Theor Appl Genet (2008) 116:373–381
DOI 10.1007/s00122-007-0675-1

ORIGINAL PAPER

Structures of the three homoeologous loci of wheat benzoxazinone biosynthetic genes *TaBx3* and *TaBx4* and characterization of their promoter sequences

Taiji Nomura · Shuhei Nasuda · Kanako Kawaura · Yasunari Ogihara · Nobuhiko Kato · Fumihiko Sato · Toshio Kojima · Atsushi Toyoda · Hajime Iwamura · Takashi R. Endo

Received: 14 May 2007 / Accepted: 9 November 2007 / Published online: 27 November 2007

Abstract Common wheat ($2n = 6x = 42$, genome formula AABBDD) accumulates benzoxazinones (Bxs) as defensive compounds. There are five Bx biosynthetic genes (*TaBx1–TaBx5*), and their homoeologous alleles are located on all three homoeologous chromosomes of the A, B and D genomes. Here the molecular structures of the *TaBx3* and *TaBx4* loci, both of which are located on chromosomes 5A, 5B and 5D, were revealed by sequencing transformation-competent artificial chromosome (TAC) clones. In all homoeologous chromosomes, *TaBx3* existed downstream of *TaBx4* in a tail-to-head manner, and the two genes were separated from each other by 9.0 kb in 5A, 7.3 kb in 5B and 11.3 kb in 5D. Among the three homoeologs of *TaBx3* and *TaBx4*, the promoter sequences were less conserved than the coding sequences. The promoter sequences of *TaBx3* and *TaBx4* were highly similar to those of their respective orthologs in the diploid progenitors of common wheat, but were not similar to those of the maize orthologs. Sequence similarity was found between the *TaBx3* and *TaBx4* coding sequences, but not between their promoter sequences despite their similar transcription pattern at the seedling stage. Some putative *cis*-elements were found to be shared by all *TaBx3* and *TaBx4* promoter regions. These results imply that stage-specific transcription of *TaBx3* and *TaBx4* is not controlled by global sequence similarity of their promoters but by some essential *cis*-elements. The promoter activity measured by transient assays in wheat protoplasts was similar among the three homoeologs of *TaBx3* and *TaBx4* in spite of their differential transcript levels in wheat seedlings.

Communicated by E. Guiderdoni.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-007-0675-1) contains supplementary material, which is available to authorized users.

T. Nomura · S. Nasuda · T. R. Endo
Division of Applied Biosciences, Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan

K. Kawaura · Y. Ogihara
Kihara Institute for Biological Research, Yokohama City University, Yokohama 244-0813, Japan

N. Kato · F. Sato
Division of Integrated Life Science, Graduate School of Biostudies, Kyoto University, Kyoto 606-8502, Japan

T. Kojima · A. Toyoda
RIKEN Genomic Sciences Center, Yokohama 230-0045, Japan

H. Iwamura
Department of Bio-Technology, School of Biology-Oriented Science and Technology, Kinki University, Kinokawa, Wakayama 649-6493, Japan

T. Nomura (✉)
Donald Danforth Plant Science Center, 975 North Warson Road, St Louis, MO 63132, USA
e-mail: tnomura@danforthcenter.org

Introduction

Benzoxazinones (Bxs) are major secondary metabolites in gramineous plants, including common wheat (*Triticum aestivum*), rye (*Secale cereale*) and maize (*Zea mays*). They are considered to play important roles in the chemical defense in plants against insect pests and pathogens (Niemeyer 1988; Sicker et al. 2000). Representative Bxs are 2,4-dihydroxy-1,4-benzoxazin-3-one (DIBOA) and its 7-methoxy derivative (DIMBOA), and they are stored in vacuoles in the form of 2-*O*-*β*-glucosides, DIBOA-Glc and DIMBOA-Glc.

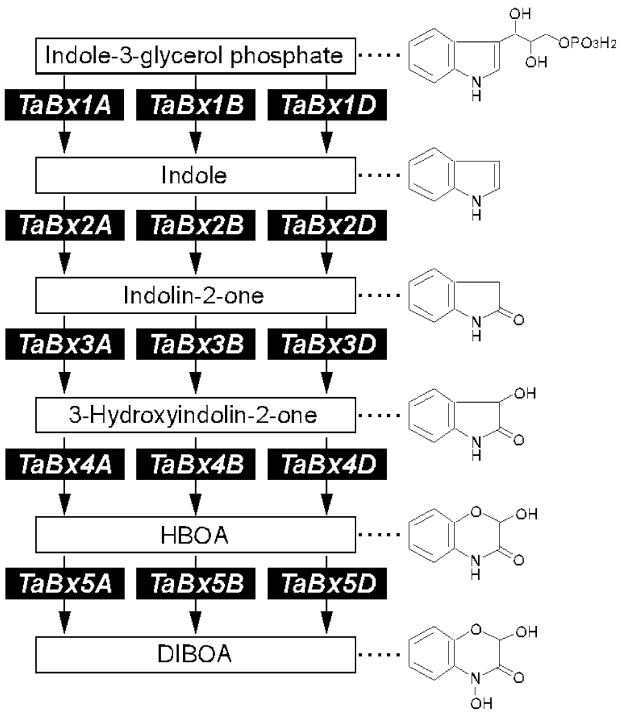

Fig. 1 Bx biosynthetic pathway showing the enzymatic reactions controlled by the *TaBx1–TaBx5* genes in the A, B and D genomes of common wheat

The biosynthetic pathway of DIBOA branches off from the tryptophan pathway at indole-3-glycerol phosphate (Frey et al. 1997, 2000; Melanson et al. 1997). Five sequential reactions from indole-3-glycerol phosphate lead to the formation of DIBOA (Fig. 1). The genes involved have been isolated in maize (*ZmBx1–ZmBx5* genes; Frey et al. 1995, 1997), wild barley, *Hordeum lechleri*, (*HlBx1–HlBx5*; Grün et al. 2005), and common wheat (*TaBx1A–TaBx5A*, *TaBx1B–TaBx5B*, and *TaBx1D–TaBx5D*; Nomura et al. 2002, 2003, 2005; Fig. 1). Of the five biosynthetic genes, the *Bx2–Bx5* genes encode cytochrome P450 monooxygenases belonging to a CYP71C subfamily.

Common wheat ($2n = 6x = 42$, genome formula AABBDD) is an allohexaploid that was formed through successive chromosome doubling of hybrids involving three ancestral diploid species ($2n = 14$), *T. urartu* (AA), *Aegilops speltoides* (SS ≈ BB), and *Ae. tauschii* (DD) (Huang et al. 2002b; Feldman and Levy 2005). Nomura et al. (2002, 2003) demonstrated that the five Bx biosynthetic genes are all present in the three diploid progenitors. It has been demonstrated that Bxs increase in amount to a maximum soon after germination and then decrease to a constant level in wheat (Nakagawa et al. 1995), maize (Ebisui et al. 1998) and rye (Sue et al. 2000). A similar pattern of Bx increase and decrease was observed in the three diploid progenitors of common wheat (Nomura et al. 2005). Although the transcript levels of each homoeolog of the *TaBx1–TaBx5* genes differ among the three genomes of common wheat and among the three diploid species, the highest transcription was observed at the seedling stage in all of them, which fully accorded with the Bx content changes (Nomura et al. 2005). These facts led us to hypothesize that the high transient level of transcription of the Bx biosynthetic genes in the juvenile stage of growth is controlled by a transcriptional mechanism that is common to the five Bx biosynthetic genes in all three genomes. In fact, it has been demonstrated that genes involved in a particular metabolic pathway are co-regulated (Gachon et al. 2005; Kato et al. 2007), and that functionally related genes are clustered on a chromosome, as well as being co-expressed (Ren et al. 2005; Zhan et al. 2006). In maize, the *ZmBx1–ZmBx5* genes are clustered on the short arm of chromosome 4 (Frey et al. 1995, 1997). In common wheat, the five *TaBx* genes are separately located on the homoeologous group-4 and group-5 chromosomes, and *TaBx1* and *TaBx2* exist in the same chromosomal bin on the group-4 chromosomes, while *TaBx3-TaBx5* exist in the same chromosomal bin on the group-5 chromosomes (Nomura et al. 2003).

The Bx biosynthetic genes have been isolated from several plant species, but no molecular mechanism for their transcriptional regulation has been elucidated. The wheat *TaBx3* and *TaBx4* genes, whose loci are thought to exist very closely to each other (Nomura et al. 2007), are excellent targets for studying the transcriptional co-regulation of the Bx biosynthetic genes. In this study, first we isolated genomic clones containing both the *TaBx3* and *TaBx4* genes from the A, B and D genomes of common wheat and sequenced them to reveal the molecular structures of those loci. Then, we analyzed the promoter sequences of the *TaBx3* and *TaBx4* alleles and also the orthologous sequences in the three diploid progenitors of common wheat.

Materials and methods

Plant materials

A common wheat (*T. aestivum*, $2n = 6x = 42$, genome formula AABBDD) cultivar, Chinese Spring (CS) was germinated and grown as described by Nomura et al. (2002), and 5-day-old shoots were used for preparation of the protoplasts. Diploid progenitors ($2n = 2x = 14$) of common wheat, *T. urartu* (accession KU199-6, genome formula AA), *Ae. speltoides* (KU5727, SS), and *Ae. tauschii* (KU20-9, DD), were used for cloning the promoter sequences of the *TaBx3* and *TaBx4* orthologs by genomic PCR. All seed stocks used in this study were obtained from the National BioResource Project-Wheat, Japan (http://www.nbrp.jp).

Isolation of TAC clones carrying the *TaBx3* and *TaBx4* loci of common wheat

Genomic clones carrying the *TaBx3* and *TaBx4* loci were screened from a transformation-competent artificial chromosome (TAC) library of CS by the pooled PCR method (Liu et al. 2000). The first screening was conducted with primer sets specific to each of the *TaBx4A*, *TaBx4B* and *TaBx4D* sequences. Forward (F) and reverse (R) primers were as follows: Bx4A-F (5′-TCGATACTAAGATTATT TGTGTCA-3′) and Bx4com-R1 (5′-CCTCCTTGATGGT CGCCTTG-3′) for TAC clone TaBx3A/TaBx4A; Bx4B-F (5′-TAAGGGAGTCCACACCCCAC-3′) and Bx4B-R (5′-GGCACTCAACTAACAGACCAT-3′) for TaBx3B/ TaBx4B; Bx4D-F (5′-ATGTAGGTAAATAATCGATAG CC-3′) and Bx4com-R1 for TaBx3D/TaBx4D. After the second screening by colony hybridization with the ^{32}P-labeled *TaBx4A* cDNA probe, positive clones were isolated for each of the three genomes. A shotgun library was constructed for each TAC clone and sequenced. Gene prediction was performed using the program "Rice Genome Automated Annotation System (RiceGAAS)" (http://RiceGAAS.dna.affrc.go.jp).

Cloning of the promoter sequences of the *TaBx3* and *TaBx4* orthologs in diploid progenitors

Total DNA was isolated from 10-day-old shoots of *T. urartu*, *Ae. speltoides* and *Ae. tauschii* by the CTAB method. Genomic PCR was performed in a 50-µl reaction mixture containing 50 ng of template DNA, each primer at 0.5 µM, 0.2 mM dNTPs, 1.5 mM $MgSO_4$, 1× reaction buffer, and 1 U of KOD-plus DNA polymerase (Toyobo). After an initial incubation for 2 min at 94°C, 35 cycles of amplification were performed (30 s at 96°C, 30 s at 64°C and 1 min at 68°C). Primers were designed based on the nucleotide sequences of the TAC clones isolated from CS to amplify sequences containing the 5′-upstream region and part of the coding region. Forward (F) and reverse (R) primers were as follows: Bx3A-proF (5′-CTAGCATTGTC CTGTTTGATTG-3′) and Bx3-comR1 (5′-TGCTGCTTG CCGAATCTTGTT-3′) for the *TaBx3A* ortholog of *T. urartu*; Bx3B-proF (5′-CTAGCATTGTCCTGTTTGGT TG-3′) and Bx3-comR1 for the *TaBx3B* ortholog of *Ae. speltoides*; Bx3D-proF (5′-CTAGCATTTGTCCTGTTTT-ATTG-3′) and Bx3-comR1 for the *TaBx3D* ortholog of *Ae. tauschii*; Bx4A-proF (5′-GAAACATCCTCCATGTTGGA CT-3′) and Bx4-comR2 (5′-GAACGGTGAGTAGTA GTGCTTCT-3′) for the *TaBx4A* ortholog of *T. urartu*; Bx4B-proF (5′-GAAACATCCTCCATGTGGGACT-3′) and Bx4B-R2 (5′-GGAGCTGCTGCTTGCATTTTG-3′) for the *TaBx4B* ortholog of *Ae. speltoides*; and Bx4D-proF (5′-GAAACATCCTCCATGTGGGATT-3′) and Bx4-comR2 for the *TaBx4D* ortholog of *Ae. tauschii*. After 5′-phosphorylation, the PCR products were ligated into the *Hin*cII site of pUC118 and sequenced.

Alignment of the three homoeologs and search for putative *cis*-regulatory elements within the promoter regions

The sequences of the three homoeologs of each of the *TaBx3* and *TaBx4* genes were aligned using ClustalW. The nucleotide identities between the homoeologs and orthologs of the *TaBx3* and *TaBx4* genes were calculated based on pairwise alignments using the program BioEdit. The promoter regions of respective homoeologs and orthologs were analyzed using the PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/; Lescot et al. 2002) to search for putative *cis*-regulatory elements.

Plasmid construction

The reporter plasmid pPluc (An et al. 2003), which has a cauliflower mosaic virus (CaMV) 35S promoter, a firefly (*Photinus pyralis*) *luciferase* (*LUC*) gene and a *nopaline synthase* (*NOS*) terminator, was digested with *Hin*dIII and *Sal*I, treated with T4 DNA polymerase, and self-ligated to produce a 35S-less pPluc vector. The 5′-upstream regions of the *TaBx3A* (−1025 to −1: nucleotide positions relative to the translation start site), *TaBx3B* (−1039 to −1), *TaBx3D* (−901 to −1), *TaBx4A* (−1756 to −1), *TaBx4B* (−1845 to −1) and *TaBx4D* (−1913 to −1) were respectively amplified by PCR using the TAC clones as template. The PCR products were 5′-phosphorylated and ligated into the blunt-ended *Hin*dIII site of the 35S-less pPluc vector. The phRluc plasmid (An et al. 2003), which comprises a CaMV35S promoter, a sea pansy (*Renilla reniformis*) *LUC* gene and a *NOS* terminator, was used as an internal standard in the protoplast transformation. The CaMV35S-sGFP (S65T)-NOS plasmid (Chiu et al. 1996) was used to check the transformation efficiency of the wheat protoplasts.

Preparation of wheat protoplasts

Protoplasts were prepared from 5-day-old shoots of CS as described by Sarhan and Cesar (1988) with some modifications. All operations were done at room temperature unless otherwise noted. The upper 5 cm part of the shoots was cut into 1 mm slices with a razor blade. Approximately 0.5 g of the sliced tissue was put in 10 ml of enzyme solution (2% Cellulase Onozuka RS, 0.5% Macerozyme R10, 5 mM MES, 1 mM NaH_2PO_4, 0.6 M sorbitol, 5 mM $CaCl_2$, 2 mM $MgCl_2$, 2 mM $MnCl_2$, 1 mM arginine, 1 mM DTT, 0.1% PVP, 2 mM glutathione, and 2 mM ascorbate) in a Petri dish (9-cm diameter). After 5-min vacuum infiltration, the

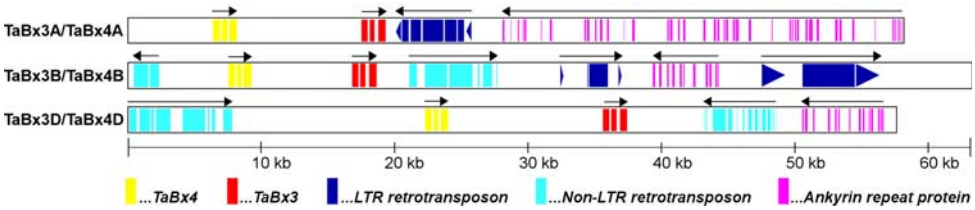

Fig. 2 Gene composition of the three TAC clones, TaBx3A/TaBx4A, TaBx3B/TaBx4B and TaBx3D/TaBx4D, predicted by the RiceGAAS program. Arrows indicate the gene direction

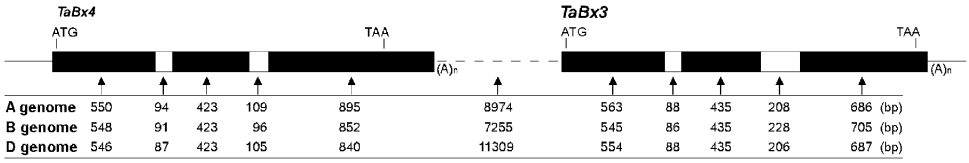

Fig. 3 Exon-intron structures of the *TaBx3* and *TaBx4* genes in A, B and D genomes of common wheat. Exons and introns are represented by black and white boxes, respectively. ATG, TAA and $(A)_n$ represent translation start codon, stop codon and poly-adenylate addition site, respectively. The size difference in the first exon and the third exon of the three homoeologs is due to the difference in the length of the 5′- and 3′-untranslated regions

tissue was digested for 3 h at 28°C in the dark. Then, protoplasts were released by shaking at 80 rpm for 1 min, and cell debris was removed by filtering through a nylon filter (60-µm pore size). Protoplasts were washed three times with washing solution (enzyme solution without enzymes and PVP), and resuspended in MaCaM solution (Dubouzet et al. 2005). The yield and quality of the protoplasts were checked under a microscope with a haemocytometer.

Transient-expression assay in protoplasts

Protoplast transformation and measurements of LUC activity were performed according to Dubouzet et al. (2005) with some modifications. Plasmid DNAs were purified with a QIAGEN Plasmid Maxi Kit (Qiagen). Approximately 1×10^5 protoplasts in 200 µl of MaCaM solution were mixed gently with 20 µl of plasmid solution containing 40 µg of reporter plasmid and 1 ng of phRluc plasmid, and then 220 µl of PEG solution was added to the mixed solution. After 20-min incubation at 28°C, the protoplast suspension was mixed with 800 µl of W5 solution, and then protoplasts were spun down, washed once with 1 ml of W5 solution, and resuspended in 1.2 ml of W5 solution. The protoplasts were transferred into a 12-well plastic plate and incubated for 24 h at 25°C in the dark without shaking. After the incubation, protoplasts were collected and homogenized in 100 µl of Passive Lysis buffer. Cell residues were removed by centrifugation (15,000 rpm, 5 min), and 10 µl of the supernatant was mixed with 50 µl of LUC Assay Reagent II to measure Pp-LUC activity, followed by the measurement of Rr-LUC activity by adding 50 µl of Stop & Glo reagent. Four independent transformations were conducted for each experiment, and each experiment was repeated three times using newly isolated protoplasts. Relative Pp-LUC activities (compared to Rr-LUC activities) were calculated to correct for transformation efficiency, and the mean of the 12 replicates was calculated.

Results

Molecular structure of the *TaBx3* and *TaBx4* loci in common wheat

We isolated three genomic clones from the TAC library of CS using the *TaBx4* cDNA as a probe. As speculated by Nomura et al. (2007), both the *TaBx3* and *TaBx4* genes were included in all three TAC clones. The nucleotide sequences of these genes in the three clones corresponded respectively to the *TaBx3* and *TaBx4* cDNA sequences in the three genomes of common wheat (GenBank accession no. AB298184 for TAC clone TaBx3A/TaBx4A, AB298185 for TaBx3B/TaBx4B and AB298186 for TaBx3D/TaBx4D). The insert sizes of TaBx3A/TaBx4A, TaBx3B/TaBx4B and TaBx3D/TaBx4D were 58,073, 63,149 and 57,580 bp, respectively. In addition, the presence of the *ankyrin repeat protein* gene, an LTR retrotransposon and a non-LTR retrotransposon in the TAC clones was predicted by the RiceGAAS program (Fig. 2). In all three clones *TaBx3* was located downstream of *TaBx4* in a tail-to-head manner. The two genes were separated by 8,974, 7,255 and 11,309 bp in the A, B and D genomes, respectively (Fig. 3).

The exon-intron structures of *TaBx3* and *TaBx4* were conserved among the three homoeologs (Fig. 3). Both genes comprised three exons and two introns in conserved positions. The maize orthologs *ZmBx3* and *ZmBx4* have two introns and one intron, respectively (Frey et al. 1997). Since the *ZmBx4* intron corresponds to the second intron of *TaBx4*, the first intron of *TaBx4* may have arisen after the

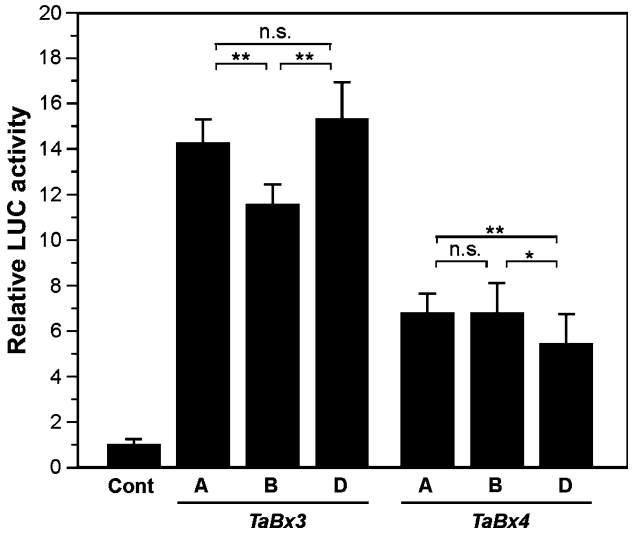

Fig. 4 Promoter activity of the A, B and D-genome homoeologs of the *TaBx3* and *TaBx4* genes. The Pp-LUC reporter was transiently expressed in wheat protoplasts under the control of the promoter of each homoeolog. The Pp-LUC activity was normalized by the Rr-LUC activity (internal standard), setting the relative activity for the promoter-less pPluc construct (Cont) at 1.0. Four independent transformations were conducted for each experiment, and the experiment was repeated three times using newly isolated protoplasts. The results are expressed as the mean ± SD ($n = 12$). The difference in promoter activity between the homoeologs was tested using Student's t test. * Significant at the 0.05 level; ** significant at the 0.01 level; *NS* not significant

differentiation of wheat and maize or the corresponding intron may have been lost in maize.

Promoter activity of the 5′-upstream regions of the *TaBx3* and *TaBx4* genes

The promoter activity of the 5′-upstream regions was measured using a transient expression system with firefly LUC as a reporter. First we used particle bombardment to introduce the reporter plasmids into 5-day-old shoots of CS, but no LUC activity was detected. Since the same experiment was successful when the CaMV35S promoter was fused (data not shown), the failure seemed to be due to a lack of transcription of the *TaBx* genes in the epidermis, as shown by RNA in situ hybridization for the *ZmBx4* gene in maize (Frey et al. 1995). Then we introduced the reporter plasmids into mesophyll protoplasts prepared from 5-day-old shoots of CS to successfully detect the LUC activity (Fig. 4). This suggested that the 5′-upstream regions of all three homoeologs of *TaBx3* and *TaBx4* possessed promoter activity. The transformation efficiency of the wheat protoplasts (about 70%) was estimated by fluorescence microscopy after introducing the GFP expression plasmid CaMV35S-sGFP (S65T)-NOS.

The promoter activity increased in the order of *TaBx3D*, *TaBx3A*, and *TaBx3B* (Fig. 4). The promoter activities of *TaBx4A* and *TaBx4B* were the same, and were slightly higher than that of *TaBx4D*. On average, the promoter activity of the *TaBx3* homoeologs was approximately twice as high as that of the *TaBx4* homoeologs.

Table 1 Nucleotide identity (%) between the coding regions of the A, B and D-genome homoeologs of the *TaBx3* and *TaBx4* genes and their orthologs in maize

	B	D	Maize
TaBx3			
A	96.3	96.5	77.3
B	–	95.8	77.6
D	–	–	77.3
TaBx4			
A	96.7	96.1	78.9
B	–	96.6	79.1
D	–	–	78.6

Sequence diversity in coding sequences and promoter regions of the *TaBx3* and *TaBx4* genes

The coding regions of the three *TaBx3* homoeologs shared a nucleotide sequence with 95.8–96.5% identity and those of the three *TaBx4* homoeologs shared a nucleotide sequence with 96.1–96.7% identity (Table 1). The nucleotide sequences of the 5′-upstream regions from the translation start site of the *TaBx3* homoeologs were able to be aligned within −1025 to −1 (*TaBx3A*), −1039 to −1 (*TaBx3B*), and −901 to −1 (*TaBx3D*) (Supplemental Fig. 5). For the *TaBx4* homoeologs, the 5′-upstream regions, −3693 to −1 (*TaBx4A*), −4455 to −1 (*TaBx4B*), and −4532 to −1 (*TaBx4D*) could be aligned, but there were long inserts with 16-bp terminal inverted repeats in *TaBx4B* (612 bp: −2586 to −1975) and in *TaBx4D* (616 bp: −2596 to −1981). Therefore, shorter sequences, −1756 to −1 (*TaBx4A*), −1845 to −1 (*TaBx4B*) and −1913 to −1 (*TaBx4D*) were used for further analyses of the *TaBx4* homoeologs (Supplemental Fig. 6). In the promoter regions, the nucleotide identity of the three homoeologs was lower than that in coding regions, 74.0–88.2% identity for *TaBx3* and 74.1–89.0% for *TaBx4* (Table 2). The nucleotide identity in the coding region between the respective *TaBx3* homoeologs and the maize ortholog *ZmBx3* was 77.3–77.6% and that between the respective *TaBx4* homoeologs and the maize ortholog *ZmBx4* was 78.6–79.1% (Table 1), but no significant alignment was possible in the promoter region between the *TaBx3/TaBx4* homoeologs and *ZmBx3/ZmBx4* orthologs (Table 2).

We isolated the corresponding promoter sequences from the three diploid progenitors of common wheat, *T. urartu* (AA), *Ae. speltoides* (SS) and *Ae. tauschii* (DD) (GenBank

Table 2 Nucleotide identity (%) between the promoter regions of the A, B and D-genome homoeologs of the *TaBx3* and *TaBx4* genes and their orthologs in the three diploid progenitors and maize

	B	D	*T.urartu*	*Ae. speltoides*	*Ae. tauschii*	Maize
TaBx3						
A	88.2	76.3	97.1	85.8	76.3	NS
B	–	74.0	87.6	86.4	74.5	NS
D	–	–	75.2	72.1	99.6	NS
T. urartu	–	–	–	85.2	75.7	NS
Ae. speltoides	–	–	–	–	72.1	NS
Ae. tauschii	–	–	–	–	–	NS
	B	D	*T. urartu*	*Ae. speltoides*	*Ae. tauschii*	Maize
TaBx4						
A	74.1	76.6	98.9	71.4	76.7	NS
B	–	89.0	74.2	83.1	89.2	NS
D	–	–	76.7	83.5	99.4	NS
T. urartu	–	–	–	71.6	76.7	NS
Ae. speltoides	–	–	–	–	83.6	NS
Ae. tauschii	–	–	–	–	–	NS

NS stands for no significant alignment

accession nos. AB297462 to AB297467), to examine to what extent the sequence similarity is preserved among the orthologs in common wheat and its diploid progenitors. The *TaBx3* and *TaBx4* promoters of common wheat showed high levels of nucleotide identity with the promoters of orthologous genes of the diploid species, 72.1–99.6% for *TaBx3* and 71.4–99.4% for *TaBx4* (Table 2). For both *TaBx*3 and *TaBx4*, the nucleotide identity was the lowest between the B genome of common wheat and the S genome of *Ae. speltoides*.

We also examined the sequence similarity between the *TaBx3* and *TaBx4* genes, which are considered to have originated by gene duplication, as in maize (Frey et al. 1997). In the coding regions, the *TaBx3* homoeologs retained 74.9–75.3% nucleotide identity with the corresponding *TaBx4* homoeologs, but in the promoter regions no significant sequence similarity was found between the two genes.

Cis-regulatory elements in the *TaBx3* and *TaBx4* promoters

Common *cis*-regulatory elements were searched for in the promoter regions of all homoeologs of the *TaBx3* and *TaBx4* genes of common wheat and the diploid progenitors using the PlantCARE program. In addition to TATA- and CAAT-boxes, all *TaBx3* and *TaBx4* homoeologs in common wheat and their orthologs in diploid progenitors had three putative *cis*-regulatory elements in common (Table 3, Supplemental Figs. 5, 6). They were the G-box and the GAG-motif, both of which are elements mainly involved in light responsiveness, and the MBS, a MYB transcription-factor binding site. In addition to these elements, Box 4, Skn-1-like motif and W box were shared by all *TaBx3* homoeologs and orthologs, of which Box 4 was specific to *TaBx3* (Supplemental Table 4). In all *TaBx4* homoeologs and orthologs, C-repeat/DRE, CAT-box, CATT-motif, GCC box, OCT and Circadian were present in common (Supplemental Table 5). GCC box was specifically found in *TaBx4*. Box 4 in *TaBx3* and CATT-motif in *TaBx4* are related to the light responsiveness, and W box in *TaBx3* and GCC box in *TaBx4* are the elements involved in the wounding and pathogen responsiveness. In these *TaBx* genes the loss of one element seems to be compensated by the other element with a similar function.

Discussion

Comparisons of the nucleotide sequences among the three homoeologs of the *TaBx* genes revealed that the sequence similarity was lower in promoter regions than in coding regions. The *TaBx* genes of common wheat and their maize orthologs had fairly similar nucleotide sequences in their coding regions, but not in their promoters. However, the *TaBx* promoters retained sequence similarity to their orthologs of the diploid wheat species. These facts suggest that the promoter sequences of the *TaBx* homoeologs and orthologs have drastically diverged during the separation of wheat and maize. This is not beyond expectation because nucleotides change more frequently in promoter regions than in coding regions (Wray et al. 2003) and because wheat and maize separated about 60 MYA (Huang et al. 2002a) while the diploid progenitors of wheat radiated about 2.5–4.5 MYA (Huang et al. 2002b). Such a rapid change in the promoter sequences was also observed between the *TaBx3* and *TaBx4* genes, which are supposed

Table 3 Putative *cis*-regulatory elements in the promoter sequences shared by all *TaBx3* and *TaBx4* homoeologs and orthologs in the diploid progenitors

Motif (sequence)	Species	Genome	Distance from ATG	
			TaBx3	*TaBx4*
G-box* (GACATGTGGT[a], GACACGTAGT[b], CACATGG[c], GTACGTG[d], CACGAC[e], CACGTA[f], CACGTT[g], CACGTC[h], CACGTG[i], TACGTT[j], TACGTG[k])	*T. aestivum* cv. CS	A	−183 (+)[g], −392 (−)[e]	−97 (−)[e], −695 (+)[a], −1355 (+)[a]
		B	−224 (−)[k], −229 (+)[f], −378 (+)[i], −381 (−)[e]	−93 (−)[e], −1483 (+)[b], −1691 (−)[e], −1829 (−)[c]
		D	−168 (+)[g], −232 (−)[k], −237 (+)[f]	−93 (−)[f], −99 (+)[d], −673 (+)[a], −1277 (+)[e], −1509 (+)[b], −1897 (−)[c]
	T. urartu	A	−186 (+)[g], −243 (−)[k], −248 (+)[f], −395 (−)[e]	−97 (−)[e], −695 (+)[a], −1348 (−)[h]
	Ae. speltoides	S	−158 (+)[g], −278 (−)[k], −283 (+)[f]	−93 (−)[f], −98 (+)[k], −546 (+)[a], −1175 (+)[e], −1408 (+)[b], −1645 (−)[e], −1788 (−)[c]
	Ae. tauschii	D	−168 (+)[g], −232 (−)[k], −237 (+)[f]	−93 (−)[f], −99 (+)[d], −672 (+)[a], −1277 (+)[e], −1509 (+)[b], −1897 (−)[c]
GAG-motif** (AGAGATG[l], AGAGAGT[m], GGAGATG[n])	*T. aestivum* cv. CS	A	−72 (+)[l], −210 (+)[l], −335 (+)[m], −629 (+)[l]	−536 (−)[l], −1356 (+)[l]
		B	−63 (+)[l], −187 (+)[l], −324 (+)[m], −639 (+)[l]	−388 (−)[n]
		D	−195 (+)[l]	−395 (−)[n]
	T. urartu	A	−72 (+)[l], −213 (+)[l], −338 (+)[m], −632 (+)[l]	−536 (−)[l]
	Ae. speltoides	S	−62 (+)[l], −185 (+)[l], −373 (+)[m], −674 (+)[l]	−392 (−)[n]
	Ae. tauschii	D	−72 (+)[l], −195 (+)[l]	−395 (−)[n]
MBS*** (TAACTG[o], CAACTG[p])	*T. aestivum* cv. CS	A	−11 (−)[p], −816 (+)[o]	−644 (+)[o]
		B	−11 (−)[p], −830 (+)[o]	−581 (+)[o]
		D	−11 (−)[p], −569 (−)[o], −694 (+)[p]	−623 (+)[o]
	T. urartu	A	−11 (−)[p], −446 (+)[o], −820 (+)[o]	−644 (+)[o]
	Ae. speltoides	S	−11 (−)[p], −860 (+)[o]	−496 (+)[o]
	Ae. tauschii	D	−11 (−)[p], −569 (−)[o], −694 (+)[p]	−622 (+)[o]
TATA-box**** (TATATAA)	*T. aestivum* cv. CS	A	−76 (−)	−55 (−)
		B	−67 (−)	−55 (−)
		D	−76 (−)	−55 (−)
	T. urartu	A	−76 (−)	−55 (−)
	Ae. speltoides	S	−66 (−)	−55 (−)
	Ae. tauschii	D	−76 (−)	−55 (−)

(+) and (-) indicate strands on which each element is present

Note: Many putative CAAT-boxes found in each sequence were not listed in this table, because it was impossible to identify an authentic one among them

[a–p] Several types of sequences are present for each motif in PlantCARE database; a–p indicates each of motif sequences found in promoter regions

* Menkens et al. (1995); ** Orozco and Ogren (1993); *** Urao et al. (1993); **** Grace et al. (2004)

to have evolved by duplication, as hypothesized from the deduction that the four maize P450 genes for Bx biosynthesis, *ZmBx2–ZmBx5*, have evolved by duplication (Frey et al. 1997). The *TaBx3* genes shared the coding sequences with about 75% nucleotide identity with the corresponding *TaBx4* genes from the same genomes. On the other hand, there was no sequence similarity in the promoter sequence between the *TaBx3* and *TaBx4* genes, although both genes are transcribed at high levels specifically at the seedling stage (Nomura et al. 2005). These results imply the presence of some essential *cis*-regulatory elements in the promoter regions, rather than their overall sequence similarity, for the regulation of the *TaBx3* and *TaBx4* transcription. This also accords with a previous report that the overall substitution and insertion/deletion frequencies are generally higher in promoter sequences than in coding sequences because most nucleotide changes in promoter regions do not affect transcription (Wray et al. 2003).

It has been described that orthologous genes usually retain the same *cis*-regulatory elements (Rombauts et al. 2003). The transient, high-level transcription in seedlings was observed for all homoeologs of the five Bx biosynthetic genes of common wheat and their orthologs of the ancestral tetraploid wheat and diploid species (Nomura et al. 2005). This fact implies the presence of some underlying unified regulatory mechanisms for regulating the transcription of the homoeologs and orthologs of the five Bx biosynthetic genes. In addition to the TATA- and CAAT-boxes, three putative *cis*-regulatory elements, MBS, G-box and GAG-motif, were common to all *TaBx3* and *TaBx4* homoeologs and orthologs. These three elements were also present in maize, within a 2 kb 5′-upstream region of *ZmBx4* (GenBank accession no. X81828). The 379-bp sequence of the 5′-upstream region of *ZmBx3* (Y11404) contained the G-box. It is well known that the expression of multiple biosynthetic genes for a secondary metabolite is often regulated by a common transcription factor; e.g., MYB transcription factors regulate most of the genes involved in the biosynthesis of anthocyanin (Gantet and Memelink 2002; Vom Endt et al. 2002). Since MYB transcription factors function in the regulation of a wide range of plant genes, including genes for secondary metabolism, development, signal transduction and disease resistance (Jin and Martin 1999), one of them might be a common transcription factor in Bx biosynthesis. Another mechanism might be involved in the co-expression of the *TaBx3* and *TaBx4* genes. Zhan et al. (2006) demonstrated that neighboring genes are more frequently co-expressed than expected by chance. They suggested the presence of chromosomal domain-level regulation of neighboring genes with the same chromatin status. The same regulatory mechanism may be functioning in the co-expression of the *TaBx3* and *TaBx4* genes, whose loci were found to be adjacent in the present study.

Although the B-genome homoeologs of *TaBx3* and *TaBx4* are preferentially transcribed in 5-day-old shoots (Nomura et al. 2005), there was only a slight difference in promoter activity among the three homoeologs (Fig. 4). Although we cannot exclude the possibility of the differential regulation of transcription between wheat seedlings and protoplasts, this discrepancy between the transcript levels and the promoter activity may be explained differently in terms that the transgenes in naked plasmids, not integrated into a chromosome and without forming a chromatin structure, were transcribed in the transient expression system used in the present study. It has been reported that the gene silencing of particular homoeologs in common wheat is associated with an epigenetic mechanism (Bottley et al. 2006; Shitsukawa et al. 2007), suggesting the involvement of chromatin structure in the transcriptional regulation of wheat homoeologs. A similar mechanism may be involved in the regulation of the *TaBx* homoeologs.

Acknowledgments We thank Dr. Ei-ichiro Fukusaki (Osaka University, Japan) for providing pPluc and phRluc vectors. We also acknowledge Dr. Yasuo Niwa (University of Shizuoka, Japan) for a gift of CaMV35S-sGFP (S65T)-NOS plasmid. This work was supported by a Grant-in-Aid for Scientific Research (No. 16000377) from the Japan Society for the Promotion of Science (to T.N.). Contribution No. 594 from the Laboratory of Plant Genetics, Graduate School of Agriculture, Kyoto University, Japan.

References

An CI, Sawada A, Fukusaki E, Kobayashi A (2003) A transient RNA interference assay using *Arabidopsis* protoplasts. Biosci Biotechnol Biochem 67:2674–2677

Bottley A, Xia GM, Koebner RMD (2006) Homoeologous gene silencing in hexaploid wheat. Plant J 47:897–906

Chiu W, Niwa Y, Zeng W, Hirano T, Kobayashi H, Sheen J (1996) Engineered GFP as a vital reporter in plants. Curr Biol 6:325–330

Dubouzet JG, Morishige T, Fujii N, An CI, Fukusaki E, Ifuku K, Sato F (2005) Transient RNA silencing of scoulerine 9-*O*-methyltansferase expression by double stranded RNA in *Coptis japonica* protoplasts. Biosci Biotechnol Biochem 69:63–70

Ebisui K, Ishihara A, Hirai N, Iwamura H (1998) Occurrence of 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one (DIMBOA) and a β-glucosidase specific for its glucoside in maize seedlings. Z Naturforsch 53c:793–798

Feldman M, Levy AA (2005) Allopolyploidy - a shaping force in the evolution of wheat genomes. Cytogenet Genome Res 109:250–258

Frey M, Kliem R, Saedler H, Gierl A (1995) Expression of a cytochrome P450 gene family in maize. Mol Gen Genet 246:100–109

Frey M, Chomet P, Glawischnig E, Stettner C, Grün S, Winklmair A, Eisenreich W, Bacher A, Meeley RB, Briggs SP, Simcox K, Gierl A (1997) Analysis of a chemical plant defense mechanism in grasses. Science 277:696–699

Frey M, Stettner C, Paré PW, Schmelz EA, Tumlinson JH, Gierl A (2000) An herbivore elicitor activates the gene for indole emission in maize. Proc Natl Acad Sci USA 97:14801–14806

Gachon CMM, Langlois-Meurinne M, Henry Y, Saindrenan P (2005) Transcriptional co-regulation of secondary metabolism enzymes

in *Arabidopsis*: functional and evolutionary implications. Plant Mol Biol 58:229–245

Gantet P, Memelink J (2002) Transcription factors: tools to engineer the production of pharmacologically active plant metabolites. Trends Pharmacol Sci 23:563–569

Grace ML, Chandrasekharan MB, Hall TC, Crowe AJ (2004) Sequence and spacing of TATA box elements are critical for accurate initiation from the β-phaseolin promoter. J Biol Chem 279:8102–8110

Grün S, Frey M, Gierl A (2005) Evolution of the indole alkaloid biosynthesis in the genus *Hordeum*: Distribution of gramine and DIBOA and isolation of the benzoxazinoids biosynthesis genes from *Hordeum lechleri*. Phytochemistry 66:1264–1272

Huang S, Sirikhachornkit A, Faris JD, Su X, Gill BS, Haselkorn R, Gornicki P (2002a) Phylogenetic analysis of the acetyl-CoA carboxylase and 3-phosphoglycerate kinase loci in wheat and other grasses. Plant Mol Biol 48:805–820

Huang S, Sirikhachornkit A, Su X, Faris J, Gill B, Haselkorn R, Gornicki P (2002b) Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum/Aegilops* complex and the evolutionary history of polyploid wheat. Proc Natl Acad Sci USA 99:8133–8138

Jin H, Martin C (1999) Miltifunctionality and diversity within the plant *MYB*-gene family. Plant Mol Biol 41:577–585

Kato N, Dubouzet E, Kokabu Y, Yoshida S, Taniguchi Y, Dubouzet JG, Yazaki K, Sato F (2007) Identification of a WRKY protein as a transcriptional regulator of benzylisoquinoline alkaloid biosynthesis in *Coptis japonica*. Plant Cell Physiol 48:8–18

Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S (2002) PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucl Acids Res 30:325–327

Liu Y-G, Nagaki K, Fujita M, Kawaura K, Uozumi M, Ogihara Y (2000) Development of an efficient maintenance and screening system for large-inserted genomic DNA libraries of hexaploid wheat in a transformation-competent artificial chromosome (TAC) vector. Plant J 23:687–695

Melanson D, Chilton M-D, Masters-Moore D, Chilton WS (1997) A deletion in an indole synthase gene is responsible for the DIMBOA-deficient phenotype of *bxbx* maize. Proc Natl Acad Sci USA 94:13345–13350

Menkens AE, Schindler U, Cashmore AR (1995) The G-box: a ubiquitous regulatory DNA element in plants bound by the GBF family of bZIP proteins. Trends Biochem Sci 20:506–510

Nakagawa E, Amano T, Hirai N, Iwamura H (1995) Non-induced cyclic hydroxamic acids in wheat during juvenile stage of growth. Phytochemistry 38:1349–1354

Niemeyer HM (1988) Hydroxamic acids (4-hydroxy-1,4-benzoxazin-3-ones): defence chemicals in the Gramineae. Phytochemistry 27:3349–3358

Nomura T, Ishihara A, Imaishi H, Endo TR, Ohkawa H, Iwamura H (2002) Molecular characterization and chromosomal localization of cytochrome P450 genes involved in the biosynthesis of cyclic hydroxamic acids in hexaploid wheat. Mol Genet Genomics 267:210–217

Nomura T, Ishihara A, Imaishi H, Ohkawa H, Endo TR, Iwamura H (2003) Rearrangement of the genes for the biosynthesis of benzoxazinones in the evolution of Triticeae species. Planta 217:776–782

Nomura T, Ishihara A, Yanagita RC, Endo TR, Iwamura H (2005) Three genomes differentially contribute to the biosynthesis of benzoxazinones in hexaploid wheat. Proc Natl Acad Sci USA 102:16490–16495

Nomura T, Ishihara A, Iwamura H, Endo TR (2007) Molecular characterization of benzoxazinone-deficient mutation in diploid wheat. Phytochemistry 68:1008–1016

Orozco BM, Ogren WL (1993) Localization of light-inducible and tissue-specific regions of the spinach ribulose bisphosphate carboxylase/oxygenase (rubisco) activase promoter in transgenic tobacco plants. Plant Mol Biol 23:1129–38

Ren X-Y, Fiers MWEJ, Stiekema WJ, Nap J-P (2005) Local coexpression domains of two to four genes in the genome of *Arabidopsis*. Plant Physiol 138:923–934

Rombauts S, Florquin K, Lescot M, Marchal K, Rouzé P, Van de Peer Y (2003) Computational approaches to identify promoters and *cis*-regulatory elements in plant genomes. Plant Physiol 132:1162–1176

Sarhan F, Cesar D (1988) High yield isolation of mesophyll protoplasts from wheat, barley and rye. Physiol Plant 72:337–342

Shitsukawa N, Tahira C, Kassai K, Hirabayashi C, Shimizu T, Takumi S, Mochida K, Kawaura K, Ogihara Y, Murai K (2007) Genetic and epigenetic alteration among three homoeologous genes of a class E MADS box gene in hexaploid wheat. Plant Cell 19: 1723–1737

Sicker D, Frey M, Schulz M, Gierl A (2000) Role of natural benzoxazinones in the survival strategy of plants. Int Rev Cytol 198:319–346

Sue M, Ishihara A, Iwamura H (2000) Purification and characterization of a β-glucosidase from rye (*Secale cereale* L.) seedlings. Plant Sci 155:67–74

Urao T, Yamaguchi-Shinozaki K, Urao S, Shinozaki K (1993) An *Arabidopsis myb* homolog is induced by dehydration stress and its gene product binds to the conserved MYB recognition site. Plant Cell 5:1529–1539

Vom Endt D, Kijne JW, Memelink J (2002) Transcription factors controlling plant secondary metabolism: what regulates the regulators? Phytochemistry 61:107–114

Wray GA, Hahn MW, Abouheif E, Balhoff JP, Pizer M, Rockman MV, Romano LA (2003) The evolution of transcriptional regulation in eukaryotes. Mol Biol Evol 20:1377–1419

Zhan S, Horrocks J, Lukens LN (2006) Islands of co-expressed neighbouring genes in *Arabidopsis thaliana* suggest higher-order chromosome domains. Plant J 45:347–357