Research Article

Identification of novel cis-elements bound by BplMYB46 involved in abiotic stress responses and secondary wall deposition

Huiyan Guo¹ †, Liuqiang Wang² †, Chuanping Yang¹, Yiming Zhang¹, Chunrui Zhang¹, Chao Wang¹ *

1. State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University, Harbin 150040, China
2. State Key Laboratory of Tree Genetics and Breeding, Research Institute of Forestry, Chinese Academy of Forestry, Beijing 100091, China
† These authors contributed equally to this work
*Correspondence: Chao Wang (wangchao@nefu.edu.cn)

Edited by: Jianmin Wan, Institute of Crop Science, CAAS, China
Running Title: Identification of novel cis-elements bound by BplMYB46

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1111/jipb.12671]

This article is protected by copyright. All rights reserved
Received: February 19, 2018 / Accepted: June 1, 2018
Abstract

Transcription factors (TFs) play vital roles in various biological processes by binding to cis-acting elements to control expressions of their target genes. The MYB TF BplMYB46, from Betula platyphylla, is involved in abiotic stress responses and secondary wall deposition. In the present study, we used a TF-centered yeast one-hybrid technology (TF-centered Y1H) to identify the cis-acting elements bound by BplMYB46. We screened a short-insert random library and identified three cis-elements bound by BplMYB46: an E-box (CA(A/T/C)(A/G/C)TG) and two novel motifs, a TC-box (T(G/A)TCG(C/G)) and a GT-box (A(G/T)T(A/C)GT(T/G)C). Chromatin immunoprecipitation (ChIP) and effector-reporter coexpression assays in Nicotiana tabacum confirmed binding of BplMYB46 to the TC-box, GT-box, and E-box motifs in the promoters of the phenylalanine ammonia lyase (PAL), peroxidase (POD), and superoxide dismutase (SOD) genes, which function in abiotic stress tolerance and secondary wall biosynthesis. This finding improves our understanding of potential regulatory mechanisms in the response to abiotic stress and secondary wall deposition of BplMYB46 in B. platyphylla.

INTRODUCTION

In plants, adaptation to diverse environments involves complex and dynamic physiological, molecular, and biochemical changes that require modulation of the expression of multiple genes (Su et al. 2014). Identification of the cis-acting elements, bound by specific transcription factors, can reveal the transcriptional regulatory mechanisms and gene expression patterns involved in environmental adaptation, and further improve our understanding of the functions of transcription factors (TFs) in plants.

Ongoing studies have identified the targets of many TFs and characterized the pathways they affect. For instance, the TF BhWRKY1, from B. hygrometrica, participates in the resistance to dehydration by binding to the W-box of the galactinol synthase promoter (Wang et al. 2009). Many bZIP and MYB TFs function in the
regulation of auxin-regulated genes, via binding to the bZIP or MYB response elements (Berendzen et al. 2012). In rice (Oryza sativa), OsNAC2 negatively regulates plant height and flowering time, via binding to the promoters of OsEATB and OsKO2, which encode proteins that repress gibberellin biosynthesis in rice (Chen et al. 2015).

In Arabidopsis thaliana, the octadecanoid-responsive AP2/ERF-domain TF 47 plays a role in the biosynthesis of jasmonic acid and abscisic acid, via binding to the (NC/GT)CGNCCA cis-element in the promoter of abscisic acid insensitive 2 (Chen et al. 2016). In addition, ThERF1 from the halophyte Tamarix hispida regulates its target genes by binding to GCC-box, DRE, and TTG motifs in response to salt stress (Wang et al. 2015). By contrast, ThbHLH1 in T. hispida regulates expression of target genes, involved in stress tolerance, by binding to the G-box (CACGTG) (Ji et al. 2016).

Although the molecular mechanisms of some plant TFs involved in the regulation of stress responses, growth, and development of have been characterized, the interactions among these genes remain unclear, and the regulatory network of gene expression, in plants, remains to be fully elucidated. Therefore, the interaction between upstream TFs and their target cis-acting elements needs further study.

In plants, the MYB proteins constitute a large TF family that have diverse functions in growth and development, as well as responses to environmental stresses. These MYB family members regulate anthocyanin biosynthesis (Dubos et al. 2008), participate in the auxin response pathway (Shin et al. 2007), promote hypocotyl elongation (Kwon et al. 2013), increase the tolerance to drought and salt stresses (Xiong et al. 2014), and affect secondary wall biosynthesis (McCarthy et al. 2009; Yoon et al. 2015).

Among the cis-elements, previous research has demonstrated that the MYBCORE and AC-box motifs are crucial for the regulatory function of MYB TFs. For instance, rice OsC1-MYB enhances tolerance to dehydration, high salt, and ABA, via binding to the MYBCORE motif (C(G/A)GT(T/G)) in the flavonoid biosynthesis gene OsDfr, which is related to drought stress responses (Ithal and Reddy 2004). In Arabidopsis, AtMYB46 and AtMYB83 promote secondary wall biosynthesis, by binding to the AC-box (ACC(A/T)A(A/C)(T/C)) in the promoters of downstream TFs (Zhong and Ye 2012).
The development and application of new methods has enabled the identification of additional DNA motifs bound by MYB transcription factors. In Arabidopsis, AtMYB46 promotes the biosynthesis of the secondary wall, via binding to an 8-bp cis-element (AGTAGGTG) in the promoters of the cellulose biosynthesis genes CESA4, CESA7, and CESA8 (Kim et al. 2012). High-throughput expression profiling showed that the putative targets of OsMYBs from rice harbor MYB-binding cis-elements in the promoter region, including a light-responsive motif (MRE; AACCTAA) and flavonoid biosynthetic gene regulation-related motifs (MBSII; AAAAGTTAGTTA) (Smita et al. 2015). In sweet cherry (Prunus avium), PavMYB10.1 participates in anthocyanin biosynthesis and determines fruit skin color, via interaction with the AE-box light-responsive elements (AGAAACAA) in the promoter region of the anthocyanin biosynthesis genes PavANS and PavUFGT (Jin et al. 2016).

Although the interactions of MYB proteins and DNA motifs have been studied, the cis-acting elements recognized by MYBs are not yet completely understood. Moreover, an outstanding question is whether any additional motifs are bound by MYBs. Therefore, a comprehensive characterization of the DNA motifs bound by MYB TFs remains an important avenue of research.

In a previous study, we showed that BplMYB46 of the Asian white birch Betula platyphylla, an important economically hardwood tree species that is used primarily for lumber, building and furniture industries, could enhance stress tolerance and promote secondary cell wall biosynthesis in transgenic plants (Guo et al. 2017). BplMYB46 acts by regulating the expression of its target genes, via binding to MYBCORE (CAGTTA) and AC-box (ACCACCT) motifs in their promoters (Guo et al. 2017). However, it is not clear whether BplMYB46 can also bind to other cis-acting elements. Here, we used the TF-centered Y1H to investigate the interaction of BplMYB46 with various DNA motifs. We identified an E-box and two novel DNA motifs bound by BplMYB46, and further confirmed these by promoter-reporter and chromatin immunoprecipitation (ChIP) assays. Thus, this study further elaborated the transcriptional regulatory mechanism underlying BplMYB46 function in response to abiotic stress and in the regulation of secondary wall deposition.
RESULTS

Identification of DNA motifs recognized by BplMYB46

To identify the DNA motifs recognized by BplMYB46, we carried out TF-centered Y1H by constructing a short-insert library of random DNA sequences and screening it with an effector construct containing the BplMYB46 full-length open reading frame. Five positive clones were identified that could grow on the selective synthetic defined medium SD/-His/-Leu/-Trp (TDO) with 3-amino-1, 2, 4-triazole (3-AT, 30 mM). These clones were further selected on high-stringency TDO medium/3-AT (50 mM), with three displaying high binding affinities to BplMYB46.

To identify the BplMYB46 binding motifs, we then sequenced the inserts of these positive clones. Since the two flanking sequences of the insertion might also be part of the cis-acting element (motif) recognized by the TF, two bases of each flanking sequence, together with the insertion sequence, were used for further analysis. The insertion DNA sequences were analyzed by PLACE and PlantCARE, as shown in Table 1. The short DNA sequence “CCAAATGAGCGG” (the underlined bases referred to the two flanking sequences from pHIS2) contained the E-box element “CAAATG”. The other insertion sequences “CCTGTCGCCCGG” and “CCAGTAGTTCGG” did not match any known motif, and represent candidate novel DNA motifs bound by MYB that have not been identified previously.

Core sequence analysis of the two novel motifs bound by BplMYB46

Y1H assays were used to determine the core sequence of the two novel motifs (Figure 1). On TDO medium with 50 mM 3-AT, cells containing the reporter construct with the element having the third DNA base (T) missing from the left border of novel motif 1 (CCTGTCGCCCGG) could not grow, indicating that BplMYB46 bound to the sequence “TGTCGCCCGG”, but not to “GTCGCCCGG”. Moreover, cells with the reporter lacking the fifth DNA base (C) from the right border did not grow, suggesting that BplMYB46 can bind to the sequence “CCTGTCGC”, but not to “CCTGTCG”
These observations demonstrated that the third DNA base (T) from the left border and the fifth DNA base (C) from the right border of novel motif 1 are essential for binding by BplMYB46. Therefore, we named the core sequence “TGTCGC” of novel motif 1 the TC-box.

Similarly, BplMYB46 could bind to the sequence “AGTAGTTCCGG”, but failed to bind to “GTAGTTCCGG” (Figure 1C), illustrating that the third DNA base (A) from the left border of novel motif 2 (CCAGTAGTTCCGG) was essential. In addition, BplMYB46 bound to the sequence “CCAGTAGTTTC” but not “CCAGTAGTTT”, suggesting that the third DNA base (C) from the right border of novel motif 2 was mandatory for binding by BplMYB46. These observations demonstrated that the core sequence of this motif was “AGTAGTTTC”, and we named it the GT-box.

**Specificity analysis of the three motifs bound by BplMYB46**

To define the specificity of the two novel motifs, bound by BplMYB46, we made single-base mutations for each DNA base in the core sequences of the TC-box and GC-box motifs, and tested BplMYB46 binding using the Y1H method. Each DNA base was substituted, individually, by the other three bases to generate a total of 18 and 24 different combinations for the TC-box and GT-box, respectively. The constructs with a substitution of the second (G) with an “A”, or the sixth (C) with a “G”, retained the same binding efficiency, as shown by growth on TDO medium with 50 mM 3-AT, when compared with the original TC-box sequence. Conversely, the other base substitutions resulted in the absence of growth on the TDO medium with 50 mM 3-AT, demonstrating that the consensus TC-box sequence for BplMYB46 binding was “T(G/A)TCG(C/G)” (Figure 2A).

For the GT-box, the substitutions of the second (G) with a “T”, the fourth (A) with a “C”, and the seventh (T) with a “G”, retained a similar efficiency as the original GT-box sequence. Conversely, the other base substitutions resulted in no growth on the TDO medium with 50 mM 3-AT, revealing that the consensus GT-box sequence for BplMYB46 binding was “A(G/T)T(A/C)GT(T/G)C” (Figure 2B).

The CAAATG of the “CCAAATGAGCGG” sequence bound by BplMYB46 in the
Y1H assay (Table 1) was annotated as the E-box, which has a canonical sequence of CANNTG (N: A/T/C/G). To determine the binding specificity of BplMYB46 to different E-box sequences, the third and fourth DNA bases in “CAAATG” were replaced with “T/C/G”, and the resulting constructs were assessed by Y1H assay. The results indicated that the substitutions of the third (A) with “T” or “C” and the fourth (A) with “G” or “C”, retained the same efficiency as the original “CAAATG” sequence (Figure 3). Thus, the consensus sequence for BplMYB46 binding was determined to be “CA(A/T/C)(A/G/C)TG”, a traditional E-box sequence.

To further confirm the relevance of BplMYB46 binding to different motifs, three tandem copies of E-box, GT-box, and TC-box, with their corresponding mutants, were ligated into pCAMBIA-1301 and used as effectors, with pROKII-BplMYB46 as the reporter (Figure 4A). The effectors with reporters were co-transformed into tobacco leaves and analyzed for GUS activity and staining. Firefly luciferase (LUC) was used as the control for normalization of the transformation efficiency. The GUS/LUC activity ratios of BplMYB46 with the E-box, GT-box, and TC-box were significantly higher than with the corresponding mutated forms (>10-fold) (Figure 4B). In accordance, the GUS staining was much stronger for binding of BplMYB46 to the E-box, GT-box, and TC-box motifs, as compared to the mutants, further demonstrating that BplMYB46 can bind to the E-box, GT-box, and TC-box motifs.

**Interaction between BplMYB46 and promoters containing the E-box, GT-box, and TC-box elements**

Our previous study demonstrated that BplMYB46 could regulate the expression of *PAL*, *POD* and *SOD*, which are associated with resistance to stress and with secondary wall biosynthesis in plants. Therefore, we speculated that BplMYB46 might bind to the promoters of these genes. Indeed, we detected E-box, GT-box, and TC-box motifs in the *PAL*, *POD*, and *SOD* promoter regions, prompting us to evaluate the specific interaction between BplMYB46 and these motifs in these three genes.

To test for this interaction, the truncated promoter regions containing only one kind of the E-box, GT-box, or TC-box, and truncated promoters lacking these three
motifs and also lacking the MYBCORE and AC-box, were cloned into pHIS2 and tested for interaction with BplMYB46 using the Y1H assay. BplMYB46 bound to the truncated promoters with the E-box, GT-box, or TC-box; however, it failed to bind to the truncated promoters without these three motifs (Figure 5).

To further substantiate the interaction between BplMYB46 and the promoters of the target genes, we next tested the interaction, in planta, with a GUS reporter assay. We co-transformed pROKII-BplMYB46 and the truncated promoters of PAL with the E-box, POD with the GT-box, and SOD with the TC-box into tobacco. Truncated promoters, without the motifs, were designated as controls (Figure 6A). The GUS/LUC activity ratio was much higher in the transformed lines harboring the E-box, GT-box, and TC-box than the control, thereby indicating that BplMYB46 could bind to the E-box, GT-box, and TC-box in the promoters of PAL, POD, and SOD (Figure 6B).

**Distribution of E-box, GT-box, and TC-box motifs in promoters of genes regulated by BplMYB46**

The BplMYB46-overexpressing lines showed significantly induced expression of 196 genes, compared with BplMYB46-silenced lines, as assessed by RNA-seq. We analyzed the distribution of the E-box, GT-box, and TC-box motifs in the promoters (−1 to −1,500 bp) of these genes (Table S1) and determined that the promoter regions of 98%, 61%, and 84% of the target genes harbored the E-box, GT-box, and TC-box motifs, respectively. The maximum number of repetitions of the E-box, GT-box, and TC-box motifs in the promoter of a gene was 9, 3, and 8, respectively. These results suggested that E-box, GT-box and TC-box motifs might play critical roles in the regulation of gene expression by BplMYB46.

**Interaction of BplMYB46 with E-box, GT-box, and TC-box occurs, in planta**

The binding between BplMYB46 and the specific motifs, in planta, was analyzed using ChIP, wherein the 35S::BplMYB46-GFP construct was transformed into birch (B. platyphylla) plants and the promoter fragments bound by BplMYB46 were captured by ChIP using a GFP antibody. ChIP-PCR results showed that the promoter fragments,
including the E-box, GT-box, and TC-box motifs, could be enriched by ChIP, using the GFP antibody (Figure 7). However, the truncated promoters lacking the E-box, GT-box, and TC-box motifs were not enriched by ChIP with the GFP antibody (ChIP+), compared to the positive control input and the negative control (ChIP-) (Figure 7B). ChIP-qPCR showed that the promoter regions of PAL, POD, and SOD, including the E-box, GT-box, and TC-box motifs, were significantly enriched in ChIP+, as compared to ChIP- (Figure 7C). These results demonstrated that BplMYB46 could regulate the expression of PAL, POD, and SOD genes by binding to the E-box, GT-box, and TC-box motifs.

**In vivo analysis of E-box, GT-box, and TC-box in response to abiotic stress**

To analyze the roles of the E-box, GT-box, and TC-box elements in the response to stress, we co-transformed the effector vector 35S::BplMYB46 and different reporters into tobacco and subjected leaves to salt or osmotic stress (200 mM NaCl for salt stress or 200 mM mannitol for osmotic stress). Transgenic tobacco leaves were collected at different time points, and the GUS/LUC activity ratio of constructs containing the E-box, GT-box, and TC-box were determined (Figure 8).

The GUS/LUC activity ratios of E-box constructs, at different treatment times, did not differ from controls as a result of salt or osmotic treatment (Figure 8A). However, the GUS/LUC activity ratios of constructs with the GT-box and TC-box were enhanced significantly by salt or osmotic treatment, as compared to the control. Specifically, the GUS/LUC activity ratios of the GT-box reached a peak at 3 h of salt treatment and at 6 h of osmotic treatment (Figure 8B). The GUS/LUC activity ratios of TC-box constructs were maximal at 1 h after salt treatment, and 0.5 h after osmotic treatment (Figure 8C). Our results suggested that GT-box and TC-box respond to abiotic stress *in vivo*, but the E-box does not.

**DISCUSSION**

Studying TF binding to *cis*-acting elements can provide insight into transcriptional
regulation and reveal, in depth, the functions of TFs. The MYB TF family, a critical plant family, has been widely studied, especially R2R3-MYB (Li et al. 2006; Dubos et al. 2010; Prouse and Campbell 2012; Kim et al. 2015). R2R3-MYB affects abiotic stress tolerance and secondary wall development via binding to the MYB-core type I sequences “CNGTT(G/A))”, MYB-core type II sequences “TNGTT(G/A))”, or AC-rich sequences “ACC(A/T)A(A/C)(T/C) and ACC(A/T)(A/C/T)(A/C/T))”, as described previously (Ithal and Reddy 2004; Zhong et al. 2013). AtMYB46 promotes secondary wall biosynthesis via binding to the cis-regulatory element (AGTAGGTG) and the SMRE sites “ACC(A/T)A(A/C)(T/C)” in Arabidopsis (Kim et al. 2012; Zhong and Ye 2012). AtMYB61 can regulate plant transpiration, xylem cell differentiation, lateral root growth, and stomatal aperture through interaction with the specific DNA sequences “ACCTAC” (Liang et al. 2005; Romano et al. 2012; Prouse and Campbell 2013). PtrMYB2/3/20/21 can bind to SMRE sites “ACC(A/T)A(A/C)(T/C)” and activate secondary wall biosynthesis (Zhong et al. 2013).

Our previous study indicated that BplMYB46, which belongs to the R2R3-MYB subfamily, can regulate the tolerance to abiotic stress and secondary wall biosynthesis, via interacting with the MYBCORE and AC-box (Guo et al. 2017). In the present study, we identified novel motifs bound by BplMYB46; these will provide further insight into the molecular mechanism underlying transcriptional control by BplMYB46.

The TF-centered Y1H (protein-to-DNA) method based on the Y1H system (Ji et al. 2014a) allowed us to screen for new motifs bound by BplMYB46. This system does not require a specific antibody against the protein of interest, nor does it use genomic methods, such as microarrays or high-throughput sequencing or removal of non-binding sequences by affinity chromatography. In the current study, we constructed a random insertion library with short DNA sequences and used this library to identify novel motifs bound by BplMYB46. In addition to the MYBCORE and AC-box, BplMYB46 also bound to a known motif, the E-box (Figure 3) and two novel motifs (the GT-box and TC-box) (Figure 1 and 2). Moreover, to confirm the regulation of gene expression by BplMYB46 binding to E-box, GT-box and TC-box elements, truncated promoters of stress-tolerance genes and secondary cell wall-specific genes were used,
with or without these motifs, in GUS assays. The results showed that BplMYB46 could activate *GUS* expression by binding to the E-box, GT-box, or TC-box motifs in tobacco (Figure 4, 5, and 6). ChIP assays further confirmed that BplMYB46 binding to E-box, GT-box, or TC-box motifs occurred in birch (Figure 7).

Our previous study indicated that BplMYB46 enhances abiotic stress tolerance in birch, via increasing scavenging of reactive oxygen species (ROS) and regulating expression of *SOD* and *POD*, which encode key ROS-scavenging enzymes. The activities of SOD and POD differed significantly in *BplMYB46* transgenic plants when exposed to NaCl or mannitol. The BplMYB46-overexpressing plants displayed significantly higher SOD and POD activities than control plants, and the *BplMYB46*-silenced plants exhibited significantly lower POD and SOD activities than control plants (Guo et al. 2017).

In the present study, these two stress-related genes, *POD* and *SOD*, were further employed for investigation of their binding and regulation by BplMYB46. We identified TC-box and GT-box motifs located in the promoters of *SOD* and *POD*, respectively (Figure 5A). The Y1H assay and GUS activities indicated that BplMYB46 could bind to the TC-box or GT-box (Figures 5B, 6B). Together with our previous study (Guo et al. 2017), our results suggest that BplMYB46 can enhance ROS scavenging by regulating the expression of *SOD* and *POD*, via an interaction with the TC-box and GT-box motifs in their promoters.

The relative GUS activities of constructs containing TC-box and GT-box elements were significantly higher under salt and osmotic treatment than under control conditions (Figure 8). This pattern suggests that, in plants, the TC-box and GT-box motifs are associated with stress tolerance, thereby further clarifying the potential transcriptional mechanism of BplMYB46 underlying resistance to adverse environmental conditions.

The E-box recognized by R2R3-MYB, bZIP, and BHLH factors has been deemed a *cis*-element associated with the response to light and the phenylpropanoid biosynthesis pathway (Hartmann et al. 2005). Phenylalanine ammonia lyase (PAL) is a key enzyme in lignin biosynthesis (Zhang et al. 2013). Interestingly, our current results
showed that the E-box is present in the \textit{PAL} promoter (Figure 5A) and did not respond to stress, \textit{in vivo} (Figure 8A). Our previous study (Guo et al. 2017) showed that \textit{BplMYB46} promoted secondary wall deposition and the expression of xylem development-related genes, including the \textit{PAL} genes. Therefore, here, we speculate that the E-box might be the \textit{cis}-acting element involved in secondary cell wall biosynthesis pathways. The present study demonstrates the complex transcriptional program of \textit{BplMYB46} in regulating tolerance to stress and secondary wall development.

In summary, we identified TC-box, GT-box, and E-box motifs as sequence elements that are bound by \textit{BplMYB46}. Furthermore, in addition to the MYBCORE and AC-box, these motifs may play important roles in the expression of target genes regulated by \textit{BplMYB46}. Our identification of the \textit{cis}-acting elements, recognized by \textit{BplMYB46}, provides valuable information that contributes to improving our understanding of the regulatory mechanisms mediated by \textit{BplMYB46}, in response to abiotic stress and in secondary cell wall deposition.

\section*{MATERIALS AND METHODS}

\textbf{Screening a random short-insertion DNA library for sequences bound by \textit{BplMYB46}}

A random short-insertion DNA library was constructed, in the pHIS2 vector, and used as the prey library, \textit{i.e.}, the reporter vector. The \textit{BplMYB46} open reading frame (ORF), without the termination codon, was constructed into the vector pGADT7-Rec2 as the effector. The primers are listed in Table S2. The effector construct and reporter library were co-transformed into Y187 cells (Ji et al. 2014a). The clones were first selected on SD/-His/-Leu/-Trp (TDO) medium/3-amino-1, 2, 4-triazole (3-AT, 30 mM), and then positive clones were subjected to the highly stringent selection of TDO medium with 50 mM 3-AT. Inserts in the pHIS2 plasmids were sequenced from colonies that grew on the selective medium.

The insertion DNA sequences, together with the left and right insertion flanking sequences (“CCC” and “GGG”), were analyzed by PLACE
(http://www.dna.affrc.go.jp/PLACE/, which is not available now) and PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) to identify known and unknown motifs.

**Core sequence determination and base mutation analysis of the motifs**

Based on the motifs identified above, DNA bases were deleted from the left or right flanking regions to confirm the core sequences of the two novel motifs. Oligonucleotides containing triple tandem repeats of the DNA sequences were inserted into the pHIS2 vector to generate reporter vectors. Briefly, pHIS2 was cleaved with EcoRI and SacI to generate cohesive ends, and the forward and reverse oligonucleotides containing three tandem repeats of the motif were annealed to generate identical cohesive ends, followed by ligation with pHIS2. The *BplMYB46* ORF was ligated into the pGADT7-Rec2 vector as an effector. To further investigate the base specificity of the three motifs bound by *BplMYB46*, tandem repeats of DNA sequences of the two novel motifs and the E-box motif, with mutations in the core sequences, were inserted into the pHIS2 vector. The reporters and the effector were co-transformed into Y187 cells.

To further investigate the base specificity of the three motifs bound by *BplMYB46*, tandem repeats of DNA sequences of the two novel motifs and the E-box motif with mutations in the core sequences were inserted into the pHIS2 vector. The primers are listed in Tables S3 and S4.

**Binding analysis of BplMYB46 to target gene promoters**

Interactions of *BplMYB46* with the specific motifs on the promoters of the target genes, including *SOD*, *POD*, and *PAL*, were investigated by Y1H. The pHIS2 vector harboring the truncated promoter regions of the genes with additional copies of the TC-box, GT-box, or E-box motifs (termed pHIS2-TC-boxp+, pHIS2-GT-boxp+, and pHIS2-E-boxp+, respectively) and the truncated promoter regions of the genes lacking the TC-box, GT-box, or E-box motifs (termed pHIS2-TC-boxp-, pHIS2-GT-boxp-, and pHIS2-E-boxp-, respectively), driving the *HIS3* gene, were established as reporter constructs. The
reporters and the effector of pGADT7-Rec2-BplMYB46 were co-transformed into Y187 cells. The primers are listed in Table S5.

**Validation by transient expression assay**

The results of the Y1H assay were validated, *in planta*, by examining the interaction between BplMYB46 and *cis*-elements using a beta glucuronidase (*GUS*) reporter gene assay. The three tandem copies of the motifs, mutated motifs (A, T to C; G, C to A), and the truncated promoters including or lacking the motifs were fused with the CaMV 35S minimal promoter (46 bp to +1, replaced 35S promoter) to drive the *GUS* gene. The effector vector pROKII-35S::BplMYB46 was co-transformed with each reporter into tobacco (*Nicotiana tabacum*) seedlings by *Agrobacterium tumefaciens*-mediated transient expression (Ji et al. 2014b). Tobacco seedlings were further subjected to abiotic stress by treatment with 200 mM NaCl or 200 mM mannitol, with water as the control. Tobacco leaves with the truncated promoters and pROKII-35S::BplMYB46 were collected at 0.5, 1, 3, 6, 9, 12, and 24 h after NaCl or mannitol treatment. The firefly luciferase (*LUC*) gene driven by the CaMV 35S promoter was also co-transformed as the control for normalization of the transformation efficiency. *GUS* and *LUC* activities were determined (Gampala et al. 2001). Data are represented as the mean of three biological replicates. The primers are listed in Tables S6 and S7.

**Searching for DNA motifs in the promoters of BplMYB46 target genes**

The expression profiles of genes regulated by BplMYB46 were analyzed using RNA-seq. FDR ≤ 0.001 and the absolute log2 ratio ≥1 were used as the threshold for significant differences in gene expression. To determine the distribution of the BplMYB46 binding motifs, the promoter regions (−1 to −1,500 bp) of the identified genes were retrieved from the birch database (http://birch.genomics.cn/page/species/index.jsp) for analysis.

**ChIP analysis**

The gene encoding a BplMYB46-green fluorescent protein (GFP) fusion
(35S::BplMYB46-GFP) was constructed into the pROKII vector and transformed into EHA105 competent cells by electroporation, followed by transient transfection into birch (B. platyphylla) plants (Ji et al. 2014b) for ChIP assays, as described previously (Ji et al. 2014a). Briefly, the birch seedlings (1–5 g) were collected and crosslinked with 1% formaldehyde for 10 min, at room temperature. The crosslinking was quenched by 180 mM glycine for 6 min, at room temperature, followed by three washes with deionized water. Tissue was then ground into a fine powder using liquid nitrogen. The purified cross-linked nuclei were sonicated to shear the chromatin into 0.2–0.8 kb fragments, and 1/10 volume was saved as the input control. The sonicated chromatin was immunoprecipitated with GFP antibody (Abmart) (ChIP+), and chromatin immunoprecipitated with anti-hemagglutinin (HA) antibody was used as a negative control (ChIP-). The DNA fragments containing or lacking the E-box, TC-box, and GT-box from the different promoter regions (−1 to −1,000 bp) of PAL, SOD, and POD were selected for amplification. The PCR amplification was as follows: 94°C for 2 min; 30 cycles at 94°C for 12 s, 55°C for 30 s, and 72°C for 45 s; 72°C for 5 min. The PCR products were analyzed by gel electrophoresis. Enrichment of truncated promoters in the immunoprecipitated samples was determined by real-time quantitative PCR (qPCR). The amplification protocol was as follows: 94°C for 30 s; 45 cycles at 94°C for 12 s, 58°C for 30 s and 72°C for 45 s; and 79°C for 1 s for plate reading. Tubulin was used as an internal control. Three biological replicates were included for statistical analysis. The primers are listed in Table S8.

ACKNOWLEDGMENTS

The authors thank Dr. Li Tan for the suggestion on the manuscript revision. This study was supported by two grants from the National Natural Science Foundation of China (31470671 and 31700587).

AUTHOR CONTRIBUTIONS

H.G., C.Y. and C.W. conceived and designed the experiments. H.G., L.W., Y.Z. and C.Z.
participated in the experiments. H.G. and C.W. analyzed the data. H.G. and C.W. drafted the manuscript. H.G. and C.W. provided the reagents and analysis tools.

REFERENCES


BZIP, and BHLH factors control light-responsive and tissue-specific activation of phenylpropanoid biosynthesis genes. **Plant Mol Biol** 57: 155-171

Ithal N, Reddy AR (2004) Rice flavonoid pathway genes, OsDfr and OsAns, are induced by dehydration, high salt and ABA, and contain stress responsive promoter elements that interact with the transcription activator, OsC1-MYB. **Plant Sci** 166: 1505-1513


Kim JH, Hyun WY, Nguyen HN, Jeong CY, Xiong L, Hong SW, Lee H (2015) AtMyb7, a subgroup 4 R2R3 Myb, negatively regulates ABA-induced inhibition of seed germination by blocking the expression of the bZIP transcription factor ABI5. **Plant Cell Environ** 38: 559-571


factor participates in dehydration tolerance in *Boea hygrometrica* by binding to the W-box elements of the galactinol synthase (*BhGolS1*) promoter. *Planta* 230: 1155-1166


SUPPORTING INFORMATION

Table S1. Distribution of motifs in the promoters of genes regulated by BplMYB46

Table S2. Primer sequences used for the construction of random cis-acting elements library and the effector of pGADT7-Rec2-BplMYB46

Table S3. Primer sequences used in reporter constructs of deletions of pHIS2-novel elements

Table S4. Primer sequences used in reporter constructs of mutations of the three elements

Table S5. Primer sequences used in the Y1H assay

Table S6. Primer sequences used in the reporter constructs analyzed in tobacco plants

Table S7. Primer sequences used in the reporter constructs to verify the results of the Y1H assay

Table S8. Primers sequences used in the ChIP-PCR analysis
Table 1 Analysis of the insertion sequences bound by BplMYB46

<table>
<thead>
<tr>
<th>Clones number</th>
<th>Random DNA insertion sequence (5′–3′)</th>
<th>Motif prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CCAAAATGAGCGG</td>
<td>E-box: “CAAATG”</td>
</tr>
<tr>
<td>2</td>
<td>CCTGTCGCCCGG</td>
<td>No result</td>
</tr>
<tr>
<td>3</td>
<td>CCAGTAGTTCGG</td>
<td>No result</td>
</tr>
</tbody>
</table>
Figure legends

Figure 1. Serial deletion of the inserted DNA motifs to determine the core sequences of two novel motifs recognized by BplMYB46

(A) Schematic representations of the reporter and effector vectors. The serially deleted insertion sequences of two novel motifs were cloned, separately, into the pHIS2 reporter vector. The BplMYB46 ORF was cloned into pGADT7-Rec2 as an effector. (B) pGADT7-BplMYB46/pHIS2-novel 1 (unknown 1): the serially deleted insertion sequences (L1–L3, R1–R5) bound by BplMYB46. L1–L3: the sequences of serially deleted insertions (left panel), R1–R5: the sequences of serially deleted insertions (right panel). (C) pGADT7-BplMYB46/pHIS2-novel 2 (unknown 2): the serially deleted sequences (L1–L3, R1–R3) bound by BplMYB46. L1–L3: the sequences of serially deleted insertions (left panel), R1–R3: the sequences of serially deleted insertions (right panel). Positive control: pGADT7-p53 binding to p53HIS2 (pGADT7-p53/p53HIS2); negative control: pGADT7-BplMYB46 binding to p53HIS2 (pGADT7-BplMYB46/p53HIS2).

Figure 2. DNA base mutation in the core sequences to determine the specificity of the two novel motifs recognized by BplMYB46

(A, B) Analysis of the specific bindings of the two novel motifs to BplMYB46 using Y1H assay. The right panel shows the sequences of the two novel motifs and their respective mutants. Positive control: pGADT7-p53/p53HIS2; negative control: pGADT7-BplMYB46/p53HIS2.

Figure 3. Specificity determination of E-box member bound by BplMYB46

Analysis of specific bindings of the BplMYB46 to E-box using Y1H assay. Right panel shows the sequences of E-box and respective substitutions. Positive control: pGADT7-p53/p53HIS2; negative control: pGADT7-BplMYB46/p53HIS2.
Figure 4. Analysis of BplMYB46 binding to different motifs in tobacco plants

(A) Schematic illustration of the motifs: E-box, GT-box, and TC-box or their complete mutation sequences used in the co-expression of reporter and effector plasmids in the tobacco leaves. (B) Transient co-transformation of reporter and effector constructs into tobacco leaves to study the bindings of E-box, GT-box, and TC-box and their mutants to BplMYB46. GUS activity and staining indicate the binding affinity between different motifs and BplMYB46. Data represent the mean (n = 3) + SD.

Figure 5. Y1H assay of the binding of E-box, GT-box, and TC-box on the promoters of target genes bound by BplMYB46

(A) Positions of the truncated promoters including E-box, GT-box, and TC-box, and lacking E-box, GT-box, and TC-box in the promoters of target genes. (B) Reporter constructs and the effector plasmid, pGAD-BplMYB46 (pGADT7-Rec2 harboring BplMYB46), were co-transformed into yeast Y187 cells, and the positive transformants determined by spotting the serial dilutions (1:1, 1:10, 1:100, 1:1000) of yeast onto DDO and TDO plates supplemented with 3-AT. pGADT7-p53/p53HIS2: positive control; pGADT7-BplMYB46/ p53HIS2: negative control.

Figure 6. Analysis of BplMYB46 binding to the truncated promoters of PAL, POD, and SOD genes in tobacco plants.

(A) Schematic representation of truncated promoter, including or lacking the E-box, GT-box, or TC-box motifs used in the co-expression of effector and reporter plasmids in the tobacco leaves. (B) Transient co-transformation of effector and reporter constructs into tobacco leaves to study the binding of E-box, GT-box, and TC-box to BplMYB46. GUS activity indicated the binding affinity of BplMYB46 to the truncated promoters. Data represent the mean (n = 3) ± SD.

Figure 7. BplMYB46 binding to different motifs using ChIP assay

(A) Positions of the truncated promoters containing and lacking E-box, GT-box, and
TC-box, respectively, in the promoters of the target genes. (B) ChIP products obtained from the promoters of \textit{PAL}, \textit{POD}, and \textit{SOD} genes analyzed by gel electrophoresis after PCR amplification. (C) Real-time quantitative PCR analysis showing enrichment of the promoter sequences of \textit{PAL}, \textit{POD}, and \textit{SOD} genes after ChIP. Input, Input DNA (positive control); CHIP+: chromatin immunoprecipitation with the anti-GFP antibody; CHIP−: chromatin immunoprecipitation with HA antibody (negative control).

**Figure 8. Relative GUS activity of E-box, GT-box and TC-box response to stress**

(A) Relative GUS activity of E-box under salt and osmotic stress. (B) Relative GUS activity of GT-box under salt and osmotic stresses. (C) Relative GUS activity of TC-box under salt and osmotic stresses.
Figure 2
Figure 3

Figure 4
Figure. 5

Figure. 6
Figure 7
Figure. 8