif I), and flavonoid biosynthesis gene regulation (MBSI) were only found in the promoter regions of several Group I/II genes. In addition, two Group III-B genes (*BdXTH22*, *BdXTH19*) contain the highest (36) and lowest (14) numbers of *cis*-elements identified, respectively.

Expression patterns of the *BdXTH* genes in various tissues

To explore the temporal and spatial expression patterns of the *BdXTH* genes in *Brachypodium*, RNA-seq data including expression profiles from various tissues and developmental stages (leaf, inflorescence, anther, pistil, seed, embryo, and endosperm) were downloaded from the NCBI database (SRP008505). Based on the clusters in the phylogenetic tree, the expression patterns of genes from the same group tended to be different, and most of the *BdXTH* genes exhibited distinct tissue-specific expression patterns (Fig. 6). For example, the *BdXTH21/26/27* genes are mainly expressed in the endosperm, *BdXTH1/2* are mainly expressed in the seed, *BdXTH8/9* are mainly expressed in the pistil, *BdXTH7/13/17* are mainly expressed in the anther, and *BdXTH3/4/14/19/23/24/29* are mainly expressed in the inflorescence, whereas expression levels were found to be relatively low in other organs, suggesting that the functions of some genes are redundant and that certain genes participate in the developmental progress of specific tissues. In addition, three genes (*BdXTH11/20/28*) all showed lower expression levels at almost all developmental stages and in almost all tissues included in the study.

Expression profiling of *BdXTH* genes in response to abiotic stress and phytohormone treatments using qRT-PCR

Previous studies have shown that the expression of *XTH* family genes can respond to multiple abiotic stresses and phytohormones. To investigate the effect of different treatment conditions on *BdXTH* gamily genes, qRT-PCR assays were performed to study the expression patterns of 29 *BdXTH* genes in response to drought (PEG), salinity (NaCl), and four plant hormones (ABA, BR, IAA, and GA3) (Fig. 7). Analysis of the data showed that the expression levels of many genes were affected by the different treatments. For the PEG and NaCl treatments, the expression patterns of the *BdXTH* genes were similar. Among the up-regulated genes, *BdXTH7/11/13/15/19/21/25* showed relatively higher mRNA levels, in which *BdXTH11* (Ancestral Group) had the highest expression level under both stress treatments. In addition, almost all of the genes that were down-regulated in response to PEG treatment, such as *BdXTH1-6/12/16/17/20/23/27*, also showed relatively lower expression levels under NaCl treatment. For the phytohormone treatments, *BdXTH19* from Group III-B showed the highest mRNA levels in response to BR, IAA, and GA3 treatments, and *BdXTH7* from Group I/II showed the highest level when treated with ABA. The altered expression patterns of the *BdXTH* genes suggests that they might be involved in adapting to adverse environmental factors, or are regulated by diverse plant hormones.

Discussion

Xyloglucan endotransglucosylase/hydrolases (XTHs), one family of cell-wall-modifying enzymes, can cut and/or rejoin xyloglucan molecules to regulate the composition and organization of the cell wall [34]. In this study, 29 *BdXTH* genes were identified in the *Brachypodium* genome, and a phylogenetic analysis showed that they are organized into three major groups (Group I/II, Group III, and the Ancestral Group). A similar XTH protein classification was also reported in strawberry (*Fragaria vesca*) [9], barley (*Hordeum vulgare*) [13], and pineapple (*Ananas comosus*) [12]. The number of *BdXTH* genes in *Brachypodium* was similar to

the number of *XTH* genes identified previously in other species, such as *Arabidopsis* [5] , rice (29) [14], and strawberry (26) [9], but was much lower than the number of *XTH* genes in tobacco (56) [10], soybean (61) [11], and wheat (57) [35]. Duplication events (tandem duplication or segmental duplication events) resulting from whole genome duplication (WGD) via polyploidization or local chromosomal rearrangement played an important role in *XTH* gene family expansion and evolution. Duplication events are likely to have arisen in an immediate common ancestor, and duplicated genes tend to have a closer relationship. The phylogenetic tree showed that all of the genes arising from tandem or segmental duplication events are clustered together and impact genes belonging to Group I/II, which is the largest group compared to the other groups. For example, the tandemly duplicated gene pairs *BdXTH1/2* and *BdXTH5/6/7/8* on chromosome 1, *BdXTH15/16* on chromosome 3, *BdXTH25/26* and *BdXTH27/28* on chromosome 5, and the segmentally duplicated gene pairs *BdXTH15/25*, *BdXTH17/27*, and *BdXTH18/27* are all members of Group I/II.

The BdXTH proteins in each group contain relatively conserved motifs and the genes structures are also conserved, suggesting that BdXTH proteins in the same group may perform similar functions (Fig. 2; Fig. 4). The results of secondary structure prediction for the translated BdXTH sequences confirms the existence of a highly conserved domain (DEIDFEFLG) that is the catalytic site for both XET and XEH activities [4], especially the three absolutely conserved catalytic residues (ExDxE) in all BdXTH proteins. Also, the *N*-glycosylation site in BdXTH proteins from Group I/II is adjacent to the catalytic domain, but tended to be located towards the carboxyl terminus in Group III-B proteins. In addition, loop 2 in Group III-A protein is longer compared to that in Group III-B members, which has been proposed to be a major structural change responsible for the endo-hydrolase activity of these proteins [15]. Previous studies have reported that XTH proteins in Group III-A mainly display XEH activity [16, 17, 36], while Group I/II and Group III-B proteins mainly shown XET activity [37].

A comparison of the predicted *XTH* gene amino acid sequences from *B. distachyon*, rice (*Oryza sativa*), tomato (*Solanum lycopersicum*), and Arabidopsis (*A. thaliana*)

revealed the high degree of conservation of the XTH sequences among various plant species, implying a general functional conservation of these proteins in the plant kingdom. Our phylogenetic analysis revealed that all clades contained proteins from both the monocot (*Brachypodium* and rice) and dicot (tomato and Arabidopsis) species included in the analysis. This suggests that the precursor genes were present in the most recent common ancestor of monocots and dicots, and that closely related proteins might perform similar functions in the different species, although xyloglucan makes up a relatively smaller fraction of the cell wall in Poales species [6, 38].

Analysis of the genomic sequences upstream of the *BdXTH* genes using the PlantCARE tool revealed that a series of *cis*-elements involved in the responses to hormones and abiotic stresses, and also in growth and development, are present in the promoter regions, indicating that *BdXTH* genes play diverse roles in *Brachypodium*. Clearly, *BdXTH* genes respond to drought, salinity, and several plant hormones (ABA, BR, IAA, and GA3) based on the changes in gene expression patterns observed in the qRT-PCR experiment (Fig. 7), which corroborate the results of the analyses of *BdXTH* gene promoter regions. Expression profiling suggests that *BdXTH* genes might be involved in adaptation to adverse environmental conditions, or that they are regulated by many plant hormones. In addition, transcriptional analysis of *BdXTH* genes in several tissues and at different development stages from an existing RNA-seq dataset showed tissue, organ, and temporal specificity, but correlations between phylogenetic groups and specific tissue expression patterns were weak (Fig. 6). For example, *BdXTH1*, *8*, *17*, *23*, and *27* are all members of Group I/II, but they are expressed at high levels in seeds, pistils, anthers, early inflorescences, and the endosperm, respectively. Also, many of the segmentally-duplicated gene pairs with close evolutionary relationships have different expression patterns. For example, the expression of *BdXTH15* is at its highest in the inflorescence and seed, but the highest expression levels of *BdXTH25* are found in the anther and endosperm. *BdXTH27* showed its highest expression level in the endosperm, but the expression of *BdXTH28* was low in all tissues examined in this study. These results suggest that genes in the *BdXTH* family might have undergone neofunctionalization or sub-functionalization

during evolution. The specific function and molecular analysis of each gene will require further study.

Conclusions

In this study, we performed a genome-wide analysis of the *BdXTH* gene family in *Brachypodium* and investigated the expression profiles of all 29 genes in different tissues and developmental stages, as well as in response to various stress conditions. Many of the *BdXTH* genes exhibit distinct tissue-specific expression patterns, indicating that various *BdXTH* genes may play a universal cell wall modification function in specific tissues and during specific stages of development. In addition, the changes in the relative expression of most *BdXTH* genes in response to diverse abiotic stresses provides an indication of the possible functions of *BdXTH* genes in response to drought, salinity, and several phytohormones. These results may advance our understanding of the role of *BdXTH* genes in the regulation of *Brachypodium* growth and development and also in its response to abiotic stresses.

Methods

Identification of the *XTH* family genes in *B. distachyon*

The latest version of the *Brachypodium distachyon* (v3.1) genome annotations was downloaded from the Phytozome database (https://phytozome.jgi.doe.gov/pz/portal.html) [39]. The Hidden Markov Model (HMM) profiles of the XTH protein domains, PF00722 and PF06955, were downloaded from the Pfam database [40], and were used as queries to search the database using the program HMMER3.0 with the default E-value. The online program SMART (http://smart.embl-heidelberg.de/) [41] and the PFAM databases (https://pfam.xfam.org) [42] were used to identify the conserved domains of candidate *Brachypodium* XTH proteins. Only proteins containing both the PF00722 and PF06955 domains were retained for further study. ProtParam (http://web.expasy.org/protparam/) was used to predict the physical and chemical

features of the BdXTH proteins. The subcellular locations of the BdXTH proteins were predicted using the online website Plant-mPLoc in Cell-PLoc 2.0 (http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc-2/) [43]. All of the XTH protein sequences from tomato, *Arabidopsis*, and rice were also downloaded from the Phytozome database (https://phytozome.jgi.doe.gov/pz/portal.html). The corresponding gene IDs of the XTH protein family members are given in Table S1.

Phylogenetic tree construction

A phylogenetic tree was constructed using the Neighbor Joining (NJ) method as implemented in MEGA-X software, and the branch support was estimated by bootstrapping with 1,000 replicates [44].The NJ tree was constructed from an alignment of 128 amino acid sequences of predicted XTH proteins, of which 29 BdXTHs were from *B. distachyon*, 36 SIXTHs were from *S. lycopersicum*, 30 OsXTHs were from *O. sativa*, and 33 AtXTHs were from *A. thaliana* (Table S1). The phylogenetic tree was then visualized using iTol (https://itol.embl.de/) [45].

Gene structures, conserved protein motifs, and *cis*-acting regulatory element analysis

The *BdXTH* gene structures were displayed using the Gene Structure Display Server (GSDS) tool (http://gsds.cbi.pku.edu.cn/) [46] by aligning the cDNA sequences and the corresponding genomic DNA sequences. The Multiple EM for Motif Elicitation (MEME, http://meme-suite.org/) [47] was used to search for possible conserved motifs in the complete amino acid sequences of predicted BdXTH proteins using the default settings. Additionally, the *cis*-elements in the *BdXTH* gene promoter regions (2,000-bp of genomic sequence upstream of the coding sequences) were analyzed using the PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) [33].

Chromosomal location and gene duplication

The chromosomal positions of *BdXTH* genes were acquired from the *Brachypodium distachyon* genome (v3.1). MCScanX (default parameters) [48] was used to analyze gene duplications using the amino acid sequences and chromosomal location data. The chromosomal locations and gene duplication relationships of the *BdXTH* genes

were displayed using TBtools (https://github.com/CJ-Chen/TBtools) [49].

Structural-based sequence alignment analysis

Alignment of the identified BdXTH protein sequences with two other proteins, TmNXG1 (PDB id: 2UWA) and PttXET16-34 (PDB id: 1UN1), for which the structures have been experimentally determined, was performed to identify common structural elements. The crystal structures of TmNXG1and PttXET16-34 were obtained from the PDB databank (www.pdb.org). The secondary structures of the BdXTH proteins were then predicted using the online website ESPript (http://espript.ibcp.fr/ESPript/ESPript/) [50].

Expression pattern analysis of *BdXTH* genes using RNA-seq data

RNA-seq data from various *Brachypodium* tissues and developmental stages (leaves, pre- and post-emergence flowers, anthers, pistils, whole seeds at 5 days after pollination, whole seeds at 10 days after pollination, embryos at 25 days after pollination, and endosperm at 25 days after pollination) were downloaded from the NCBI database (SRP008505) [51] to analyze the expression patterns of the different *BdXTH* genes. The heat maps and hierarchical clustering were created by TBtools (https://github.com/CJ-Chen/TBtools) [49] using FPKM values extracted from the RNA-seq dataset.

Plant materials and treatments

Seeds of *Brachypodium distachyon* (ecotype Bd21) provided by Professor Hailong An of Shandong Agricultural University were placed on wet filter paper and maintained for 3 days at 4℃ in the dark, after which they were cultivated in a plant growth incubator under long-day conditions (18h light/6 h dark) at 20℃ for one week. The young seedlings were then transferred to custom-made plastic vessels with holes in the bottoms suspended on the surface of a reservoir containing 0.5X MS liquid medium and maintained in a growth chamber under the same conditions (18 h light/6 h dark, 20℃). Seedlings with three leaves were subjected to different abiotic stress and phytohormone treatments, which included 150 mM NaCl, 20% polyethylene glycol (PEG) 6000, 1 μM 3-indole acetic acid (IAA), 1 μM gibberellic acid (GA_3), 1 μM abscisic acid (ABA), and 1 μM 24-epibrassinolide (BR). Roots of the

stress-treated and control plants were collected after 2 h of treatment. All of the materials were immediately frozen in liquid nitrogen and stored at -80℃ prior to RNA extraction. Every sample comprised three independent biological replications.

RNA isolation and qRT-PCR gene expression analysis

The frozen samples were ground to powder in liquid nitrogen with a mortar and pestle. Total RNA was isolated from roots using the RNAprep Pure Plant Kit (Tiangen, Beijing, China) and first-strand total cDNA was then synthesized using the HiScript® II Q RT SuperMix for qPCR (+gDNA wiper) Kit (Vazyme, China). Beacon designer software was used to design the gene-specific primers for qRT-PCT (Table S2). Real-time qRT-PCR assays were performed using ChamQ™ SYBR® qPCR Master Mix (Vazyme, China) on a Bio-Rad CFX96 Real-time PCR System (Bio-Rad, USA). The *BdUBC18* gene was used as the internal control for normalization of gene expression. Each PCR contained 0.4 μL of each primer, 1 μL of template cDNA and 10 μL of 2XChamQ SYBR qPCR Master Mix in a final volume of 20 μL. The thermal cycling protocol was as follows: 95 °C for 30 s followed by 40 cycles of 95 °C for 10 s and 60 °C for 30 s. Subsequently, melting curves were performed to confirm the specificity of the primers. Each reaction was performed three times, and the $2^{-\Delta\Delta Ct}$ method [52] was used to calculate the relative gene expression levels.

Declarations

Abbreviations

XTH: Xyloglucan endotransglucosylase/hydrolase; IAA: Auxin; GA3: Gibberellin; ABA: Abscisic acid; BR: Brassinolide; MW: Molecular weight; MeJA: Methyl jasmonate; PEG: Polyethylene glycol; qRT-PCR: Quantitative real time polymerase chain reaction; GSDS: Gene Structure Display Server; NJ: Neighbour-Joining

Acknowledgements

We thank Professor Hailong An for providing *Brachypodium distachyon* (ecotype Bd21) seeds.

Authors' contributions

HW and XY designed the research. HS, QT, WX and WD performed the research and analyzed the data. HS wrote the manuscript. HW and XY revised the manuscript. All authors have read and approved the manuscript.

Funding

This work was financially supported by the Introduction and Training Plan of Young Creative Talents in Universities of Shandong Province: Research Group of Forest Tree Biotechnology (CN). The funding body was not involved in the design of the study; collection, analysis, or interpretation of data; or manuscript writing.

Availability of data and materials

All of the data and materials supporting our research findings are contained in the methods section of the manuscript. Details are provided in the attached additional files.

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. Van Sandt VS, Suslov D, Verbelen JP, Vissenberg K: **Xyloglucan endotransglucosylase activity loosens a plant cell wall**. *Annals of botany* 2007, **100**(7):1467-1473.
2. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B: **The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics**. *Nucleic acids research* 2009, **37**(Database issue):D233-238.
3. Eklöf JM, Brumer H: **The XTH gene family: an update on enzyme structure, function, and phylogeny in xyloglucan remodeling**. *Plant Physiol* 2010, **153**(2):456-466.
4. Campbell P, Braam J: **Xyloglucan endotransglycosylases: diversity of genes, enzymes and potential wall-modifying functions**. *Trends in plant science* 1999, **4**(9):361-366.
5. Yokoyama R, Nishitani K: **A comprehensive expression analysis of all members of a gene family encoding cell-wall enzymes allowed us to predict cis-regulatory regions involved in cell-wall construction in specific organs of Arabidopsis**. *Plant & cell physiology* 2001, **42**(10):1025-1033.
6. Hara Y, Yokoyama R, Osakabe K, Toki S, Nishitani K: **Function of xyloglucan endotransglucosylase/hydrolases in rice**. *Annals of botany* 2014, **114**(6):1309-1318.
7. Saladié M, Rose JK, Cosgrove DJ, Catalá C: **Characterization of a new xyloglucan endotransglucosylase/hydrolase (XTH) from ripening tomato fruit and implications for the diverse modes of enzymic action**. *The Plant journal : for cell and molecular biology* 2006, **47**(2):282-295.
8. Xuan Y, Zhou ZS, Li HB, Yang ZM: **Identification of a group of XTHs genes responding to heavy metal mercury, salinity and drought stresses in Medicago truncatula**. *Ecotoxicology and environmental safety* 2016, **132**:153-163.
9. Opazo MC, Lizana R, Stappung Y, Davis TM, Herrera R, Moya-León MA: **XTHs from Fragaria vesca: genomic structure and transcriptomic analysis in ripening fruit and other tissues**. *BMC genomics* 2017, **18**(1):852.
10. Wang M, Xu Z, Ding A, Kong Y: **Genome-Wide Identification and Expression Profiling Analysis of the Xyloglucan Endotransglucosylase/Hydrolase Gene Family in Tobacco (Nicotiana tabacum L.)**. *Genes* 2018, **9**(6).
11. Song L, Valliyodan B, Prince S, Wan J, Nguyen HT: **Characterization of the XTH Gene Family: New Insight to the Roles in Soybean Flooding Tolerance**. *International journal of molecular sciences* 2018, **19**(9).
12. Li Q, Li H, Yin C, Wang X, Jiang Q, Zhang R, Ge F, Chen Y, Yang L: **Genome-Wide Identification and Characterization of Xyloglucan Endotransglycosylase/Hydrolase in Ananas comosus during Development**. *Genes* 2019, **10**(7).
13. Fu MM, Liu C, Wu F: **Genome-Wide Identification, Characterization and**

Expression Analysis of Xyloglucan Endotransglucosylase/Hydrolase Genes Family in Barley (Hordeum vulgare). *Molecules* 2019, **24**(10).

14. Yokoyama R, Rose JK, Nishitani K: **A surprising diversity and abundance of xyloglucan endotransglucosylase/hydrolases in rice. Classification and expression analysis**. *Plant Physiol* 2004, **134**(3):1088-1099.
15. Baumann MJ, Eklöf JM, Michel G, Kallas AM, Teeri TT, Czjzek M, Brumer H, 3rd: **Structural evidence for the evolution of xyloglucanase activity from xyloglucan endo-transglycosylases: biological implications for cell wall metabolism**. *The Plant cell* 2007, **19**(6):1947-1963.
16. Zhu XF, Shi YZ, Lei GJ, Fry SC, Zhang BC, Zhou YH, Braam J, Jiang T, Xu XY, Mao CZ *et al*: **XTH31, encoding an in vitro XEH/XET-active enzyme, regulates aluminum sensitivity by modulating in vivo XET action, cell wall xyloglucan content, and aluminum binding capacity in Arabidopsis**. *The Plant cell* 2012, **24**(11):4731-4747.
17. Kaewthai N, Gendre D, Eklöf JM, Ibatullin FM, Ezcurra I, Bhalerao RP, Brumer H: **Group III-A XTH genes of Arabidopsis encode predominant xyloglucan endohydrolases that are dispensable for normal growth**. *Plant Physiol* 2013, **161**(1):440-454.
18. Han Y, Zhu Q, Zhang Z, Meng K, Hou Y, Ban Q, Suo J, Rao J: **Analysis of xyloglucan endotransglycosylase/hydrolase (XTH) genes and diverse roles of isoenzymes during persimmon fruit development and postharvest softening**. *PLoS One* 2015, **10**(4):e0123668.
19. Witasari LD, Huang FC, Hoffmann T, Rozhon W, Fry SC, Schwab W: **Higher expression of the strawberry xyloglucan endotransglucosylase/hydrolase genes FvXTH9 and FvXTH6 accelerates fruit ripening**. *The Plant journal : for cell and molecular biology* 2019, **100**(6):1237-1253.
20. Lee YK, Rhee JY, Lee SH, Chung GC, Park SJ, Segami S, Maeshima M, Choi G: **Functionally redundant LNG3 and LNG4 genes regulate turgor-driven polar cell elongation through activation of XTH17 and XTH24**. *Plant molecular biology* 2018, **97**(1-2):23-36.
21. Wang B, Zhang D: **Association of allelic variation in PtoXET16A with growth and wood properties in Populus tomentosa**. *International journal of molecular sciences* 2014, **15**(9):16949-16974.
22. Han Y, Han S, Ban Q, He Y, Jin M, Rao J: **Overexpression of persimmon DkXTH1 enhanced tolerance to abiotic stress and delayed fruit softening in transgenic plants**. *Plant Cell Rep* 2017, **36**(4):583-596.
23. Zhu XF, Wan JX, Sun Y, Shi YZ, Braam J, Li GX, Zheng SJ: **Xyloglucan Endotransglucosylase-Hydrolase17 Interacts with Xyloglucan Endotransglucosylase-Hydrolase31 to Confer Xyloglucan Endotransglucosylase Action and Affect Aluminum Sensitivity in Arabidopsis**. *Plant Physiol* 2014, **165**(4):1566-1574.
24. Han Y, Ban Q, Hou Y, Meng K, Suo J, Rao J: **Isolation and Characterization of Two Persimmon Xyloglucan**

Endotransglycosylase/Hydrolase (XTH) Genes That Have Divergent Functions in Cell Wall Modification and Fruit Postharvest Softening. *Frontiers in plant science* 2016, **7**:624.

25. Opanowicz M, Vain P, Draper J, Parker D, Doonan JH: **Brachypodium distachyon: making hay with a wild grass**. *Trends in plant science* 2008, **13**(4):172-177.
26. Ozdemir BS, Hernandez P, Filiz E, Budak H: **Brachypodium genomics**. *International journal of plant genomics* 2008, **2008**:536104.
27. Brkljacic J, Grotewold E, Scholl R, Mockler T, Garvin DF, Vain P, Brutnell T, Sibout R, Bevan M, Budak H *et al*: **Brachypodium as a model for the grasses: today and the future**. *Plant Physiol* 2011, **157**(1):3-13.
28. Hsia MM, O'Malley R, Cartwright A, Nieu R, Gordon SP, Kelly S, Williams TG, Wood DF, Zhao Y, Bragg J *et al*: **Sequencing and functional validation of the JGI Brachypodium distachyon T-DNA collection**. *The Plant journal : for cell and molecular biology* 2017, **91**(3):361-370.
29. Wu H, Xue X, Qin C, Xu Y, Guo Y, Li X, Lv W, Li Q, Mao C, Li L *et al*: **An Efficient System for Ds Transposon Tagging in Brachypodium distachyon**. *Plant Physiol* 2019, **180**(1):56-65.
30. International Brachypodium I: **Genome sequencing and analysis of the model grass Brachypodium distachyon**. *Nature* 2010, **463**(7282):763-768.
31. Gordon SP, Contreras-Moreira B, Woods DP, Des Marais DL, Burgess D, Shu S, Stritt C, Roulin AC, Schackwitz W, Tyler L *et al*: **Extensive gene content variation in the Brachypodium distachyon pan-genome correlates with population structure**. *Nature communications* 2017, **8**(1):2184.
32. Johansson P, Brumer H, 3rd, Baumann MJ, Kallas AM, Henriksson H, Denman SE, Teeri TT, Jones TA: **Crystal structures of a poplar xyloglucan endotransglycosylase reveal details of transglycosylation acceptor binding**. *The Plant cell* 2004, **16**(4):874-886.
33. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S: **PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences**. *Nucleic acids research* 2002, **30**(1):325-327.
34. Rose JK, Braam J, Fry SC, Nishitani K: **The XTH family of enzymes involved in xyloglucan endotransglucosylation and endohydrolysis: current perspectives and a new unifying nomenclature**. *Plant & cell physiology* 2002, **43**(12):1421-1435.
35. Liu Y, Liu D, Zhang H, Gao H, Guo X, Wang D, Zhang X, Zhang A: **The alpha- and beta-expansin and xyloglucan endotransglucosylase/hydrolase gene families of wheat: molecular cloning, gene expression, and EST data mining**. *Genomics* 2007, **90**(4):516-529.
36. Fanutti C, Gidley MJ, Reid JS: **Action of a pure xyloglucan endo-transglycosylase (formerly called xyloglucan-specific endo-(1-->4)-beta-D-glucanase) from the cotyledons of germinated nasturtium seeds**. *The Plant journal : for cell and molecular biology* 1993,

3(5):691-700.
37. Maris A, Suslov D, Fry SC, Verbelen JP, Vissenberg K: **Enzymic characterization of two recombinant xyloglucan endotransglucosylase/hydrolase (XTH) proteins of Arabidopsis and their effect on root growth and cell wall extension**. *J Exp Bot* 2009, **60**(13):3959-3972.
38. Carpita NC, Gibeaut DM: **Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth**. *The Plant journal : for cell and molecular biology* 1993, **3**(1):1-30.
39. Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N *et al*: **Phytozome: a comparative platform for green plant genomics**. *Nucleic acids research* 2012, **40**(Database issue):D1178-1186.
40. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J *et al*: **The Pfam protein families database**. *Nucleic acids research* 2012, **40**(Database issue):D290-301.
41. Letunic I, Doerks T, Bork P: **SMART 7: recent updates to the protein domain annotation resource**. *Nucleic acids research* 2012, **40**(Database issue):D302-305.
42. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A *et al*: **The Pfam protein families database in 2019**. *Nucleic acids research* 2019, **47**(D1):D427-D432.
43. Chou KC, Shen HB: **Plant-mPLoc: a top-down strategy to augment the power for predicting plant protein subcellular localization**. *PLoS One* 2010, **5**(6):e11335.
44. Kumar S, Stecher G, Li M, Knyaz C, Tamura K: **MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms**. *Molecular biology and evolution* 2018, **35**(6):1547-1549.
45. Letunic I, Bork P: **Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees**. *Nucleic acids research* 2016, **44**(W1):W242-245.
46. Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G: **GSDS 2.0: an upgraded gene feature visualization server**. *Bioinformatics* 2015, **31**(8):1296-1297.
47. Bailey TL, Johnson J, Grant CE, Noble WS: **The MEME Suite**. *Nucleic acids research* 2015, **43**(W1):W39-49.
48. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H *et al*: **MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity**. *Nucleic acids research* 2012, **40**(7):e49.
49. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R: **TBtools - an integrative toolkit developed for interactive analyses of big biological data**. *Mol Plant* 2020.
50. Robert X, Gouet P: **Deciphering key features in protein structures with the new ENDscript server**. *Nucleic acids research* 2014, **42**(Web Server

issue):W320-324.
51. Davidson RM, Gowda M, Moghe G, Lin H, Vaillancourt B, Shiu SH, Jiang N, Robin Buell C: **Comparative transcriptomics of three Poaceae species reveals patterns of gene expression evolution**. *The Plant journal : for cell and molecular biology* 2012, **71**(3):492-502.
52. 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.

Figure legends

Fig. 1 Phylogenetic relationships among XTH proteins from *Brachypodium* and three other plant species. The phylogenetic tree was constructed with the Neighbor Joining (NJ) method as implemented in MEGA-X software, and branch confidence was estimated by bootstrapping with 1,000 replicates. The open blue, gray, red, and black circles indicate proteins from *Brachypodium*, rice, tomato, and *Arabidopsis*, respectively. In addition, the XTH proteins are classified into three clades (Group I/II, Group III, and the Ancestral Group). Proteins in Group I/II and the Ancestral Group are shown with purple and blue backgrounds, respectively. Group III is further divided into two subclades, Group III-A and Group III-B, which are indicated with red and green backgrounds, respectively.

Fig. 2 Unrooted neighbor-joining phylogenetic tree, conserved protein motifs, and structural analysis of *BdXTH* genes. (**a**) Phylogenetic relationships of the XTH proteins in *Brachypodium*. Proteins from the four clades (Group I/II, Group III-A and Group III-B, and the Ancestral Group) are color coded as in Figure 1. (**b**) The structures of the 29 putative *BdXTH* genes. The UTRs, exons, and introns are represented by green boxes, yellow boxes, and black lines, respectively. (**c**) Conserved motif analysis of the BdXTH proteins. The different motifs are indicated by different colored boxes numbered motif 1 to motif 20. The structural features of the 20 motifs are shown in Fig S1.

Fig. 3 The physical locations of *BdXTH* genes on the five *Brachypodium* chromosomes. Tandemly duplicated gene pairs and segmentally duplicated genes are linked by red lines. The chromosome numbers are displayed at the top of each chromosome and the scale in megabases (Mb) is shown on the left.

Fig. 4 Structure-based sequence alignment of BdXTH proteins. The structures of two proteins (PttXET16-34, PDB id: 1UN1; TmNXG1, PDB id: 2UWA) have been experimentally determined. Proteins in Group I/II and the Ancestral Group had similar structures to 1UN1, and proteins in Group III show similar structures to 2UWA. The active site (ExDxE), and loops 1, 2, and 3 are underlined in black and green, respectively. The *N*-glycosylation site residues are indicated by asterisks.

Fig. 5 Numbers of *cis*-acting elements in the promoter regions of the 29 *BdXTH* genes. Three types of *cis*-acting elements in the 2,000 bp of DNA sequence upstream of the promoter regions are shown in the figure, including phytohormone- and environmentally-responsive elements and plant growth and development-related elements. Members of the different element classes are shown at the top of the figure in different shades of gray.

Fig. 6 Heat map showing the expression pattern of *BdXTH* genes in *Brachypodium*. Expression profiles from various tissues and developmental stages (leaves, early inflorescences, emerging inflorescences, anthers, pistils, seeds at 5 days after pollination, seeds at 10 days after pollination, embryos, and endosperm) were downloaded from the NCBI database (SRP008505). The relative expression levels are represented by the colored bars. Red and green boxes indicate high and low expression levels, respectively.

Fig. 7 Expression analysis of *BdXTH* genes under different conditions. Quantitative real-time polymerase chain reaction (qRT-PCR) analysis of *BdXTH* gene expression in response to abiotic stresses (drought and salinity) (**a**), and phytohormone treatments (ABA, BR, IAA, and GA3) (**b**). The means ± SD of three biological replicates are presented.

Fig. S1 Sequence logos of 20 conserved motifs detected in the BdXTH proteins. The overall height of each stack represents the degree of conservation at that position, and the height of the individual letter within each stack indicates the relative frequency of the corresponding amino acid.

Fig. S2 Multiple alignment of deduced amino acid sequences of BdXTH proteins from Group I/II, the Ancestral Group and 1UN1.

Fig. S3 Multiple alignment of deduced amino acid sequences of BdXTH proteins from Group III and 2UWA.

Fig. S4 *Cis*-acting elements identified in the 2,000 bp of sequence upstream of the promoter regions of the *BdXTH* genes.

Table 1 The physicochemical properties of XTH gamily genes in *Brachypodium*.

Table S1 Identification of *XTH* gene family members from three plant species.

Table S2 Oligonucleotide primers used for qRT-PCR assays in this study.

Fig. 1

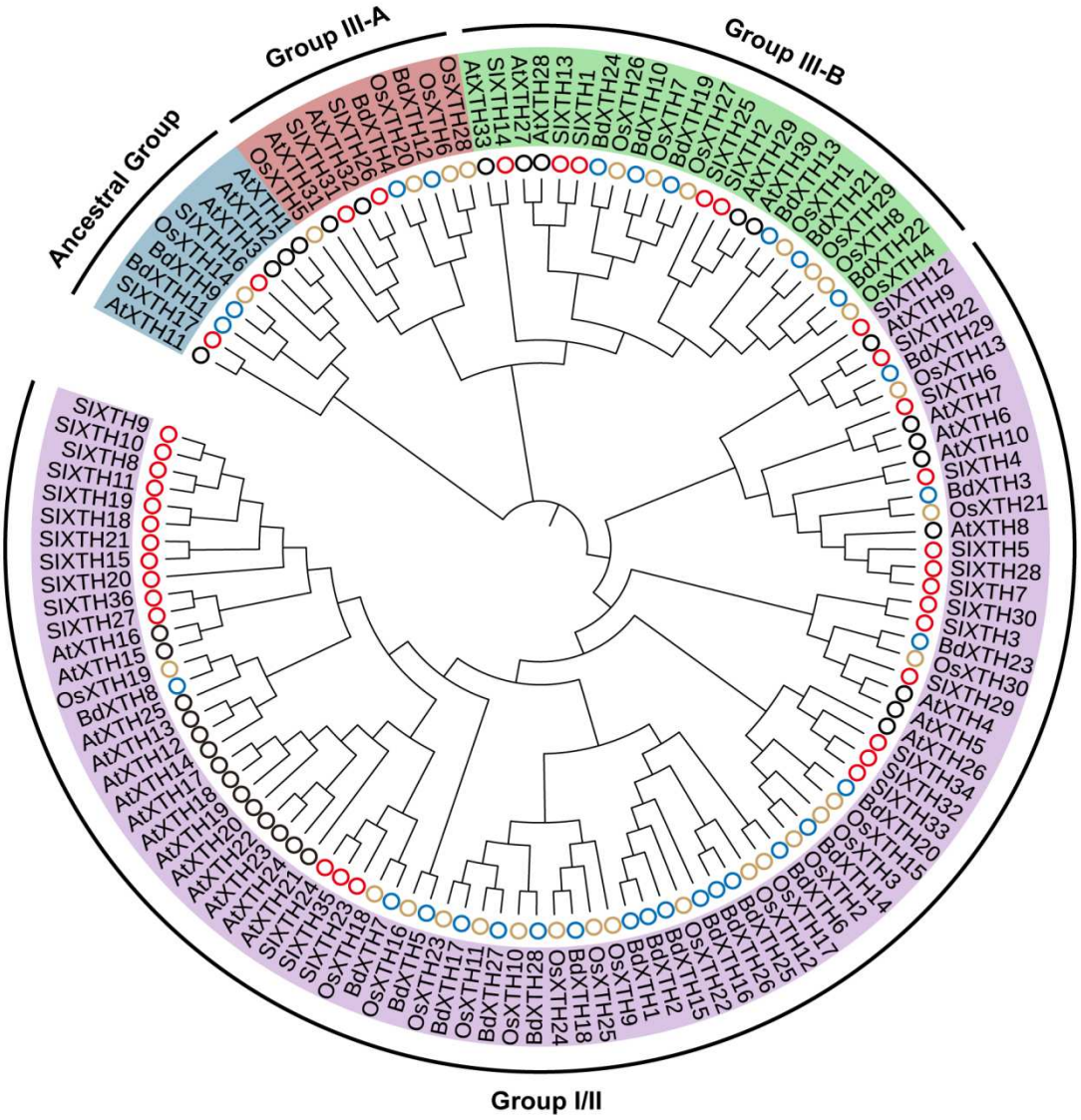


Fig. 2

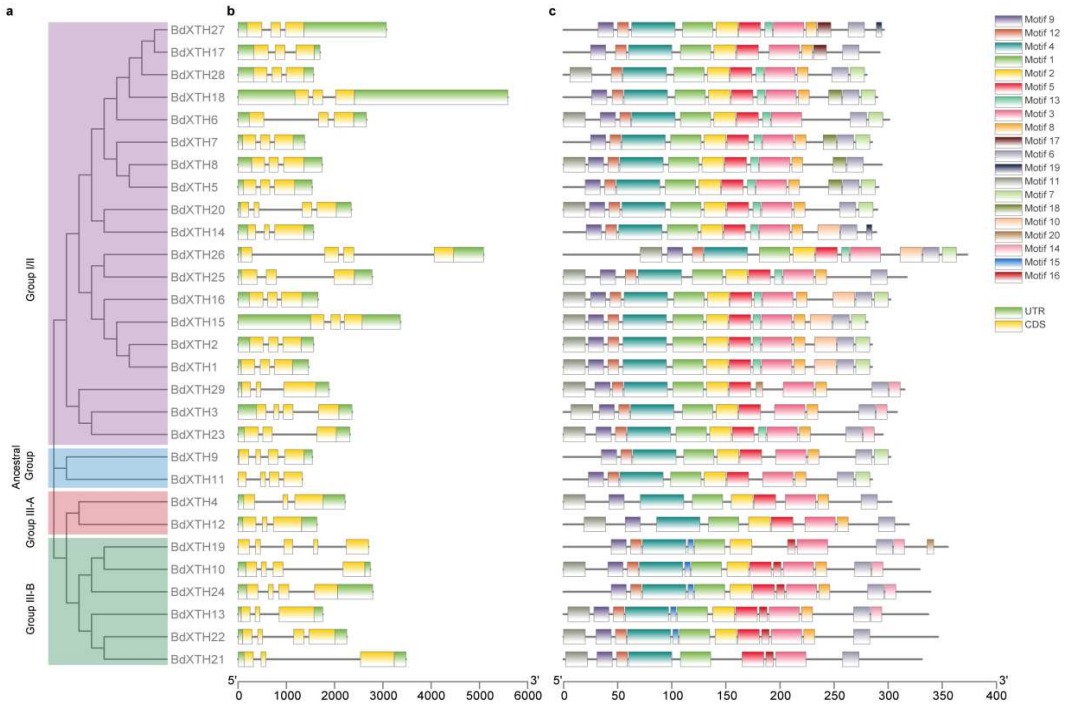

a
b
c
BdXTH27
BdXTH17
BdXTH28
BdXTH18
BdXTH6
BdXTH7
BdXTH8
BdXTH5
BdXTH20
BdXTH14
BdXTH26
BdXTH25
BdXTH16
BdXTH15
BdXTH2
BdXTH1
BdXTH29
BdXTH3
BdXTH23
BdXTH9
BdXTH11
BdXTH4
BdXTH12
BdXTH19
BdXTH10
BdXTH24
BdXTH13
BdXTH22
BdXTH21
Group I/II
Ancestral Group
Group III-A
Group III-B
5'
3'
0
1000
2000
3000
4000
5000
6000
50
100
150
200
250
300
350
400
Motif 9
Motif 12
Motif 4
Motif 1
Motif 2
Motif 5
Motif 13
Motif 3
Motif 8
Motif 17
Motif 6
Motif 19
Motif 11
Motif 7
Motif 18
Motif 10
Motif 20
Motif 14
Motif 15
Motif 16
UTR
CDS

Fig. 3

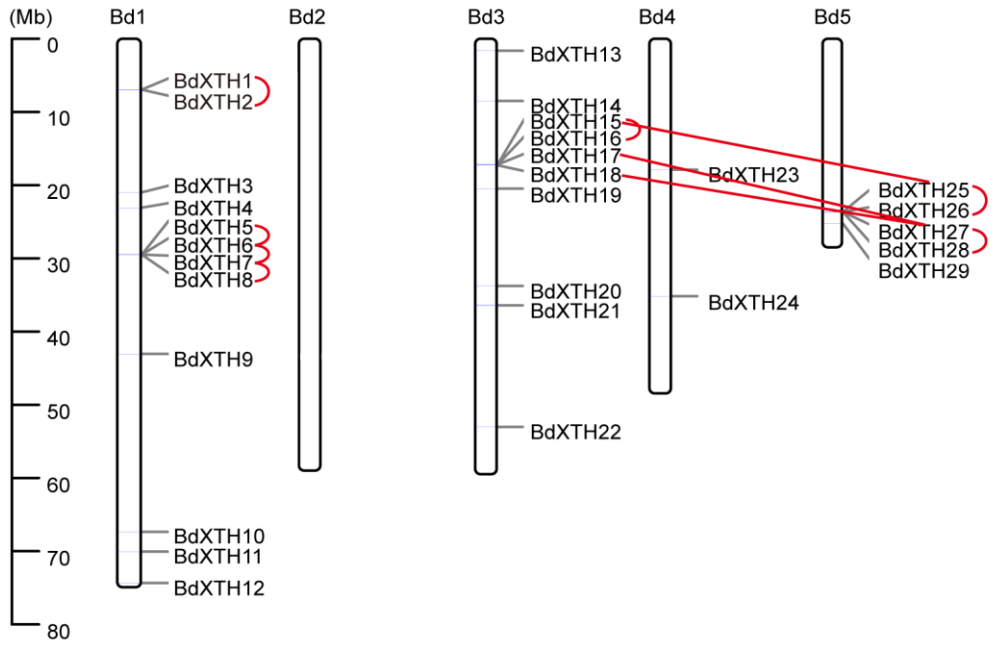

(Mb)
Bd1
Bd2
Bd3
Bd4
Bd5
0
10
20
30
40
50
60
70
80
BdXTH1
BdXTH2
BdXTH3
BdXTH4
BdXTH5
BdXTH6
BdXTH7
BdXTH8
BdXTH9
BdXTH10
BdXTH11
BdXTH12
BdXTH13
BdXTH14
BdXTH15
BdXTH16
BdXTH17
BdXTH18
BdXTH19
BdXTH20
BdXTH21
BdXTH22
BdXTH23
BdXTH24
BdXTH25
BdXTH26
BdXTH27
BdXTH28
BdXTH29

Fig. 4

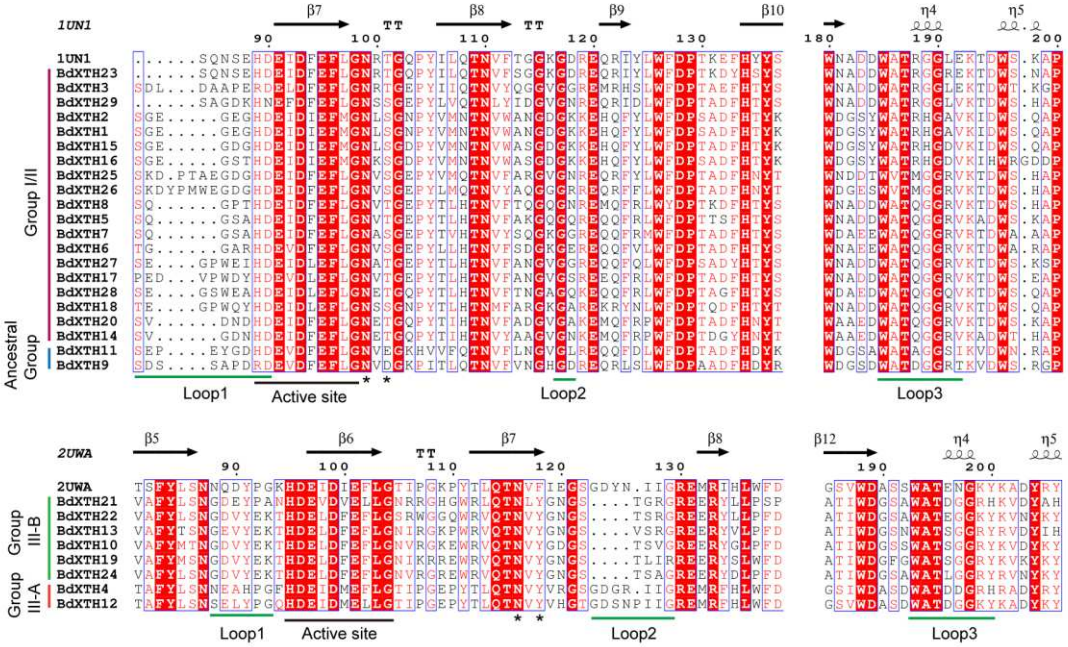
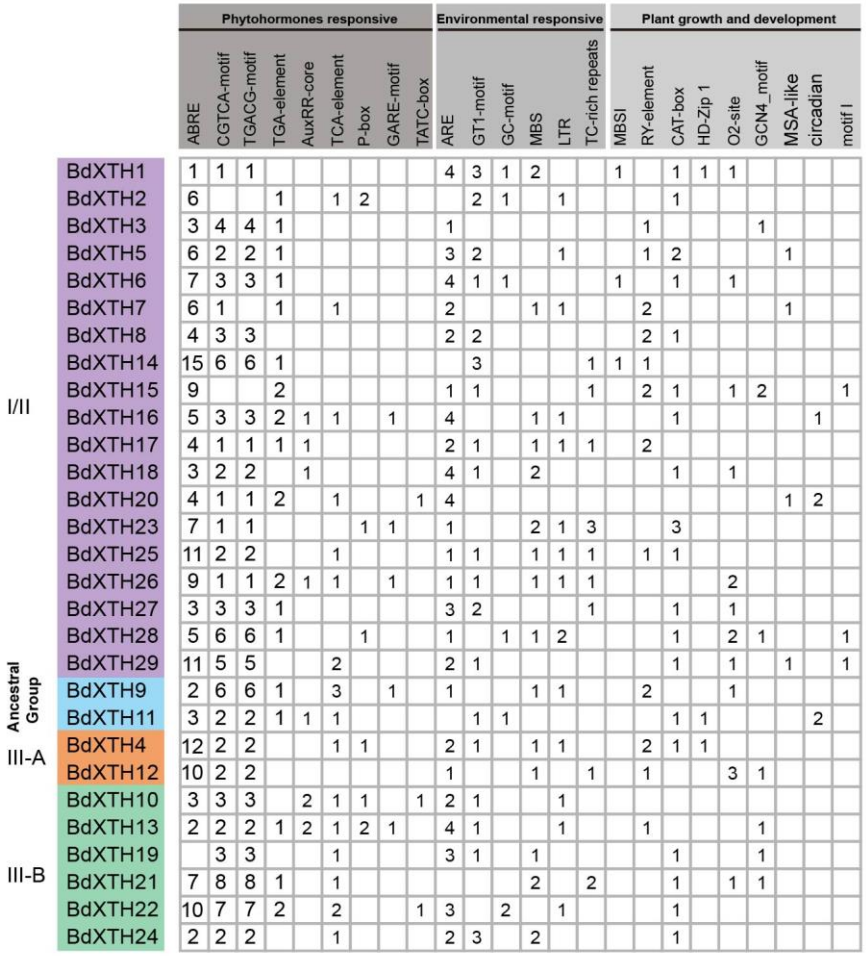

1UN1
β7
TT
β8
TT
β9
β10
η4
η5
90
100
110
120
130
180
190
200
1UN1
BdXTH23
BdXTH3
BdXTH29
BdXTH2
BdXTH1
BdXTH15
BdXTH16
BdXTH25
BdXTH26
BdXTH8
BdXTH5
BdXTH7
BdXTH6
BdXTH27
BdXTH17
BdXTH28
BdXTH18
BdXTH20
BdXTH14
BdXTH11
BdXTH9
Group I/II
Ancestral Group
Loop1
Active site
Loop2
Loop3
2UWA
β5
β6
TT
β7
β8
β12
η4
η5
90
100
110
120
130
190
200
2UWA
BdXTH21
BdXTH22
BdXTH13
BdXTH10
BdXTH19
BdXTH24
BdXTH4
BdXTH12
Group III-B
Group III-A
Loop1
Active site
Loop2
Loop3

Fig. 5



		Phytohormones responsive									Environmental responsive						Plant growth and development								
		ABRE	CGTCA-motif	TGACG-motif	TGA-element	AuxRR-core	TCA-element	P-box	GARE-motif	TATC-box	ARE	GT1-motif	GC-motif	MBS	LTR	TC-rich repeats	MBSI	RY-element	CAT-box	HD-Zip 1	O2-site	GCN4_motif	MSA-like	circadian	motif I
I/II	BdXTH1	1	1	1							4	3	1	2			1		1	1	1				
	BdXTH2	6			1		1	2				2	1		1				1						
	BdXTH3	3	4	4	1						1							1				1			
	BdXTH5	6	2	2	1						3	2			1			1	2				1		
	BdXTH6	7	3	3	1						4	1	1				1		1		1				
	BdXTH7	6	1		1		1				2			1	1			2					1		
	BdXTH8	4	3	3							2	2						2	1						
	BdXTH14	15	6	6	1							3				1	1	1							
	BdXTH15	9			2						1	1				1		2	1		1	2			1
	BdXTH16	5	3	3	2	1	1		1		4			1	1				1					1	
	BdXTH17	4	1	1	1	1					2	1		1	1	1		2							
	BdXTH18	3	2	2		1					4	1		2					1		1				
	BdXTH20	4	1	1	2		1			1	4												1	2	
	BdXTH23	7	1	1				1	1		1			2	1	3			3						
	BdXTH25	11	2	2			1				1	1		1	1	1		1	1						
	BdXTH26	9	1	1	2	1	1		1		1	1		1	1	1					2				
	BdXTH27	3	3	3	1						3	2				1			1		1				
	BdXTH28	5	6	6	1			1			1		1	1	2				1		2	1			1
	BdXTH29	11	5	5			2				2	1							1		1		1		1
Ancestral Group	BdXTH9	2	6	6	1		3		1		1			1	1			2			1				
	BdXTH11	3	2	2	1	1	1					1	1						1	1				2	
III-A	BdXTH4	12	2	2			1	1			2	1		1	1			2	1	1					
	BdXTH12	10	2	2							1			1		1		1			3	1			
III-B	BdXTH10	3	3	3		2	1	1		1	2	1			1										
	BdXTH13	2	2	2	1	2	1	2	1		4	1			1			1				1			
	BdXTH19		3	3			1				3	1		1					1			1			
	BdXTH21	7	8	8	1		1							2		2			1		1	1			
	BdXTH22	10	7	7	2		2			1	3		2		1				1						
	BdXTH24	2	2	2			1				2	3		2					1						

Fig. 6

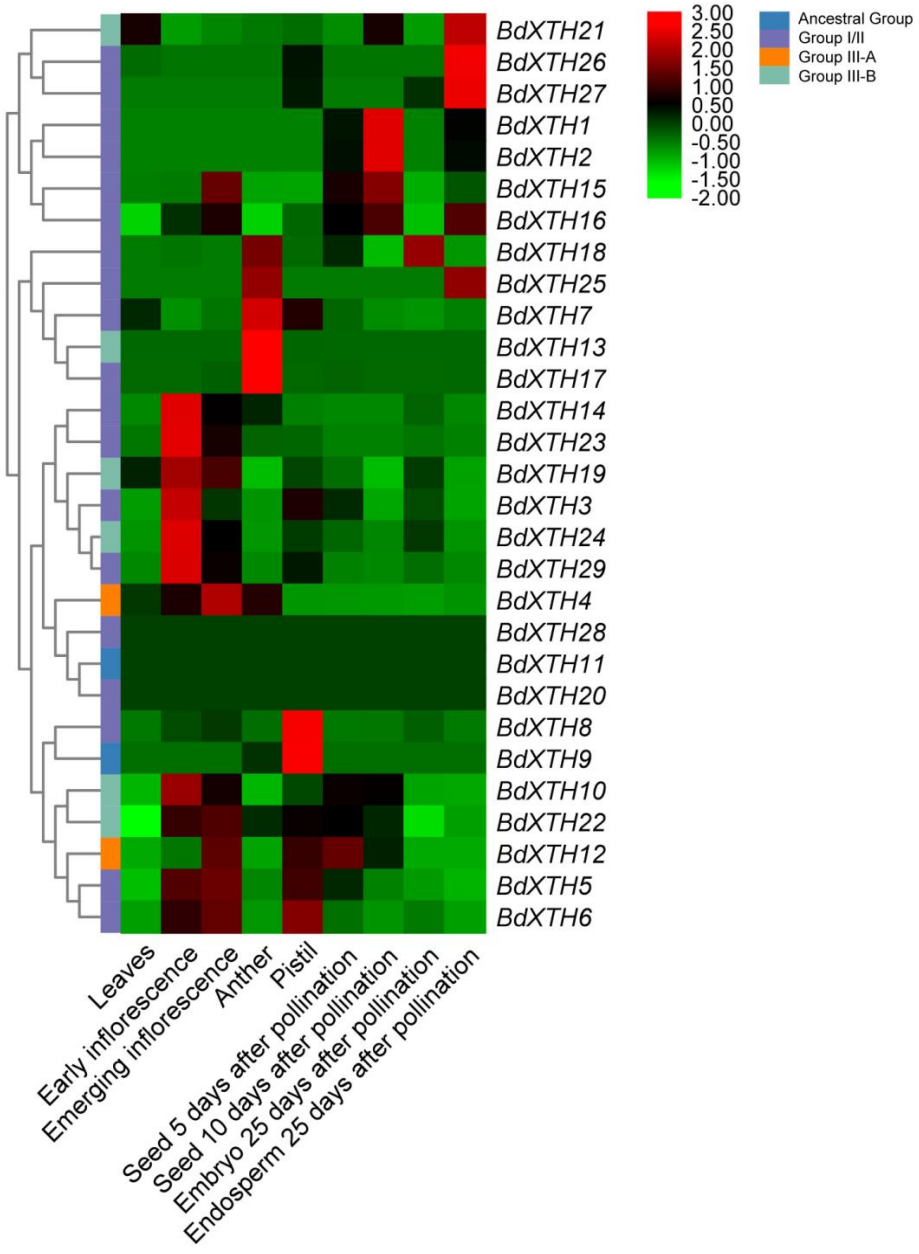

3.00
2.50
2.00
1.50
1.00
0.50
0.00
-0.50
-1.00
-1.50
-2.00
Ancestral Group
Group I/II
Group III-A
Group III-B
BdXTH21
BdXTH26
BdXTH27
BdXTH1
BdXTH2
BdXTH15
BdXTH16
BdXTH18
BdXTH25
BdXTH7
BdXTH13
BdXTH17
BdXTH14
BdXTH23
BdXTH19
BdXTH3
BdXTH24
BdXTH29
BdXTH4
BdXTH28
BdXTH11
BdXTH20
BdXTH8
BdXTH9
BdXTH10
BdXTH22
BdXTH12
BdXTH5
BdXTH6
Leaves
Early inflorescence
Emerging inflorescence
Anther
Pistil
Seed 5 days after pollination
Seed 10 days after pollination
Embryo 25 days after pollination
Endosperm 25 days after pollination

Fig. 7

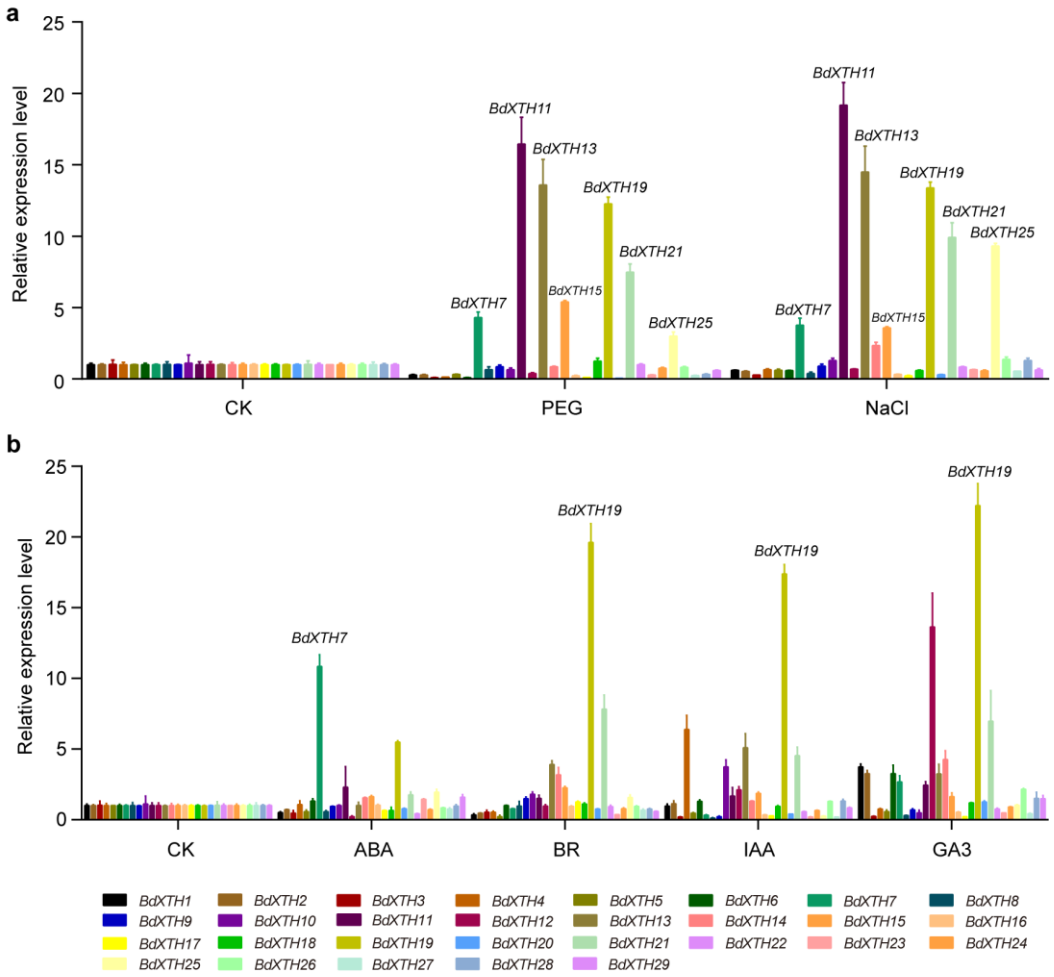
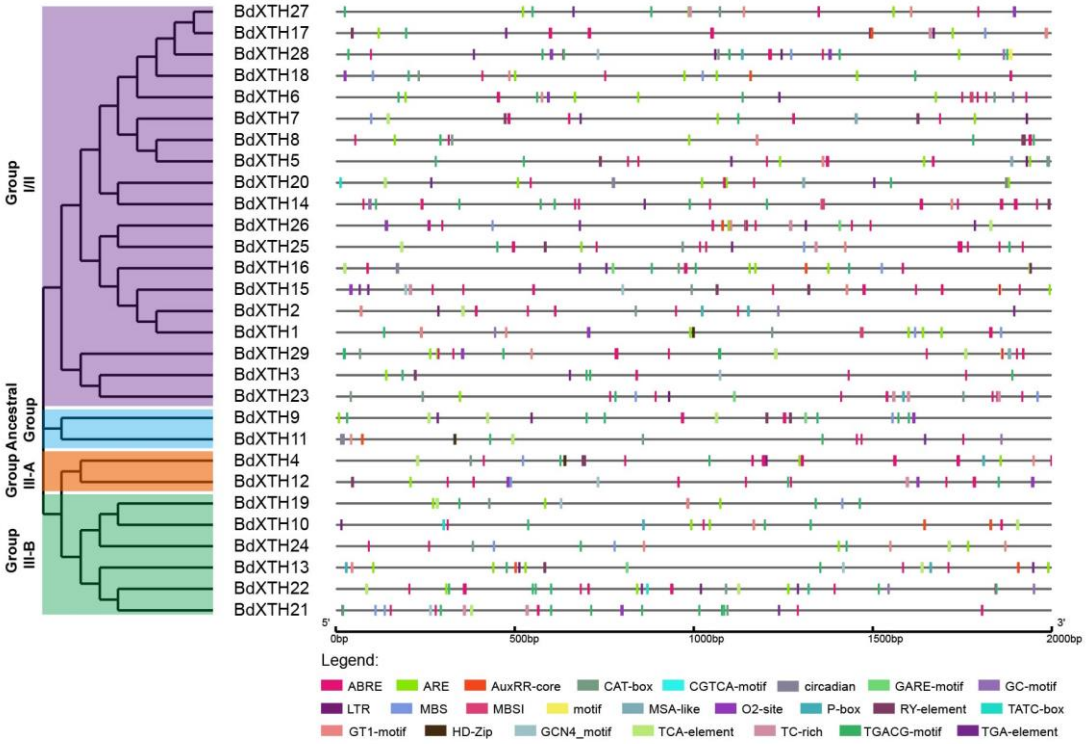

a
Relative expression level
25
20
15
10
5
0
BdXTH7
BdXTH11
BdXTH13
BdXTH15
BdXTH19
BdXTH21
BdXTH25
CK
PEG
NaCl
b
Relative expression level
BdXTH7
BdXTH19
BdXTH19
BdXTH19
CK
ABA
BR
IAA
GA3
BdXTH1
BdXTH2
BdXTH3
BdXTH4
BdXTH5
BdXTH6
BdXTH7
BdXTH8
BdXTH9
BdXTH10
BdXTH11
BdXTH12
BdXTH13
BdXTH14
BdXTH15
BdXTH16
BdXTH17
BdXTH18
BdXTH19
BdXTH20
BdXTH21
BdXTH22
BdXTH23
BdXTH24
BdXTH25
BdXTH26
BdXTH27
BdXTH28
BdXTH29

Fig. S1


Motif 1
Motif 2
Motif 3
Motif 4
Motif 5
Motif 6
Motif 7
Motif 8
Motif 9
Motif 10
Motif 11
Motif 12
Motif 13
Motif 14
Motif 15
Motif 16
Motif 17
Motif 18
Motif 19
Motif 20

Fig. S4


Group I/II
Group Ancestral Group
Group III-A
Group III-B
BdXTH27
BdXTH17
BdXTH28
BdXTH18
BdXTH6
BdXTH7
BdXTH8
BdXTH5
BdXTH20
BdXTH14
BdXTH26
BdXTH25
BdXTH16
BdXTH15
BdXTH2
BdXTH1
BdXTH29
BdXTH3
BdXTH23
BdXTH9
BdXTH11
BdXTH4
BdXTH12
BdXTH19
BdXTH10
BdXTH24
BdXTH13
BdXTH22
BdXTH21
5'
3'
0bp
500bp
1000bp
1500bp
2000bp
Legend:
ABRE
ARE
AuxRR-core
CAT-box
CGTCA-motif
circadian
GARE-motif
GC-motif
LTR
MBS
MBSI
motif
MSA-like
O2-site
P-box
RY-element
TATC-box
GT1-motif
HD-Zip
GCN4_motif
TCA-element
TC-rich
TGACG-motif
TGA-element

960 **Table S2 Oligonucleotide primers used for qRT-PCR assays in this study.**
961

Primer name	Sequence
Bradi1g09690.1F	CCGAGGTTTACCTGCGTTCT
Bradi1g09690.1R	CCTCTCCGTAAGACTGTCGC
Bradi1g09700.1F	CCGAGGTTTACCTGCGTTCT
Bradi1g09700.1R	CCTCTCCGTAAGACTGTCGC
Bradi1g25847.1F	GCGGTGTAACTATCAGAGA
Bradi1g25847.1R	CTACCCATGCTCGTGTAT
Bradi1g27867.1F	TTCGTTTCACAAGCATAGC
Bradi1g27867.1R	AGTAGCATTGAGAACATCCA
Bradi1g33810.1F	AGCTCAGCGACATGAGCTAC
Bradi1g33810.1R	AGTTTCTCTCAACGGCGGAG
Bradi1g33817.1F	CGTGGCAGTGTAAGAATC
Bradi1g33817.1R	AAACGAGTGGAGAAACCT
Bradi1g33827.1F	GAAGCCCGTGGATGTACCAG
Bradi1g33827.1R	GTGTCGGCGCAGTAGTTGTA
Bradi1g33840.1F	ATGCCTCGCTCATTCCACTC
Bradi1g33840.1R	TCGCTATTATCGCCGACACC
Bradi1g44777.1F	CCAGACCAACGTCTTCGTCA
Bradi1g44777.1R	CCTGTAGTCGTGGAAGTCGG
Bradi1g68590.1F	AAGGAAGGAACGGAACGGAC
Bradi1g68590.1R	GCTGCTGCTTCGCTAAACTC
Bradi1g71937.1F	GAGGATGAAGATACCCGGCG
Bradi1g71937.1R	AAAACGACGTGCTTGCCTTC
Bradi1g77990.1F	ACGGATCGCTCCAAGAATCG
Bradi1g77990.1R	CGTTAATCAATCGCCGCTGG
Bradi3g02700.1F	TGAGTTCCTCGGCAACATCC
Bradi3g02700.1R	GTCCACAGGATGGAGTAGCG
Bradi3g10290.1F	ATCCGGCCAGTTAAGCGAAA

Bradi3g10290.1R	GCCGGACACAAGTCCGATAA
Bradi3g18590.1F	GGCAGTTGGCACGTTAGTTG
Bradi3g18590.1R	CATCCAAATCGCCTCCCCAT
Bradi3g18600.1F	TTTTGCCACCGTCGTTGTTC
Bradi3g18600.1R	TCCTTGACACAGCAGCTACG
Bradi3g18607.1F	GCGCGAGACAAGCACATGAT
Bradi3g18607.1R	CAGTCCGAGGTGCATTCCTT
Bradi3g18690.1F	AGTGAACAACGGGACAACCA
Bradi3g18690.1R	CAGAACACCTACGGCCTCAC
Bradi3g21337.2F	GGTTACTGCTACGACCGTCT
Bradi3g21337.2R	CTGATCCCTGGACCTCGCTA
Bradi3g31767.1F	GTCCGATGTACGGGTTCTCC
Bradi3g31767.1R	GGCCTTGCTCCAATCAGTCT
Bradi3g34227.1F	TACCGCCATGGCTTCTTCAG
Bradi3g34227.1R	GACAGGTAGAAGGCGACGAC
Bradi3g52307.1F	GACGCCGGATGATAGGACTG
Bradi3g52307.1R	GCCGTGGAGGTAGTAATCCG
Bradi4g16990.1F	CGCGCCCAATTAATCCAGTG
Bradi4g16990.1R	TACGTACGCCAATCGACCAG
Bradi4g29707.1F	GATGTGATCCCGCCATGACT
Bradi4g29707.1R	AAATTCCGACCGGTCTCCAC
Bradi5g20718.1F	GAGTTCCTGGGGAACGTCAG
Bradi5g20718.1R	AACCACAGGTAGAACCGCTG
Bradi5g20726.1F	TTAACGCGTGCAGAGGCTAT
Bradi5g20726.1R	GGACAGACACACAGAGCGAA
Bradi5g20734.1F	CAAACGAGGTTCTCCCTCCC
Bradi5g20734.1R	AAATGTGATATGCGCGTGGC
Bradi5g20742.1F	AAGTCAAGACGGACTGGTCG

Bradi5g20742.1R	CACGTACTCCTGCCCGTATC
Bradi5g22907.1F	GGCGTCGTCACTGCTTTCTA
Bradi5g22907.1R	GTACAGGTTCGTCTGCACCA
BdUBC18-F	TGGAGGCACCTCAGGTCATTTTC
BdUBC18-R	GTTGCTTTGCTGGCGAGCTAGAC

962
963

Figures

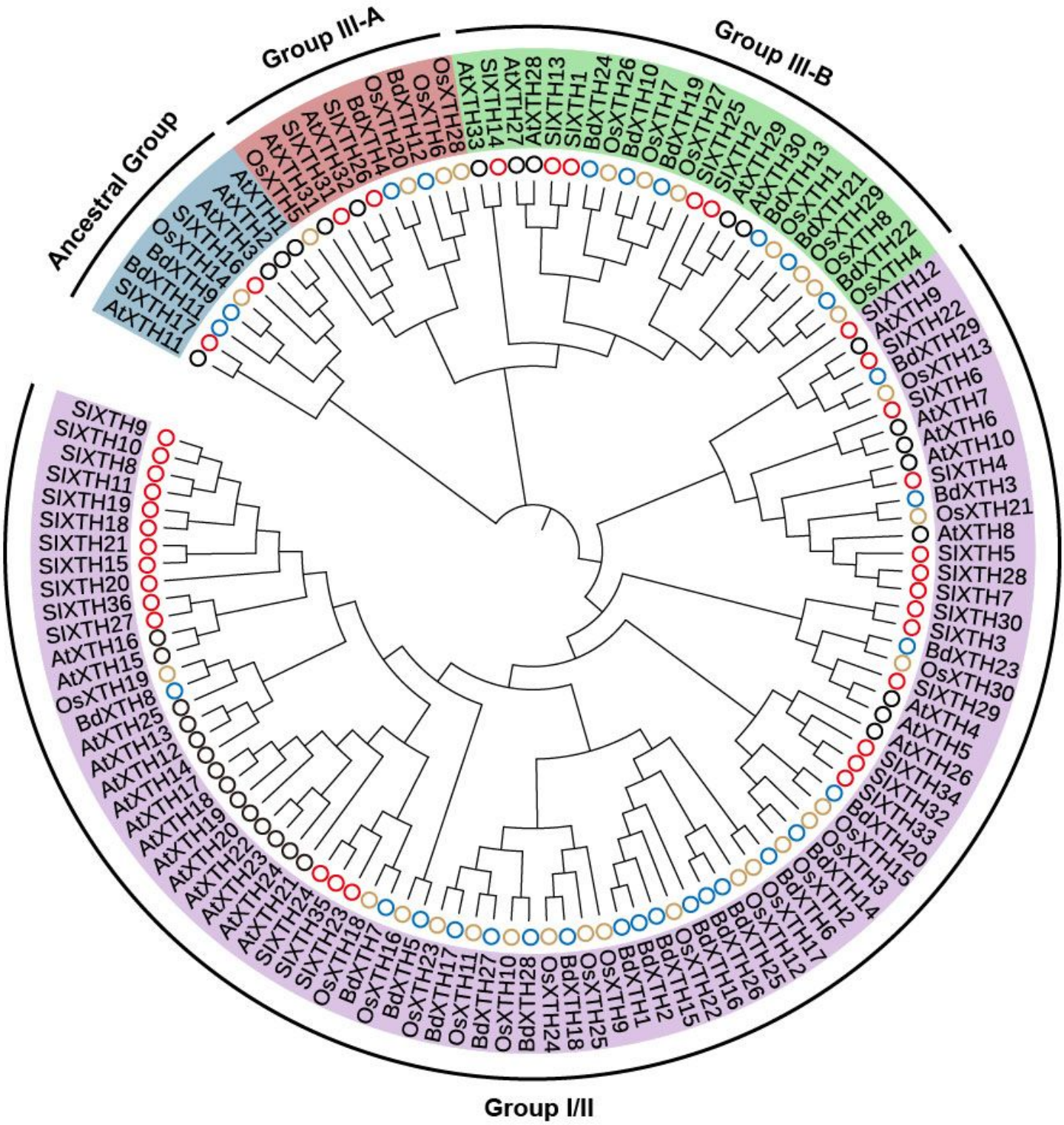


Figure 1

Phyloge netic relationships among XTH proteins from Brachypodium and three other plant species. The phylogenetic tree was constructed with the Neighbor Joining (NJ) method as implemented in MEGA X software, and branch confidence was estimated by bootstrapping with 1,000 replicates. The open blue,

gray, red, and black circles indicate proteins from Brachypodium , rice, tomato, and Arabidopsis , respectively. In addition, the XTH proteins are classified into three clades (Group I/II, Group III, and the Ancestral Group) Group). Proteins in Group I/II and the Ancestral Group are shown with purple and blue backgrounds, respectively. Group III is further divided into two subclades, Group III A and Group III B, which are indicated with red and green backgrounds, respectively.

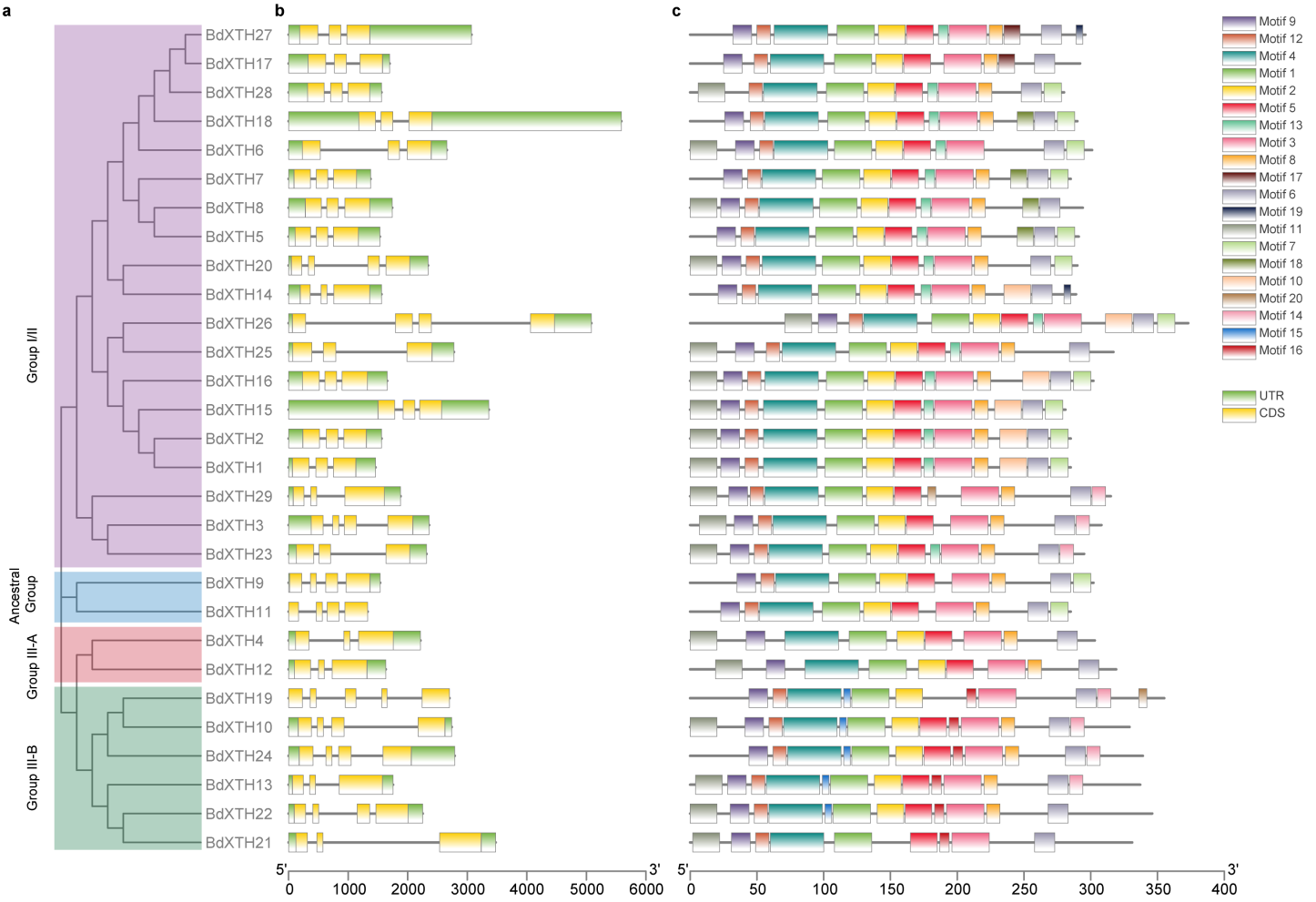


Figure 2

Unrooted neighbor joining phylogenetic tree, conserved protein motifs, and structural analysis of BdXTH genes. a) Phylogenetic relationships of the XTH proteins in Brachypodium . Proteins from the four clades (Group I/II, Group III A and Group III B, and the Ancestral Group) are color coded as in Figure 1. (b) The structures of the 29 putative BdXTH genes. The UTRs, exons, and introns are represented by green boxes, yellow boxes, and black lines, respectively. (c) Conserved motif analysis of the BdXTH proteins. The different motifs are indicated by different colored boxes numbered motif 1 to motif 20. The structural features of the 20 motifs are shown in Fig S1.

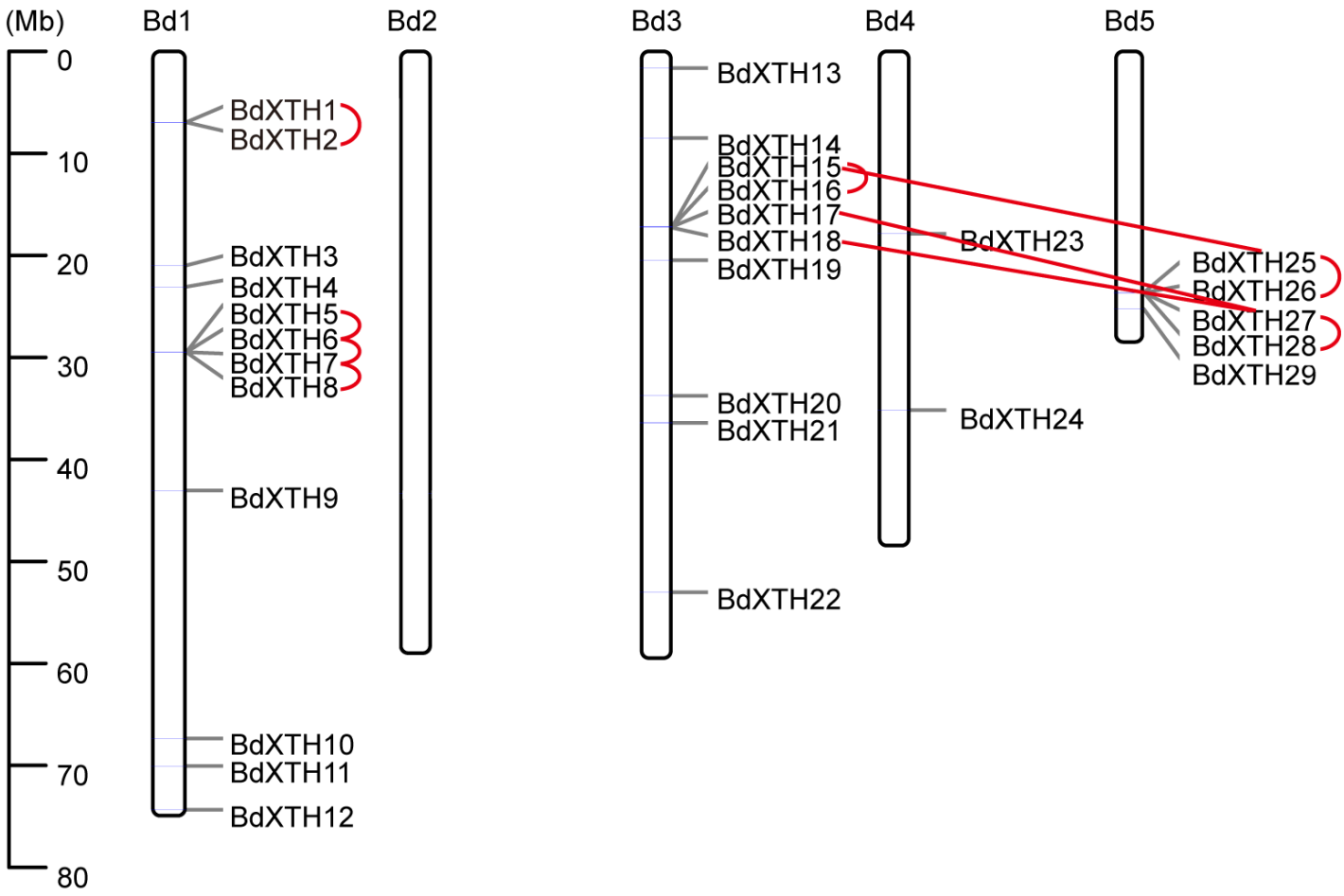


Figure 3

The physical locations of BdXTH genes on the five Brachypodium chromosomes. Tandemly duplicated gene pairs and segmentally duplicated genes are linked by red lines. The chromosome numbers are displayed at the top of each chromosome and the scale in megabases (Mb) is shown on the left.

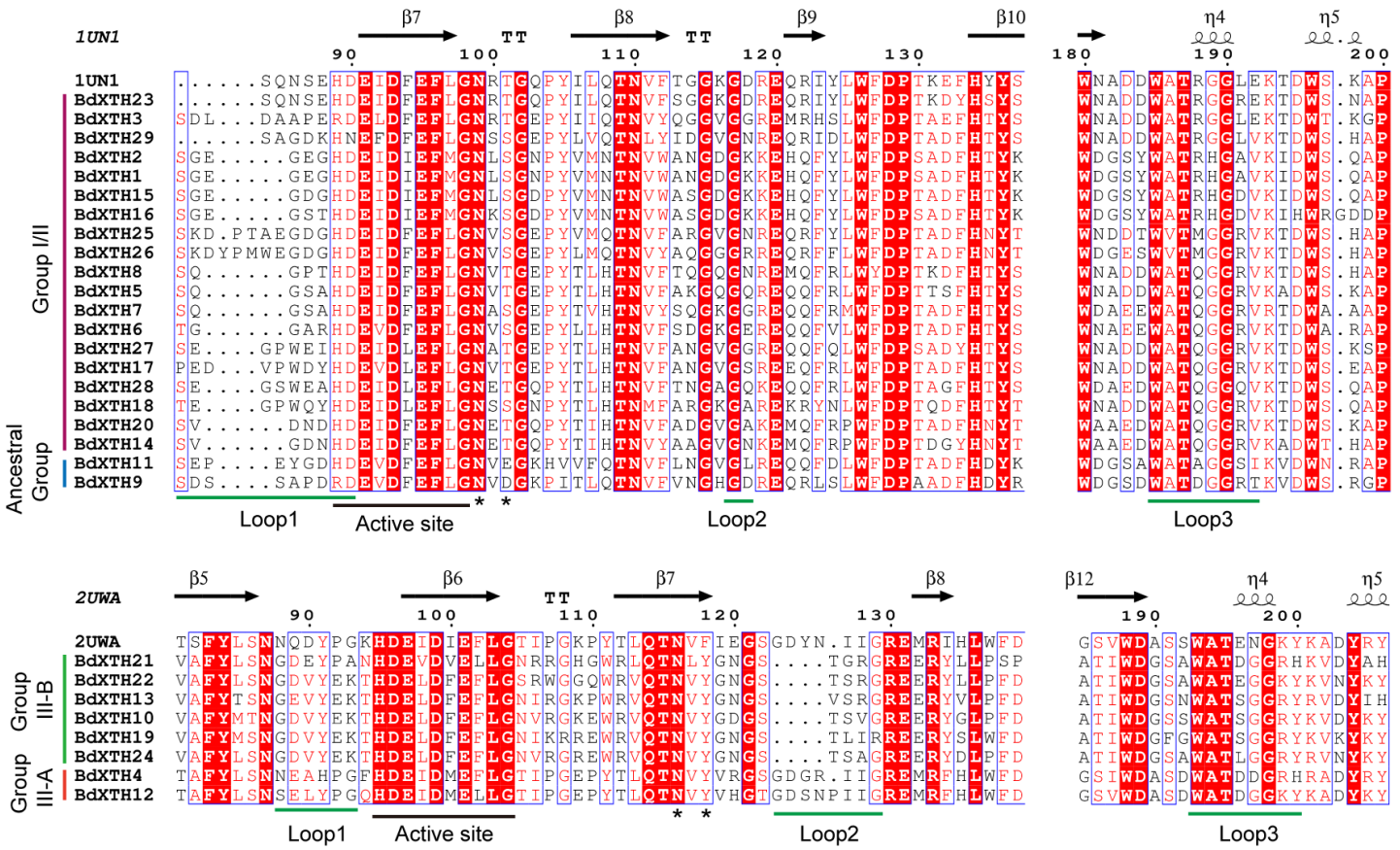


Figure 4

Structure based sequence alignment of BdXTH proteins. The structures of two proteins (PttXET16 34, PDB id: 1UN1; TmNXG1, PDB id: 2UWA) have been experimentally determined. Proteins in Group I/II and the Ancestral Group had similar structures to 1UN1, and proteins in Group III show similar structures to 2UW A. The active site (E x D x E), and loops 1, 2, and 3 are underlined in black and green , respectively . The N glycosylation site residues are indicated by asterisks.

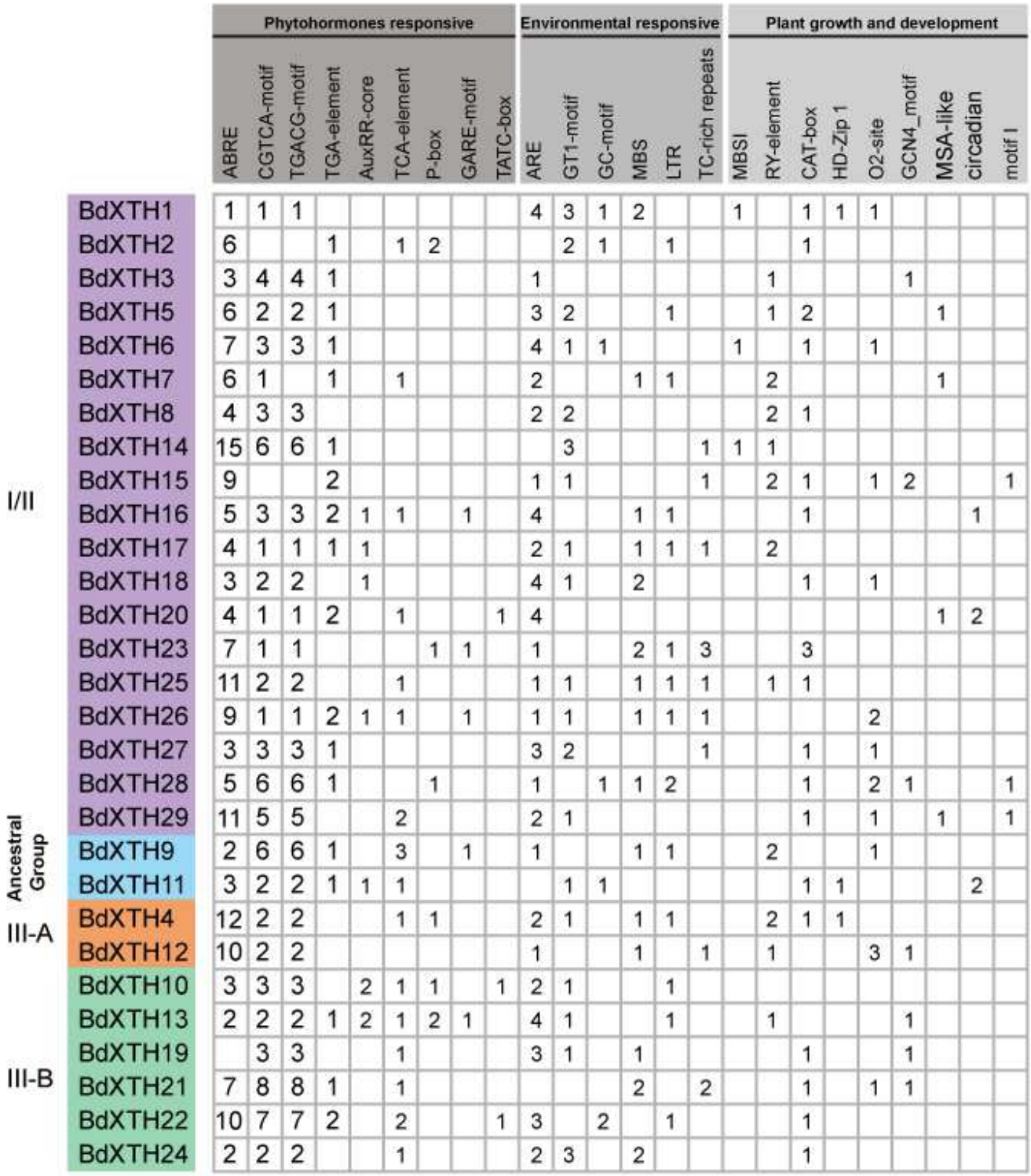


Figure 5

Numbers of cis acting elements in the promoter regions of the 29 BdXTH genes. Three types of cis acting elements in the 2,000 bp of DNA sequence upstream of the promoter regions are shown in the figure, including phytohormone and environmentally responsive elements and plant growth and development related elements. Members of the different element classes are shown at the top of the figure in different shades of gray

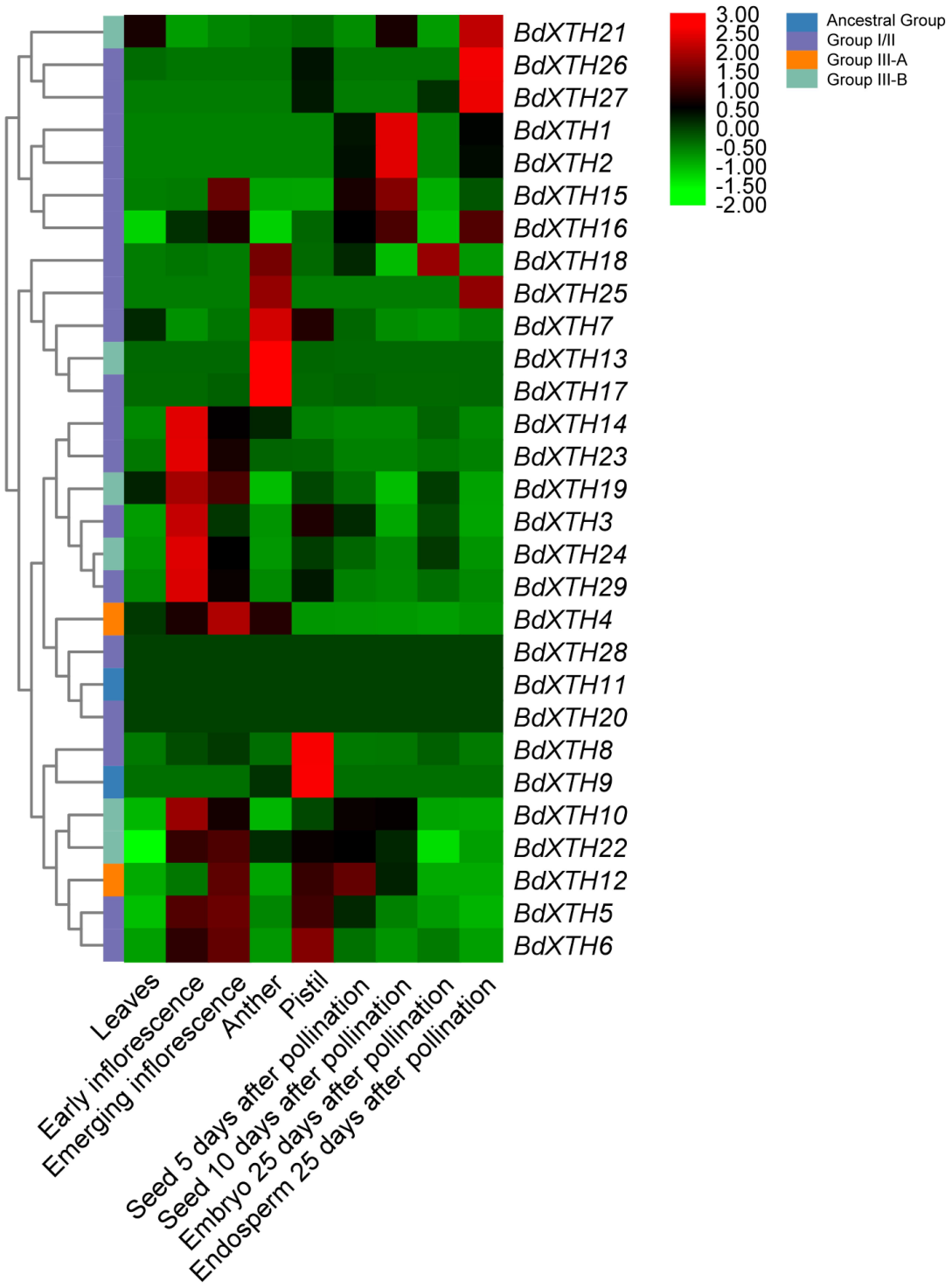


Figure 6

Heat map showing the expression pattern of BdXTH genes in Brachypodium . Expression profiles from various tissues and developmental stages (leaves, early inflorescences, e merging inflorescences, anthers, pistils, seeds at 5 days after pollination, seeds at 10 days after pollination, embryos, and endosperm) were downloaded from the NCBI database (SRP008505). The relative expression levels are represented by the colored bars. Red and green boxes indicate high and low expression levels, respectively.

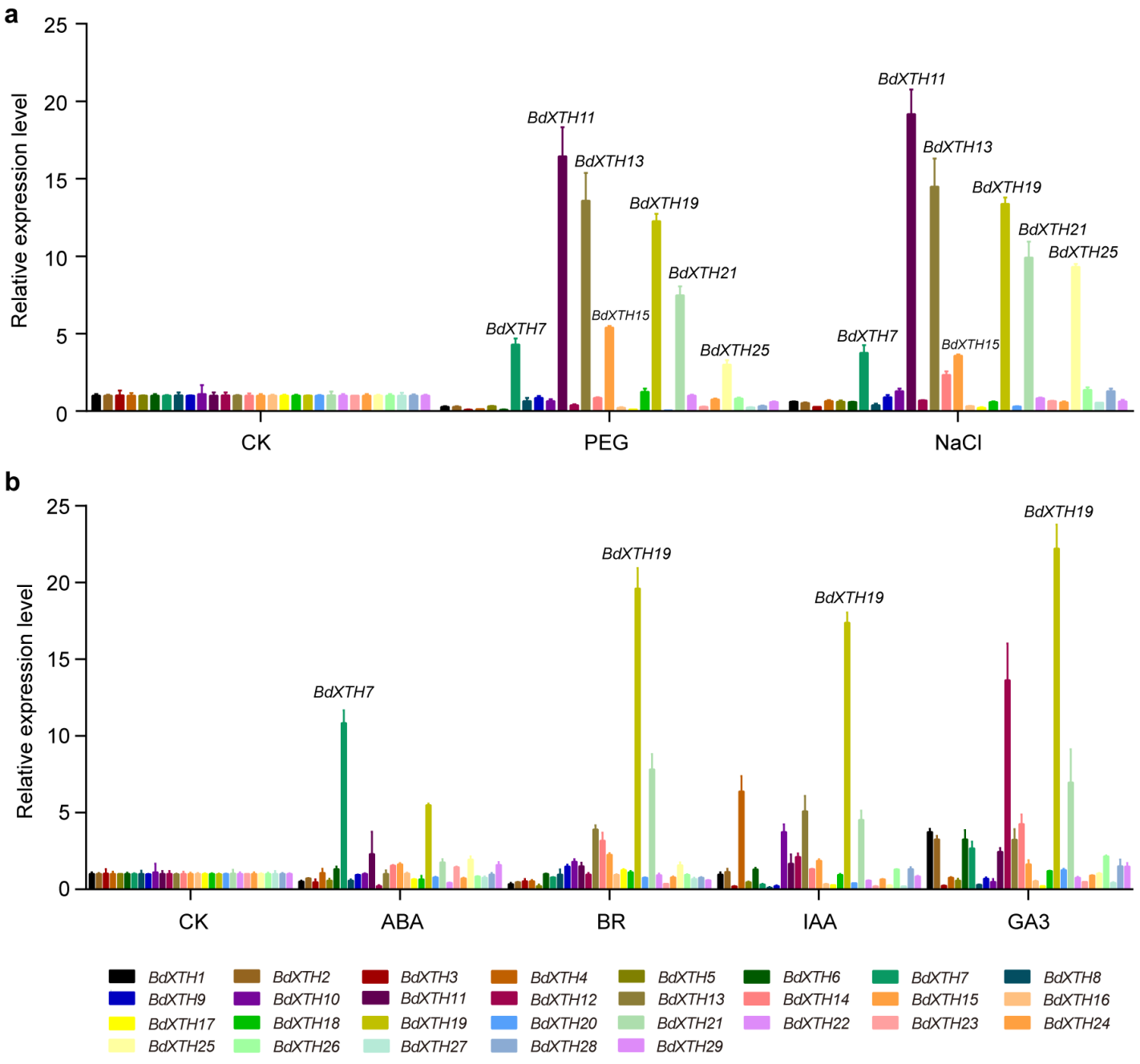


Figure 7

Expression analysis of BdXTH genes under different conditions. Quantitative real time polymerase chain reaction (qRT PCR) analysis of BdXTH gene expression in response to ab iotic stresses (drought and salinity) (a), and phytohormone treatments (ABA, BR, IAA, and GA3) (b). The means ± SD of three biological replicates are presented.

Supplementary Files

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

- Fig.S1.jpg
- Fig.S2.pdf
- Fig.S3.pdf
- Fig.S4.tif
- TableS1.xlsx
- TableS2.xlsx