**ORIGINAL ARTICLE**

# Isolation and structural analysis of the *Bx6* and *Bx7* genes controlling the biosynthesis of benzoxazinoids in rye (*Secale cereale* L.)

Beata Bakera[1] · Monika Rakoczy-Trojanowska[1]

## Abstract

Benzoxazinoids (BXs) are plant secondary metabolites, first discovered in the 1950s, which are synthesized in many monocotyledonous species from the Poaceae family and in several dicotyledonous plants. They constitute a significant element of the plant's defence mechanism against both biotic (pests and diseases) and abiotic (elevated salinity, heavy metals) stresses. The aim of this research was to gain more information on the genetic background of BXs biosynthesis in rye (*Secale cereale* L.) by isolating and sequencing the $ScBx6_{L318}$ and $ScBx7_{L318}$ genes encoding 2-oxoglutarate-dependent dioxygenase and 7-*O*-methyltransferase, respectively. Using the modified Amplicon Express method, BAC clones containing the $ScBx6_{L318}$ and $ScBx7_{L318}$ genes were isolated and sequenced. The similarity between the $ScBx6_{L318}$ and *ZmBx6* genes were, correspondingly, 78% and 68% in the coding sequence (*cds*) and the amino acid sequence (AA). A lesser similarity was found between the $ScBx7_{L318}$ and *ZmBx7* genes (72% and 45% at the *cds* and AA levels, respectively).

**Keywords** *ScBx* genes · Benzoxazinoids · Rye BAC library · Phylogenetic analysis · Polymorphisms

## Introduction

The first stage of the BXs biosynthesis pathway is the conversion of indole-3-glycerophosplast to indole catalysed by indole synthase in the chloroplast. Then, four cytochrome P450 monooxygenases catalyse the transformation of indole to DIBOA in the endoplasmic reticulum (Frey et al. 1997, 2009; Grün et al. 2005; Chu et al. 2011). Afterwards, glucosylation of DIBOA to DIBOA-Glc localised in the cytoplasm is necessary to transform toxic aglucon into a harmless, stable compound, and its increased solubility, compared to aglucon form, allows it to be stored in a vacuole (Dick et al. 2012). DIBOA-Glc hydroxylation and *O*-methylation reactions cause its transformation into DIMBOA-Glc (Frey

et al. 2003; Jonczyk et al. 2008). DIMBOA-Glc can be transformed into HDMBOA-Glc by successive *O*-methylation reactions (Meihls et al. 2013; Handrick et al. 2016; Wouters et al. 2016) or by the hydroxylation reaction to TRIMBOA-Glc, which becomes a substrate for two *O*-methylation reactions leading to $HDM_2BOA$-Glc synthesis (Handrick et al. 2016; Wouters et al. 2016). All glucosides produced are transported and stored in the vacuole (Frey et al. 2009; Niemeyer 2009; Meihls et al. 2013; Makowska et al. 2015). As a result of an increase in jasmonic acid and/or its ester in response to a pathogen or pest attack, DIBOA-Glc and DIMBOA-Glc undergo hydroxylation, and their aglucons (DIBOA and DIMBOA) are released from the vacuole into the cytoplasm (Oikawa et al. 2002; Niemeyer 2009). The BXs biosynthesis pathway is best known in maize, although its scope in other species is steadily increasing.

So far, many *Bx* genes controlling BXs biosynthesis have been found, namely 16 genes in maize (*ZmBx1–ZmBx14*, *ZmGlu1–ZmGlu2*), (Frey et al. 1997; Czjzek et al. 2000; Rad et al. 2001; Jonczyk et al. 2008; Sue et al. 2011; Meihls et al. 2013; Handrick et al. 2016), 13 genes in wheat (*TaBx1–TaBx5*, *TaGlu1a–TaGlu1d*, *TaGT1a–TaGT1d*), (Nomura et al. 2002, 2003, 2005; Sue et al. 2011), five genes in *H. lechleri* (*HlBx1–HlBx5*), (Grün et al. 2005), eight genes in rye (*ScBx1–ScBx5*; whole gene sequences), (Bakera

✉ Beata Bakera
beata_bakera@sggw.pl

1 Department of Plant Genetics, Breeding and Biotechnology, Institute of Biology, Warsaw University of Life Sciences, 166 Nowoursynowska Str., 02-787 Warsaw, Poland

et al. 2015), as well as *ScBx1–ScBx6-like* (*cds*, Tanwir et al. 2017), *ScGlu* (protein sequence, Nikus et al. 2003), and *ScGT* (mRNA sequence, Sue et al. 2011). Recently, Groszyk et al. (2017) proved the biological function of the *ScBx1* gene using the virus-induced gene silencing approach. To date, only the *ScBx6-like* gene *cds* has been isolated in rye (Tanwir et al. 2017). Moreover, the existence of the *ScBx6* and *ScBx7* genes was questioned by, for example, Frey et al. (2003), despite the confirmed synthesis of DIMBOA. Therefore, we assumed that both genes must be present in the rye genome.

## Materials and methods

### Plant and biological material

The BAC library used for isolation of the $ScBx6_{L318}$ and $ScBx7_{L318}$ genes, constructed from the DNA of rye inbred line L318, consisted of 39 superpools, each containing 2688 clones.

DNA for primer verification was isolated from the leaves of winter rye (*Secale cereale* L.) inbred line L318 using a CTAB method described by Murray and Thompson (1980). Plants were grown in a Conviron CMP 6050 climate-controlled chamber (photoperiod 16 h day/8 h night, temperature 22/20 °C, light intensity 200 μE, humidity 70%) for 2 weeks.

### Isolation of positive BAC clones and PCR

The positive clones, i.e., those containing searchable genes, were isolated using the modified Amplicon Express Strategy (https://ampliconexpress.com/products-services/screening-services/pools-andsuperpools). In the first stage, the PCR template was the gDNA (at a concentration of 100 ng/1 μl) of the 39 superpools. The next step was a PCR with plate (W), row (R), and column (T) pools as templates from the superpools selected in the first stage. The PCRs contained 500 ng of total genomic DNA, 3 μM of forward and reverse primers, 0.2 mM of dNTPs, 0.5 mM of $MgCl_2$, $1 \times PCR$ buffer, and 3 units of DreamTaq polymerase (Fermentas) in a total volume of 15 μl. Amplification was performed in a thermal cycler Gene Amp PCR System 9700 using the following conditions: (1) 94 °C for 1 min; (2) 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 60 s for 35 cycles; and (3) 72 °C for 5 min. The products were separated on a 1% agarose gel, stained with ethidium bromide, and visualised on a UV transilluminator. Selected BAC clones were suspended in LB medium with chloramphenicol and sent to the GENOMED S.A. company, where MiSeq (Illumina) PE250 and CLC Genomic Workbench v. 7.5 were used for

**Table 1** Primers used for BAC library screening

| Gene | Primer pairs |
| --- | --- |
| *ScBx6* | $F_I$: CGACAC CCTCTT CTTCCACC |
| | $R_I$: TTATTC GCACCC CTCCACAG |
| | $F_{II}$: CATCTA CAGCAG GGACCA GG |
| | $R_{II}$: TCAAGC GAGCGA CCATCAAG |
| *ScBx7* | $F_{III}$: TTACAA TTTAAT TCTCTG GCAACGC |
| | $R_{III}$: AATTCA TTAGCA CTAACT GGTCACA |
| | $F_{VI}$: TGTGTT CTAAAA TACACA AGAATGC |
| | $R_{VI}$: GCAACC TGCACC AAAAAT AA |

sequencing and assembly, respectively. The resulting contigs were analysed bioinformatically to identify the $ScBx_{L318}$ genes.

### Primer design

To isolate the $ScBx6_{L318}$ and $ScBx7_{L318}$ genes, primers were constructed based on the sequences of their maize orthologs *ZmBx6* and *ZmBx7* (GenBank Acc. No. AF540907 and EU192149, respectively). After establishing fragmentary sequences of the $ScBx6_{L318}$ and $ScBx7_{L318}$ genes and based on blast analysis against the rye genome sequence (Bauer et al. 2017, GenBank assembly: GCA_900079665.1), rye-specific primers were designed. Four primer pairs were used for BAC library screening based on two selected amplicons per gene (Table 1). The designed primers were synthesised at GENOMED S.A. in Warsaw.

### Bioinformatic analysis

Sequence analysis was performed using the Sequencher 4.5 program. BioEdit 7.0.9.0 software was used to identify the $ScBx6_{L318}$ and $ScBx7_{L318}$ genes in positive BAC clones. SoftBerry/FGENESH/Monocot plants (generic, corn, rice, wheat, barley), (https://linux1.softberry.com/), (Solovyev

2007) software was employed to determine the structure of the analysed genes. The I-TASSER program was used to create a structure model of proteins encoded by the $ScBx6_{L318}$, $ScBx6_{Lo7}$, $ScBx6\text{-}like_{PICASSO}$, $ScBx7_{L318}$, and $ScBx7_{Lo7}$ genes (https://zhanglab.ccmb.med.umich.edu/I-TASSER/), (Zhang 2008; Roy et al. 2010; Yang et al. 2015). The cDNA sequences were changed into amino acid sequences and subjected to I-TASSER analysis. Out of the five models, the greatest scoring I-TASSER model was chosen based on the C-score.[1] Phylogenetic trees were constructed based on the *cds* of genes encoding dioxygenases and *O*-methyltransferases from the NCBI and ENA databases. These were created with the use of the MEGA6 program (Tamura et al. 2013), Neighbor Joining algorithm (Saitou and Nei 1987), and the Maximum Composite Likelihood method (Nei and Kumar 2000), with a bootstrap (Felsenstein 1985) value of 1000. The other applications applied in the bioinformatic analysis were:

- BLASTN 2.7.1+ (Zhang et al. 2000)—to determine the localization of the $ScBx6_{L318}$ and $ScBx7_{L318}$ genes on rye chromosomes
- 'Clustal Omega'—for a sequence comparison to find the SNP and INDEL polymorphisms (https://www.ebi.ac.uk/Tools/msa/clustalo/)
- 'EMBOSS Transeq'—to rewrite the nucleotide sequence into an amino acid sequence (https://www.ebi.ac.uk/Tools/st/emboss_transeq/)
- ENA database (www.ebi.ac.uk/ena)—to find the *Bx6* and *Bx7* genes in wheat and to analyse the rye genome sequence
- NCBI database (https://ncbi.nlm.nih.gov/)—to analyse the rye genome sequence and to find the *Bx6* and *Bx7* genes in line Lo7
- NCBI database blast algorithm—to find highly similar sequences to the analysed sequence and to determine the degree of similarity between the compared nucleotide and the amino acid sequences (https://blast.ncbi.nlm.nih.gov/Blast.cgi)
- PCRPrimerStats program (Stothard 2000)—to assess the quality of the constructed primers (the presence of so-called "hairpins")
- 'PlantCare'—for an analysis of the $ScBx6_{L318}$ and $ScBx7_{L318}$ gene promoters available within 749 bp upstream from the ATG codon (Lescot et al. 2002), (https://bioinformatics.psb.ugent.be/webtools/plantcare/html/)

- Primer3 v. 0.4.0 (Koressaar and Remm 2007; Untergasser et al. 2012)—for the primer design
- 'showseq'—to determine the sequence length (https://emboss.bioinformatics.nl/cgi-bin/emboss/showseq)
- the application is available at: https://eu.idtdna.com/calc/analyzer—to assess the tendency to create homo- and hetero-dimers in the constructed primers

## Results

Searching through the BAC library allowed the identification of the $ScBx6_{L318}$ and $ScBx7_{L318}$ genes in two separate superpools for each gene. The plate, row, and columnar pools from four superpools were searched to determine the coordinates of the clones containing the studied genes. An analysis of the results determined on which plates the clones containing the tested genes were located.

### Structure and localization of the *ScBx6* and *ScBx7* genes

Based on the bioinformatic analysis, the length and structure of the $ScBx6_{L318}$ and $ScBx7_{L318}$ genes were predicted. The total length of $ScBx6_{L318}$ was 1558 bp and that of the coding sequence was 1125 bp. It comprised three exons (with lengths 510 bp, 327 bp, and 288 bp), two introns (143 bp and 99 bp), and a 3′ UTR (191 bp), (GenBank Acc. No. MG516219, supplementary materials, Fig. IA). The total length of $ScBx7_{L318}$ was 1836 bp, and that of the coding sequence was 1098 bp. The gene was composed of 2 exons (with lengths 789 bp and 309 bp), one intron (664 bp), and a 3′ UTR (74 bp), (GenBank Acc. No. MG519859, supplementary materials, Fig. IB). Both genes were found to have a great similarity to their orthologs from maize (supplementary materials, Table A1 and A2). An analysis of the rye genome (GCA_900002355.1 version 2017/02/08; Bauer et al. 2017) from the NCBI database enabled the location of possible orthologs of both genes in line Lo7 (from Lo7_v2_contig_75580 and Lo7_v2_contig_1354237, GenBank Acc. No. FKKI010054593.1 and FKKI010538798.1, respectively) and to predict their structure. The number of structural components in both genes was the same in both lines (three exons, two introns, and two exons and one intron, respectively) however, in genes from line L318, a 12 bp deletion in the first exon ($ScBx6_{L318}$) and a 2 bp insertion in the intron ($ScBx7_{L318}$) were found, in comparison to the Lo7 genes. Based on comparative analysis, 22 SNPs and 1 INDEL for the *ScBx6* accessions and 28 SNPs and 2 INDELs for the *ScBx7* accessions were identified (supplementary materials, Tables B and C). Polymorphisms located in the exons led to a few amino acid substitutions and caused length and

---

[1] The C-score is a confidence score used to estimate the quality of the predicted models by I-TASSER. It was calculated based on the significance of the reading template alignments and the convergence parameters of the structure assembly simulations.
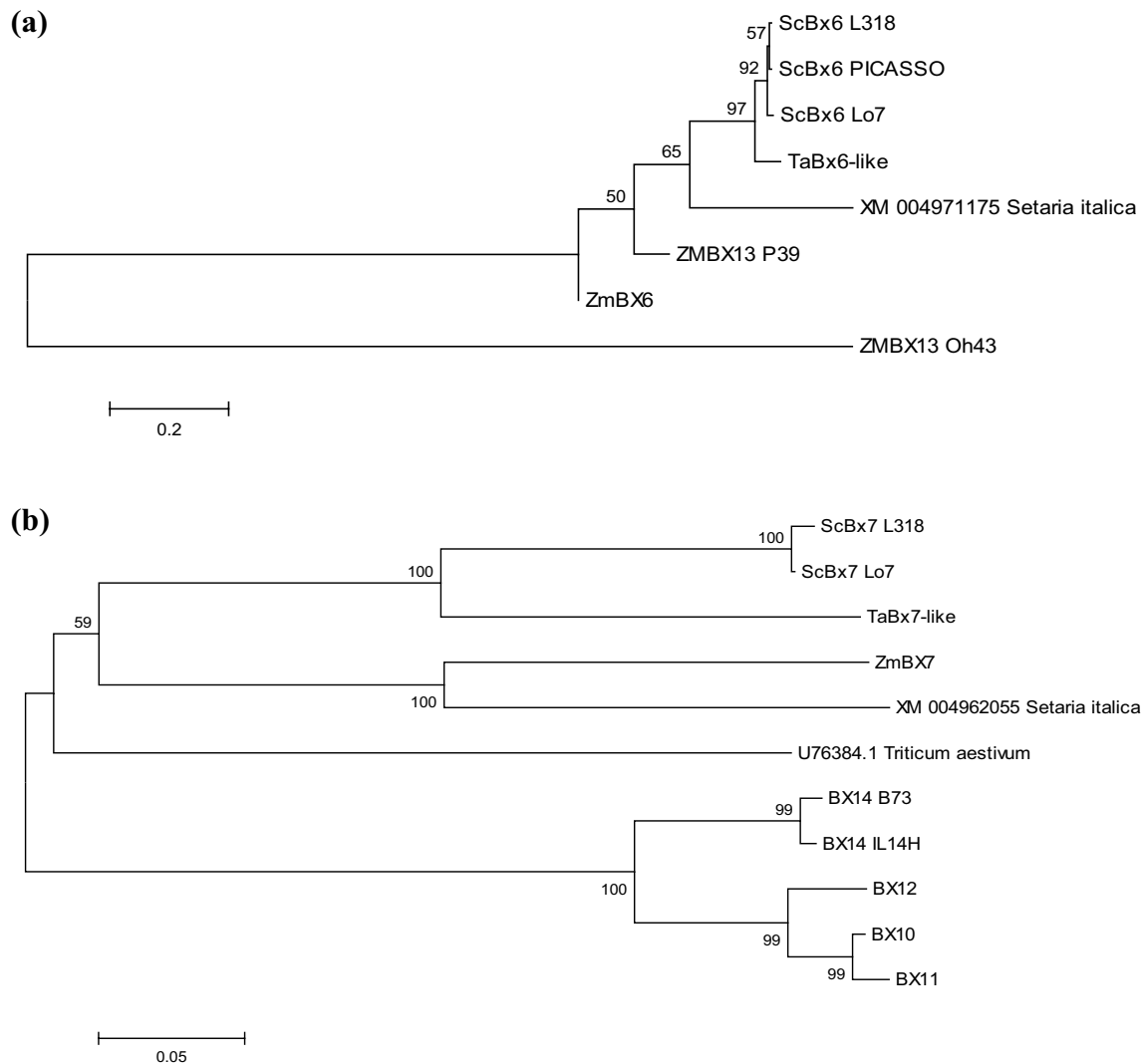
**(a)**



**(b)**



**Fig. 1** **a** Phylogenetic tree of dioxygenases sequences: *ZmBx6* (GenBank accession no. AF540907), *ZmBx13* Oh43 (GenBank accession no. KU521786), *ZmBx13* P39 (GenBank accession no. KU521787), *ScBx6*$_{L318}$ (GenBank accession no. MG516219), *ScBx6-like*$_{PICASSO}$ (GenBank accession no. HG380520), *TaBx6-like* (ENA accession no. CBD24305), and a gene from *S. italica* encoding dioxygenase (GenBank accession no. XM004971175); **b** phylogenetic tree of *O*-methyltransferase sequences: *ZmBx7* (GenBank accession no. EU192149), *ZmBx10* (GenBank accession no. KC754962), *ZmBx11* (GenBank accession no. KC754963), *ZmBx12* (GenBank accession no. KC754964), *ZmBx14* B73 (GenBank accession no. KU521788), *ZmBx14* IL14H (GenBank accession no. KU521789), *ScBx7*$_{L318}$ (GenBank accession no. MG519859), *TaBx7-like* gene (GenBank accession no. AK330822), a wheat gene encoding *O*-methyltransferase (GenBank accession no. U76384.1), and a gene from *S. italica* encoding *O*-methyltransferase (GenBank accession no. XM0049602055). Both phylogenetic trees were generated by Mega 6 software (Tamura et al. 2013) based on the Maximum Composite Likelihood method (Nei and Kumar 2000). The bootstrap values are indicated at the branch points. Scale bar indicates number of substitutions per site

arrangement mismatches, which had an influence on the structure of the encoded proteins. The protein structure of both SCBX6$_{L318}$ and SCBX6$_{PICASSO}$ had one less α-helix in relation to the protein model encoded by the *ScBx6*$_{Lo7}$ gene. Moreover, the SCBX7$_{L318}$ protein model contained differences in the arrangement of two α-helixes and two β-helixes in comparison to SCBX7$_{Lo7}$ (supplementary materials, Figs. IIA and IIB).

An alignment of *ScBx* sequences to the rye genome (Bauer et al. 2017) allowed localization of the *ScBx6*$_{L318}$ and *ScBx7*$_{L318}$ genes on chromosomes 7R and 5R, respectively.

**Table 2** Stress-specific motifs found in regulatory sequences of the $ScBx6_{L318}$ and $ScBx7_{L318}$ genes

| Gene name | Promoter sequence length (bp) | Analyzed sequence length (bp) | Stress-specific motifs/ no. of a given motif | Function | Frequency (no. of stress-specific motifs/100 bp) |
|---|---|---|---|---|---|
| $ScBx6$ L318 | 749 | 749 | CGTCA/1 | cis-Acting regulatory element involved in the MeJA-responsiveness | 1.47 |
| | | | TGACG/1 | | |
| | | | CGCACGTGTC/1 | cis-Acting element involved in the abscisic acid responsiveness | |
| | | | TACGGTC/1 | | |
| | | | ACGTG/4 | | |
| | | | CACGTG/1 | | |
| | | | CCGAAA/1 | cis-Acting element involved in low-temperature responsiveness | |
| | | | GTTTTCTTAC/1 | cis-Acting element involved in defence and stress responsiveness | |
| $ScBx7$ L318 | 3388 | 749 | CCGAAA/1 | cis-Acting element involved in low-temperature responsiveness | 0.27 |
| | | | CAACTG/1 | MYB binding site involved in drought-inducibility | |

## Phylogenetic analysis of *Bx* genes

Two phylogenetic trees were constructed based on the *cds* of genes encoding dioxygenases (Fig. 1a) and *O*-methyltransferases (Fig. 1b).

The $ScBx6_{L318}$, $ScBx6_{Lo7}$, and $ScBx6$-like$_{PICASSO}$ genes were compared to three genes from maize and one gene from *S. italica* encoding dioxygenases. Additionally, one wheat gene (Acc. No. CBD24305, ENA database), designated by us as the *TaBx6-like* gene, was included in the analysis. The homology range for the *TaBx6-like* gene to other *Bx6* genes, in *cds* and at the AA level, was 77–93% and 67–92%, correspondingly (supplementary materials, D1–D2). On the resulting phylogenetic tree, the $ScBx6_{L318}$, $ScBx6_{Lo7}$, and $ScBx6$-like$_{PICASSO}$ genes along with the wheat *TaBx6-like* gene were distributed in two strongly supported branches (bootstrap values of 92 and 97%). The remaining part of the tree was composed of three branches, with more distant sequences of dioxygenases from *S. italica* and maize.

The $ScBx7_{L318}$ and $ScBx7_{Lo7}$ genes were compared to nine genes encoding *O*-methyltransferases, (i.e., six genes from maize, two genes from wheat, and one gene from *S. italica*). A wheat gene named by us as *TaBx7-like* (GenBank Acc. No. AK330822) showed 76% identity in *cds* and 64% at the AA level to both the $ScBx7_{L318}$ and $ScBx7_{Lo7}$ genes and 69% and 44% to *ZmBx7*, respectively (supplementary materials, E1–E2). The analysed genes were grouped in four strongly supported (with bootstrap values of 100%) clusters. An apparent separation of $ScBx7_{L318}$ and $ScBx7_{Lo7}$ together with the *TaBx7-like* gene from the remaining three clusters was observed. Furthermore, one from the remaining three clusters, which contained gene from *S. italica* and *ZmBx7*,

was the least distant from $ScBx7_{L318}$, $ScBx7_{Lo7}$ and *TaBx7-like* gene cluster.

## In silico analysis

To confirm the function of the newly isolated genes of line L318, a bioinformatic analysis of the 749 bp 5′ regulatory sequence, including the core promoter, proximal and distal regulatory elements, was performed. In both genes, stress-specific motifs (SSM) were identified. For $ScBx6_{L318}$ eleven SSM were found: cis-acting regulatory elements involved in MeJA-responsiveness (2), cis-acting elements involved in abscisic acid responsiveness (7), cis-acting element involved in low-temperature responsiveness (1), and cis-acting element involved in defence and stress responsiveness (1). In turn, for $ScBx7_{L318}$ only two SSM were identified: cis-acting element involved in low-temperature responsiveness (1), and MYB binding site involved in drought-inducibility (1), (Table 2).

## Discussion

Since the discovery of BXs in 1955 (Virtanen and Hietala 1955a, b), information regarding genes encoding enzymes involved in the biosynthetic pathway of these compounds has gradually extended. Although BXs were identified first in rye, the genetic background of biosynthesis of these compounds was practically unknown until 2013, with the exception of the *ScBx1-ScBx2* (La Hovary 2011), *Scglu* (Nikus et al. 2003), and *ScGT* (Sue et al. 2011) genes. After 2013, new *Bx* genes as well as the updated version of *Bx* sequences that were known previously were published. Five *ScBx* genes

(*ScBx1*–*ScBx5*) of rye inbred line L318 were isolated and sequenced (Bakera et al. 2015). Recently, Tanwir et al. (2017) published the coding sequences of the *ScBx1*–*ScBx6* genes from cultivar PICASSO. The present study led to the isolation and prediction of the structure and function of two consecutive rye genes controlling BXs biosynthesis, namely *ScBx6* and *ScBx7*, which are orthologs of *ZmBx6* and *ZmBx7* in maize, respectively.

According to Frey et al. (2003), the orthologs of the *Bx6* and *Bx7* genes were lost in rye and wild barley species, which in turn caused the dominant aglucon in these species to become DIBOA, whereas in maize and wheat it became DIMBOA. However, Frey et al. did not present an alternative pathway for the synthesis of DIMBOA, which was proven by many authors to be present in rye (e.g., Rice et al. 2005; Zasada et al. 2007; Meyer et al. 2009; Rakoczy-Trojanowska et al. 2017). To date, TRIBOA-Glc, an intermediate product between DIBOA-Glc and DIMBOA-Glc, has not been found in rye, which some researchers (Frey et al. 2003) saw as proof that the BXs biosynthetic pathway in rye ends on BX5. It can be hypothesised that in rye the synthesis level of TRIBOA-Glc is much less than in maize, which makes its detection difficult.

The $ScBx6_{L318}$ and $ScBx6_{Lo7}$ genes, based on the bioinformatic prediction, contained three exons and two introns. However, based on Frey et al. (2003), and in accordance with our results of the maize genome analysis deposited in the NCBI database (GCF_000005005.2, 6.09.2018), the *ZmBx6* gene has no introns. The *ZmBx6* gene seemed to be unique in the *Bx6* ortholog group given that the other *Bx* genes have at least one intron. On the basis of our comparative analysis of these sequences and the genomes of several Poaceae species, it was determined that *Bx6*-like genes in *Ae. tauschii* and *S. italica* had two introns, and in *S. bicolor* none (data not shown). The *cds* of $ScBx6_{L318}$, $ScBx6_{PICASSO}$, and *ZmBx6* were identical in length (1125 bp), in contrast to $ScBx6_{Lo7}$, in which the *cds* was 12 bp longer.

Despite the very great similarity of $ScBx6_{L318}$ to the $ScBx6$-$like_{PICASSO}$ (99%) and $ScBx6_{Lo7}$ (98%) genes in terms of *cds*, significant differences between their protein models were detected. An insertion in the first exon of $ScBx6_{Lo7}$ (374 aa to 378 aa) resulted in one additional α-helix in relation to $ScBx6_{L318}$ and ScBx6-LIKE$_{PICASSO}$. However, the cause for differences in the arrangement of four α-helices and two β-helices were most likely associated with two non-conservative substitutions ($Ser_{L318}/Arg_{PIC}$ and $Ala_{L318}/Thr_{PIC}$) within the third exon of $ScBx6$-$like_{PICASSO}$ and four and two non-conservative substitutions in the first ($Met_{L318}/Leu_{Lo7}$, $Gly_{L318}/Arg_{Lo7}$, $Val_{L318}l/Thr_{Lo7}$, $Val_{L318}l/Thr_{Lo7}$), and the third ($Glu_{L318}/Asp_{Lo7}$, $Val_{L318}/Leu_{Lo7}$) exon, respectively, in the case of $ScBx6_{Lo7}$. Tanwir et al. (2017) demonstrated that the ScBx6-LIKE$_{PICASSO}$ enzyme allows for the conversion of DIBOA-Glc to TRIBOA-Glc. Despite

the structural differences and chemical properties of the ScBx6-LIKE$_{PICASSO}$ and ZMBX6 proteins, they must have the same function, which suggests that the transformation of DIBOA-Glc into TRIBOA-Glc takes place in both rye and maize (Tanwir et al. 2017). The great identity in *cds* and at the AA sequence level indicates that the $ScBx6_{L318}$ and $ScBx6_{Lo7}$ proteins have the same catalytic properties as in the PICASSO cultivar. It can be assumed that the orthologs of the *Bx6* and *Bx7* genes that control the BXs biosynthetic pathway in maize are also present in rye.

Based on the bioinformatic prediction, $ScBx7_{L318}$ $ScBx7_{Lo7}$ both have two exons and one intron. Blast analysis of the *ZmBx7* gene sequence (GenBank Acc. No. EU192149) against the maize genome in the NCBI database (GCF_000005005.2, 6.09.2018) proved that the *ZmBx7* gene also had the same structure (two exons and one intron). The gene *cds* in both rye lines was 1098 bp, while for *ZmBx7* it was 1161 bp. Both the $ScBx7_{L318}$ and $ScBx7_{Lo7}$ genes showed 72% identity to the *ZmBx7* gene at the *cds* level. The great resemblance of the coding sequence, its length, and the same structure of the $ScBx7_{L318,}$ $ScBx7_{Lo7}$, and *ZmBx7* genes allowed us to assume that these genes are orthologs and have the same function. The catalytic properties of BX7 have been known so far only in maize. Both the $ScBx7_{L318}$ and $ScBx7_{Lo7}$ genes encoded a protein 26 amino acids less than ZMBX7 (365 vs 391), and showed a similarity of 45% at 99% coverage. The SCBX7$_{L318}$ and SCBX7$_{Lo7}$ protein models had differences in the arrangement of two α- and β-helixes, the source of which were six SNPs associated with non-conservative substitutions ($Ala_{L318}/Pro_{Lo7}$, $Thr_{L318}/Ala_{Lo7}$, $Ser_{L318}/Asn_{Lo7}$, $Lys_{L318}/Arg_{Lo7}$, $Asn_{L318}/Lys_{Lo7}$, $Thr_{L318}/Ser_{Lo7}$) within the first exon. Structural elements such as elongated and/or additional α-helixes could be related to differences in activity and enzymatic efficiency in relation to their orthologs, and requires further research.

The phylogenetic analysis of the *cds* of genes encoding dioxygenases revealed that the $ScBx6_{L318}$, $ScBx6_{Lo7}$, $ScBx6$-$like_{PICASSO}$, and *TaBx6-like* genes are orthologs of *ZmBx6*, although in the case of the *TaBx6-like* genes more experimental data will be necessary to determine their function.

The phylogenetic analysis of the *cds* of genes encoding *O*-methyltransferases deposited in the NCBI database and those obtained in this study revealed that the wheat gene designated as *TaBx7-like* was located within the most distant cluster encompassing the $ScBx7_{L318}$ and $ScBx7_{Lo7}$ genes. This gene was similar to both the $ScBx7_{L318}$ and $ScBx7_{Lo7}$ genes, as well as to the other analyzed genes encoding *O*-methyltransferases. Therefore, it can be assumed that *TaBx7-like* encodes *O*-methyltransferase and, presumably, it controls the conversion of TRIBOA-Glc to DIMBOA-Glc in wheat. The second wheat gene (GenBank Acc. No. U76384.1) included in the analysis was located on a separate branch than the rest of the genes which encode

*O*-methyltransferases, and was not as strongly supported as others (59 vs. 99–100%). Moreover, it showed a lesser level of similarity, both on the *cds* and at the AA level, to the $ScBx7_{L318}$, $ScBx7_{Lo7}$, and *ZmBx7* genes (supplementary materials, F1–F2). Therefore, this gene most likely is not involved in BXs biosynthesis in wheat. On the basis of the obtained results, it can be supposed that the newly isolated $ScBx7_{L318}$ gene codes for *O*-methyltransferase, which takes part in the BXs biosynthesis pathway in rye, analogously to *ZmBx7* in maize. However, it will be necessary to study the enzyme's catalytic properties and its ability to catalyse TRIBOA-Glc and TRIMBOA-Glc conversion to DIMBOA-Glc and $DIM_2BOA$-Glc, respectively, to confirm the hypothesis formulated above.

To predict the probable functions of the newly isolated $ScBx6_{L318}$ and $ScBx7_{L318}$ genes, we have performed the promoter in silico analysis of their upstream regulatory sequences. Currently, the analysis of 5′ cis regulatory sequences is often used to gain information regarding the function of the investigated genes (e.g., Guo et al. 2017; Wang et al. 2017; Nawkar et al. 2017; Charfeddine et al. 2017). By applying such analysis, Guo et al. (2017) proved that the protein encoded by the *BplMYB46* gene, which binds to the MYBCORE and AC-box motifs, plays an important role in secondary cell wall synthesis and response to abiotic stress in *Betula platyphylla*. The identification of three cis GT-1 elements in the promoter sequence of the *Cucumis sativus CsSAMs* gene proved its role in response to elevated salinity stress (Wang et al. 2017).

In the promoter sequences of the $ScBx6_{L318}$ and $ScBx7_{L318}$ genes several stress-specific motifs were identified, which clearly indicates their roles in plant defence. However, cis-acting regulatory elements associated with abscisic acid (ABA) responsiveness, defence and stress responsiveness, and MeJA-responsiveness were found in *ScBx6* uniquely, when the MYB binding site associated with drought-inducibility was identified in the 5′ upstream region of *ScBx7* exclusively. MeJA is responsible for launching systemic resistance mechanisms. After wounding or contact of the plant with the pathogen, a significant increase in the MeJA concentration is observed (Rates et al. 2009). Part of the response to pathogens triggered by MeJA is the accumulation of secondary metabolites. The expression of genes involved in secondary metabolism is increased when the primary metabolism is inhibited (Paudel et al. 2014). According to Schaller et al. (2005) MeJA leads to degradation of chlorophyll a and b and inhibits the activity of ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO). The chlorophyll decomposition process is accompanied by lesser photosynthetic efficiency. This regulator stimulates photorespiration, peroxidise activity, and is involved in stomatal closure (Schaller et al. 2005). The motifs connected with the MeJA signalling pathway were present in the majority of *ScBx* promoter sequences cloned and analysed to date, with the exception of $ScBx3_{L318}$ (characterized previously by Bakera et al. 2015) and $ScBx7_{L318}$ (described in this work). ABA, another important hormone, affects tolerance to many abiotic stresses, such as the presence of heavy metals, drought, high temperature, coldness, elevated salinity, or irradiation, and the amount of this regulator increases in response to stress occurrence (Vishwakarma et al. 2017). Other functions exhibited by ABA and mentioned by Vishwakarma et al. (2017) are inhibition of seed germination and impact on stomatal movement. In the $ScBx6_{L318}$ gene 7 SSM associated with the ABA signalling pathway (ABRE) were found. The frequency of those SSM were greater in comparison to that located in regulatory sequences of other *Bx* genes from the Poaceae family and ranged from 1.75 up to 7.0-fold. The number of all motifs differed as well, and showed the greatest value in the $ScBx6_{L318}$ gene and ranged from 1.83 up to 11-fold (supplementary materials, Table G-S). Interestingly, no ABRE motifs were identified in the $ScBx1_{L318}$–$ScBx5_{L318}$ genes (Bakera et al. 2015) and in the $ScBx7_{L318}$ gene (described in this work). Zhang et al. (2017) reported the occurrence of at least one drought stress-responsive cis-element (including i.e., a MBS—MYB binding site, involved in drought-inducibility and low-temperature response—LTR) in the promoter sequences of maize drought-responsive genes. The $ScBx7_{L318}$ gene had either SSM types, while in other analyzed genes only MBS (*Bx3* orthologs) or LTR ($ScBx6_{L318}$, *AsBx4* and *AtBx4* genes) or neither (*ZmBx* genes) motif have been found (supplementary materials, Table G–S). The presence of motifs related to the MeJA relay pathway and other important cis-acting elements in the promoter sequence of the $ScBx6_{L318}$ and $ScBx7_{L318}$ genes clearly indicates their role in rye defence strategy. Those results ensure us in the conviction that there is potential in, for example, selection of lines with great BXs content towards more resistant cultivars to abiotic and biotic stresses. Rakoczy-Trojanowska et al. (2017) showed that polymorphisms in *ScBx* gene sequences, especially the most valuable SNPs (*ScBx4*_1702 and *ScBx5*_1105) unaffected by environmental factors, were associated with pre-harvest sprouting resistance (PHS-R) and greater GDIMBOA, MBOA, and HBOA content, respectively.

The obtained results extended the knowledge of the genetic background of BXs biosynthesis in rye by adding the characteristics of the next two genes. We predicted their probable function and we have established their chromosome localization. Plants with the ability to synthesize BXs, from the Poaceae family, differ in regards to the number of involved genes controlling the BXs pathway. The fewest number is known in *H. lechleri* (5) and the greatest in maize (16). In three species (maize, wheat, and rye) all *Bx* genes are divided between four chromosomes, however, maize possesses four clusters and wheat and rye only 2. Based on our

results we postulate that the BXs biosynthesis pathway in wheat and rye could be analogous to that in maize, not only till DIBOA synthesis but in subsequent steps as well.

## Conclusions

1. The $ScBx6_{L318}$ and $ScBx7_{L318}$ rye genes are most probably the orthologs of the $ZmBx6$ and $ZmBx7$ genes, respectively.
2. Based on the in silico analysis of $Bx6$ and $Bx7$ orthologs in wheat and maize, it can be proposed that the DIMBOA-Glc biosynthetic pathway is similar in both species.

## Compliance with ethical standards

## References

Bakera B, Makowska B, Groszyk J, Niziołek M, Orczyk W, Bolibok-Brągoszewska H, Hromada-Judycka A, Rakoczy-Trojanowska M (2015) Structural characteristics of ScBx genes controlling the biosynthesis of hydroxamic acids in rye (*Secale cereale* L.). J Appl Genet 56(3):287–298

Bauer E, Schmutzer T, Barilar I, Mascher M, Gundlach H, Martis MM, Twardziok SO, Hackauf B, Gordillo A, Wilde P, Schmidt M, Korzun V, Mayer KFX, Schmid K, Schön C-C, Scholz U (2017) Towards a whole-genome sequence for rye (*Secale cereale* L.). Plant J 89(5):853–869

Charfeddine S, Charfeddine M, Saïdi MN, Jbir R, Bouzid RG (2017) Potato dehydrins present high intrinsic disorder and are differentially expressed under ABA and abiotic stresses. Plant Cell Tissue Organ 128(2):423–435

Chu HY, Wegel E, Osbourn A (2011) The plant genome: an evolutionary view on structure and function. From hormones to secondary metabolism: the emergence of metabolic gene clusters in plants. Plant J 66:66–79

Czjzek M, Cicek M, Zamboni V, Bevan DR, Henrissat B, Esen A (2000) The mechanism of substrate (aglycone) specificity in b-glucosidases is revealed by crystal structures of mutant maize b-glucosidase-DIMBOA, -DIMBOAGlc, and -dhurrin complexes. Proc Natl Acad Sci USA 97:13555–13560

Dick R, Rattei T, Haslbeck M, Schwab W, Gierl A, Frey M (2012) Comparative analysis of benzoxazinoid biosynthesis in monocots and dicots: independent recruitment of stabilization and activation functions. Plant Cell Online 24(3):915–928

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

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 defence mechanism in grasses. Science 277:696–699

Frey M, Huber K, Park WJ, Sicker D, Lindberg P, Meeley RB, Simmons CR, Yalpani N, Gierl A (2003) A 2-oxoglutarate-dependent dioxygenase is integrated in DIMBOA biosynthesis. Phytochemistry 62:371–376

Frey M, Schullehner K, Dick R, Fiesselmann A, Gierl A (2009) Benzoxazinoid biosynthesis, a model for evolution of secondary metabolic pathways in plants. Phytochemistry 70:1645–1651

Groszyk J, Kowalczyk M, Yanushevska Y, Stochmal A, Rakoczy-Trojanowska M, Orczyk W (2017) Identification and VIGS-based characterization of Bx1 ortholog in rye (*Secale cereale* L.). PLoS ONE 12(2):e0171506

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 benzoxazinoid biosynthesis genes from *Hordeum lechleri*. Phytochemistry 66(11):1264–1272

Guo H, Wang Y, Wang L, Hu P, Wang Y, Jia Y, Zhang C, Zhang Y, Zhang Y, Wang C, Yang C (2017) Expression of the MYB transcription factor gene *BplMYB46* affects abiotic stress tolerance and secondary cell wall deposition in *Betula platyphylla*. Plant Biotechnol J 15:107–121

Handrick V, Robert CAM, Ahern KR, Zhou S, Machado RAR, Maag D, Glauser G, Fernandez-Penny FE, Chandran JN, Rodgers-Melnik E, Schneider B, Buckler ES, Boland W, Gershenzon J, Jander G, Erb M, Köllner TG (2016) Biosynthesis of 8-*O*-methylated benzoxazinoid defence compounds in maize. Plant Cell 28(7):1682–1700

Jonczyk R, Schmidt H, Osterrieder A, Fiesselmann A, Schullehner K, Haslbeck M, Sicker D, Hofmann D, Yalpani N, Simmons C, Frey M, Gierl A (2008) Elucidation of the final reactions of DIMBOA-glucoside biosynthesis in maize: characterization of *Bx6* and *Bx7*. Plant Physiol 146:1053–1063

Koressaar T, Remm M (2007) Enhancements and modifications of primer design program Primer3. Bioinformatics 23(10):1289–1291

La Hovary C (2011) Allelochemicals in *Secale cereale*: biosynthesis and molecular biology of benzoxazinones. Ph.D. dissertation. North Carolina State University

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. Nucleic Acids Res 30(1):325–327

Makowska B, Bakera B, Rakoczy-Trojanowska M (2015) The genetic background of benzoxazinoid biosynthesis in cereals. Acta Physiol Plant 37:176

Meihls LN, Handrick V, Glauser G, Barbier H, Kaur H, Haribal MM, Jander G (2013) Natural variation in maize aphid resistance is associated with 2, 4-dihydroxy-7-methoxy-1, 4-benzoxazin-3-one glucoside methyltransferase activity. Plant Cell 25(6):2341–2355

Meyer SLF, Rice CP, Zasada IA (2009) DIBOA: fate in soil and effects on root—knot nematode egg numbers. Soil Biol Biochem 41:1555–1560

Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8(19):4321–4326

Nawkar GM, Maibam P, Park J-H, Woo SG, Kim CY, Lee SY, Kang CH (2017) In silico study on *Arabidopsis* BAG gene expression in response to environmental stresses. Protoplasma 254(1):409–421

Nei M, Kumar S (2000) Molecular evolution and phylogenetics. Oxford University Press, New York

Niemeyer HM (2009) Hydroxamic acids derived from 2-hydroxy-2H-1,4-benzoxazin-3(4H)-one: key defence chemicals of cereals. J Agric Food Chem 57:1677–1696

Nikus J, Esen A, Jonsson LMV (2003) Cloning of a plastidic rye (*Secale cereale*) b-glucosidase cDNA and its expression in *Escherichia coli*. Physiol Plant 118:337–345

Nomura T, Ishihara A, Imaishi H, Endo TR, Ohkawa H, Iwamura H (2002) Molecular characterisation and chromosomal localization of cytochrome P450 genes involved in the biosynthesis of cyclic hydroxamic acids in hexaploid wheat. Mol Genet Genom 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. PNAS 102(45):16490–16495

Oikawa A, Ishihara A, Iwamura H (2002) Induction of HDMBOA-Glc accumulation and DIMBOA-Glc4-*O*-methyltransferase by jasmonic acid in poaceous plants. Phytochemistry 61:331–337

Paudel S, Rajotte EG, Felton GW (2014) Benefits and costs of tomato seed treatment with plant defence elicitors for insect resistance. Arthropod Plant Interact 8:539–545

von Rad U, Huttl R, Lottspeich F, Gierl A, Frey M (2001) Two glucosyltransferases are involved in detoxification of benzoxazinoids in maize. Plant J 28:633–642

Rakoczy-Trojanowska M, Orczyk W, Krajewski P, Bocianowski J, Stochmal A, Kowalczyk M (2017) *ScBx* gene based association analysis of hydroxamate content in rye (*Secale cereale* L.). J Appl Genet 58:1–9

Rates SMK (2009) Plants as a source of drugs. Toxicon 39:603–613

Rice CP, Park YB, Adam F, Abdul-Baki AA, Teasdale JR (2005) Hydroxamic acid content and toxicity of rye at selected growth stages. J Chem Ecol 31(8):1887–1905

Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc 5:725–738

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Schaller F, Schaller A, Stintzi A (2005) Biosynthesis and metabolism of jasmonates. J Plant Growth Regul 23:179–199

Solovyev VV (2007) Statistical approaches in eukaryotic gene prediction. In: Balding D, Cannings C, Bishop M (eds) Handbook of statistical genetics, 3rd edn. Wiley-Interscience, Hoboken

Stothard P (2000) The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. Biotechniques 28:1102–1104

Sue M, Nakamura C, Nomura T (2011) Dispersed benzoxazinone gene cluster: molecular characterization and chromosomal localization of glucosyltransferase and glucosidase genes in wheat and rye. Plant Physiol 157:985–997

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

Tanwir F, Dionisio G, Adhikari KB, Fomsgaard IS, Gregersen PL (2017) Biosynthesis and chemical transformation of benzoxazinoids in rye during seed germination and the identification of a rye *Bx6-like* gene. Phytochemistry 140:95–107

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3—new capabilities and interfaces. Nucleic Acids Res 40(15):e115

Virtanen AI, Hietala PK (1955a) 2(3)-Benzoxazolinone an anti-fusarium factor in rye seedlings. Acta Chem Scand 9:1543–1544

Virtanen AI, Hietala PK (1955b) The structure of the precursors of benzoxazolinone in rye plants II. Suomen Kemistilehti 32:252

Vishwakarma K, Upadhyay N, Kumar N, Yadav G, Singh J, Mishra RK, Kumar V, Verma R, Upadhyay RG, Pandey M, Sharma S (2017) Abscisic acid signaling and abiotic stress tolerance in plants: a review on current knowledge and future prospects. Front Plant Sci 8:161

Wang L-W, He M-W, Guo S-R, Zhong M, Shu S, Sun J (2017) NaCl stress induces *CsSAMs* gene expression in *Cucumis sativus* by mediating the binding of *CsGT-3b* to the GT-1 element within the *CsSAMs* promoter. Planta 245(5):889–908

Wouters FC, Blanchett B, Gershenzo J, Vassão DG (2016) Plant defense and herbivore counter-defense: benzoxazinoids and insect herbivores. Phytochem Rev 15(6):1127–1151

Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y (2015) The I-TASSER suite: protein structure and function prediction. Nat Methods 12:7–8

Zasada IA, Rice CP, Meyer SLF (2007) Improving the use of rye (*Secale cereale*) for nematode management: potential to select cultivars based on *Meloidogyne incognita* host status and benzoxazinoid content. Nematology 9:53–60

Zhang T, Tan D, Zhang L, Zhang X, Han Z (2017) Phylogenetic analysis and drought-responsive expression profiles of the WRKY transcription factor family in maize. Agric Gene 3:99–108

Zhang Y (2008) I-TASSER server for protein 3D structure prediction. BMC Bioinform 9:40

Zhang Z, Schwartz S, Wagner L, Webb M (2000) A greedy algorithm for aligning DNA sequences. J Comput Biol 7(1–2):203–214