Acta Physiol Plant (2017) 39:189
DOI 10.1007/s11738-017-2488-4

ORIGINAL ARTICLE

Involvement of WRKY, MYB and DOF DNA-binding proteins in interaction with a rice germin-like protein gene promoter

Farah Deeba[1] · Tasawar Sultana[1] · Tariq Mahmood[2] · Charlotte O'Shea[3] · Karen Skriver[3] · S. M. Saqlan Naqvi[1]

Received: 30 September 2016 / Revised: 19 February 2017 / Accepted: 24 July 2017 / Published online: 31 July 2017

Abstract Expression of germin-like proteins (GLPs) is reportedly modulated during exposure to pathogens and abiotic stresses. Nevertheless, little is known about the transcription factors and their modulatory role in the mechanism of the regulation of GLP genes. The promoter of *Oryza sativa* Root Expressed GLP2 (*Os*RGLP2) gene reportedly showed strong expression in transgenic tobacco during salinity, dehydration and wounding stresses. In the present study, an effort has been made to characterize the *cis*-regulatory elements of *Os*RGLP2 promoter and their binding proteins. The putative stress-responsive regulatory elements in the promoter and corresponding binding proteins (*Os*WRKY71, *Os*DOF18 and *Os*MYB1) were identified by *in silico* analysis. The DNA-binding domains of selected proteins were cloned, overexpressed and purified. Electrophoretic mobility shift assays (EMSAs) demonstrated that these recombinant domains were able to bind with DIG-labeled *Os*RGLP2 promoter fragments containing W-box, AAAG and WAACCA motifs. Binding was confirmed by competitor EMSA and EMSA with mutant oligonucleotides. These regulatory elements were also active in binding with nuclear factors from rice nuclear protein extract in vitro as confirmed by competitive EMSA. It can be concluded that *Os*WRKY71, *Os*MYB1 and *Os*DOF18 proteins are involved in transactivation of *Os*RGLP2 gene expression under different abiotic stress conditions.

Keywords DNA-binding proteins · EMSA · *Os*RGLP2 promoter · *Os*DOF18 · *Os*MYB1 · *Os*WRKY71

Introduction

Plant growth and productivity are adversely affected by several types of stresses such as drought, salinity, threshold temperatures and pathogen attack. To cope with these challenges, plants have evolved molecular systems consequently developing adaptive responses through physiological, biochemical and molecular changes (Bohnert et al. 1995). A number of stress-responsive genes have been reported to be involved in the alleviation of stress-induced cellular damage; one of the class is germin-like proteins (GLPs) which is believed to provide resistance to plants against biotic and abiotic stresses.

GLPs are widely distributed in plants and seem to perform a range of metabolic roles from structural to catalytic. The catalytic roles, mostly include oxalate oxidase (OXO) for germins and superoxide dismutase (SOD) activity for GLPs (Yasmin et al. 2015); nevertheless, serine protease inhibitory activity (Segarra et al. 2003), ADP glucose pyrophosphatase (Rodriguez-Lopez et al. 2001), polyphenol oxidase (Cheng et al. 2014) and cysteine peptidase activities (de Freitas et al. 2016) have also been associated with GLPs.

Communicated by Z.-L. Zhang.

Electronic supplementary material The online version of this article (doi:10.1007/s11738-017-2488-4) contains supplementary material, which is available to authorized users.

✉ S. M. Saqlan Naqvi
saqlan@uaar.edu.pk

[1] Department of Biochemistry, PMAS Arid Agriculture University Rawalpindi, Rawalpindi, Pakistan

[2] Department of Plant Sciences, Quaid e Azam University, Islamabad, Pakistan

[3] Department of Biology, University of Copenhagen, 5 Ole Maaloesvej, 2200 Copenhagen, Denmark

Germin and GLPs are not only expressed at specific developmental stages, but many members are also responsive to biotic and abiotic stresses (Jiang et al. 2007; Dunwell et al. 2008). GLP overexpression in transgenic plants has been associated with improved tolerance to pathogen infection and salt stress (Wang et al. 2013); thus insight into GLP promoters and their corresponding DNA-binding proteins becomes highly relevant. There are several reports regarding *in silico* analysis of GLP promoters. Sassaki et al. (2014) analyzed *Eg*GLP promoter and found *cis*-elements implicated in light, auxin, abscisic acid and salicylic acid responsiveness. Himmelbach et al. (2010) characterized the seven W-boxes found to be essential for induction by salicylic acid and providing defense against pathogen infection.

Mahmood et al. (2007) cloned *Os*RGLP2 promoter, performed heterologous expression in tobacco and observed strong GUS expression in transgenics during wounding/mechanical stress, salt and dehydration. Hussain (2015) performed 5′ deletion analysis of the *Os*RGLP2 promoter and observed that full length as well as two deleted fragments of 776 and 565 bp were responsive to wound, salinity and temperature stresses, indicating that the *Os*RGLP2 promoter may serve as a location where multiple signaling pathways might integrate to produce a response to stress stimuli. However, experimental evidence regarding the identification and functional characterization of transcription factors which might interact with *cis*-acting regulatory elements of the GLPs promoters were required. Therefore, in the present study, putative transcription factors that interact with *cis*-acting regulatory elements of *Os*RGLP2 region were identified by *in silico* analysis followed by their cloning in *E. coli*, overexpression and purification. Interaction of purified proteins with fragments of *Os*RGLP2 promoter was investigated by electrophoretic mobility shift assays (EMSAs). Results revealed that *Os*RGLP2 promoter can interact with *Os*WRKY71, *Os*DOF18 and *Os*MYB1 proteins which play a role in regulating *Os*RGLP2 gene transcription.

Materials and methods

Computational analysis of *Os*RGLP2 promoter

The *Os*RGLP2 promoter region (accession no. DQ414400.1) was screened for the presence of putative stress-responsive elements as listed in PLACE (http://www.dna.affrc.go.jp/PLACE/signalscan.html). The elements detected by PLACE were further validated by other online plant *cis*-element databases, PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/), ConSite (http://consite.genereg.net/), and JASPAR (http://jaspar.binf.ku.dk/). Target genes were investigated using the gene expression database called Genevestigator (Plant Biology) (https://genevestigator.com/gv/plant.jsp) to identify WRKY71 (Os02g0181300), MYB1 (Os01g0850400) and DOF18 (Os08g0490100) that appeared to show up-regulation during various stresses. Data were acquired for abiotic stresses including cold, heat, drought and salt stresses.

Sequence retrieval and sequence alignment for the DNA-binding domain identification

Nucleotide and protein sequences of selected stress-responsive transcription factors were obtained from NCBI (http://www.ncbi.nlm.nih.gov/). Protein sequences of *At*MYB1 and *Zm*DOF1 were subjected to BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and search was performed against a non-redundant database, but was limited to include only hits within the *O. sativa* genome. For identification of DNA-binding domains, protein sequences of various plant proteins were obtained from BLAST result and multiple alignment was performed using TCOFFEE (http://tcoffee.crg.cat/). ESPript (http://espript.ibcp.fr/ESPript/ESPript/) was then used for producing graphical representations of the multiple alignments. Secondary structures of DNA-binding domains were predicted using PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/).

Preparation of crude nuclear protein extract

To analyze the possible physical interactions of predicted sites with specific transcription factors, nuclear proteins were extracted from *Oryza sativa.* For this purpose, *O. sativa* cv KS282 seeds obtained from Rice Program, National Agricultural Research Centre, Islamabad, were germinated on half-strength MS medium at 25 °C for 10 days. Five grams of 10-day-old seedling mass was ground in an ice-cold mortar and pestle in liquid nitrogen and nuclear proteins were extracted according to the protocol of Escobar et al. (2001). The nuclear protein extract was dialyzed in the EMSA binding buffer.

Production and purification of recombinant proteins

To validate the interaction of target transcription factors with the respective putative binding sites, target proteins were produced as purified recombinant proteins. For this purpose, total RNA was isolated using RNeasy Plant Mini Kit (Qiagen) followed by cDNA synthesis using Oligo (dT) 18 Primer and RevertAid Premium Reverse Transcriptase (Thermoscientific). cDNA was then used as template for PCR amplification of DNA-binding domains of *Os*WRKY71 and *Os*MYB1, while cDNA clone

(J065152E11) acquired from NIAS was used as template for *Os*DOF18. The DNA sequences coding for DNA-binding domains, including 15–20 amino acids flanking N and C terminus, were amplified using specific primers containing *Bam*HI and *Xho*I restriction sites (Supplementary Table 1) and cloned into predigested pGEX4T-1 vector. Cloning of correct fragments in recombinant plasmids was confirmed by restriction digestion and commercial sequencing. The recombinant plasmids were transformed into BL21 (DE3) and expression was induced with 1 mM isopropyl β-D-thiogalactopyranoside (IPTG) for 3 h when OD_{600} reached about 0.6. The cells were harvested by centrifuging (3500*g*, 20 min, 4 °C) and resuspended in phosphate-buffer saline (PBS) containing 1% triton X-100. The resuspended pellet was sonicated (6 cycles, 30 s) followed by incubation on ice for 30 min. The supernatant was then collected after centrifugation for 20 min at 15,000*g* at 4 °C and adsorbed onto glutathione-Sepharose beads for 30 min at 4 °C, after which the beads were washed twice in PBS and fusion proteins were eluted using elution buffer. The apparent molecular mass and purity of eluted proteins were analyzed by 15% SDS-PAGE according to standard protocols. Semi-purified proteins were subjected to cation exchange chromatography and the eluents were dialyzed in EMSA binding buffer [50 mM HEPES; pH 7.5, 40 mM KCl, 12.5 mM $MgCl_2$, 2 mM DTT, 1 mM PMSF 0.2 mM EDTA, 20% (v/v) glycerol].

Electrophoretic mobility shift assays (EMSAs)

Oligonucleotides (30 bp) containing WRKY, MYB or DOF binding sites located in the *Os*RGLP2 promoter were annealed by heating at 95 °C for 10 min and then slowly cooled to room temperature in TEN buffer [Tris (10 mM, pH 8), 1 mM EDTA, 100 mM NaCl]. A DIG Gel Shift kit (Roche) was used to detect DNA–protein interaction. Digoxigenin (DIG) was labeled at the 3′ end of the double-stranded oligonucleotides. The labeled probes were then purified by ethanol precipitation. The pellet was air dried and resuspended in ddH_2O to a final concentration of 2.5 pmol/μL. The labeling efficiency was tested by spotting serial dilutions on nylon membrane along with control labeled oligonucleotides provided with the kit.

DIG-labeled (50 fmol), unlabeled (10 pmol) or mutated labeled oligonucleotides (50 fmol) were mixed with GST fusion proteins or a crude nuclear protein extract (NPE) in binding buffer and kept at room temperature for 30 min. The reaction products were fractionated through a pre-run 6% non-denaturing polyacrylamide gel for 90 min in 0.5× TBE buffer at 4 °C. The proteins were then electrophoretically transferred onto a positively charged nylon membrane (Roche) by applying a constant current of 300 mA for 30 min. The DNA was cross-linked to the membrane using a UV stratalinker (Stratagene). The nylon membrane was blocked in 1× blocking reagent for 1 h at room temperature and then incubated with a 1:20,000 dilution of alkaline phosphatase coupled anti-DIG antibody for 30 min. The membrane was washed twice with washing buffer [0.1 M maleic acid (pH 7.5), 0.15 M NaCl, 0.3% (v/v) Tween 20] for 20 min each. After 5 min equilibration in detection buffer (100 mM Tris/HCl, pH 9.5, 100 mM NaCl), the membrane was carefully placed on a plastic sheet and 1:100 dilution of the CSPD substrate was then added on the membrane dropwise near the edges and then the whole membrane was covered by tilting it. The membrane was then incubated at room temperature for 5 min and subsequently placed in a plastic folder and incubated at 37 °C for another 15 min. Finally, the membrane was exposed to X-ray film for 3 h to capture the chemiluminescent signal.

Results

In silico analysis of the *Os*RGLP2 promoter

In the *Os*RGLP2 gene promoter region, many key *cis*-elements that were responsive to abiotic stresses were identified by PLACE, for instance MYB binding site, MYC binding site, DRE elements (dehydration-responsive element), DOF binding site, ABRE, GT-1 element, BIHD1 binding site, ANAERO2CONSENSUS, CBF (C-repeat binding site) and W-Box (Supplementary Table 2). Putative *cis*-elements were further validated by analyzing the promoter sequence in JASPAR, PlantCare and Consite tools. All the tools shared the presence of MYB, W-box and DOF elements. JASPAR also predicted the presence of bZIP and NAC025 binding sites. Fourteen putative W-boxes, 14 DOF and 4 MYB1 binding sites were found common in the *Os*RGLP2 promoter region on plus and negative DNA strands (Fig. 1). Therefore, these three sites were chosen for further analysis and it was assumed that WRKY, DOF and MYB proteins may be involved in transcriptional regulation of *Os*RGLP2 gene under abiotic stresses.

*Os*WRKY71 was predicted by PLACE as an interaction partner of W-box located in the *Os*RGLP2 promoter. To better understand the secondary structure of the WRKY71 DNA-binding domain, multiple alignments were carried out with WRKY proteins from different plants. Results revealed that it contains a single, highly

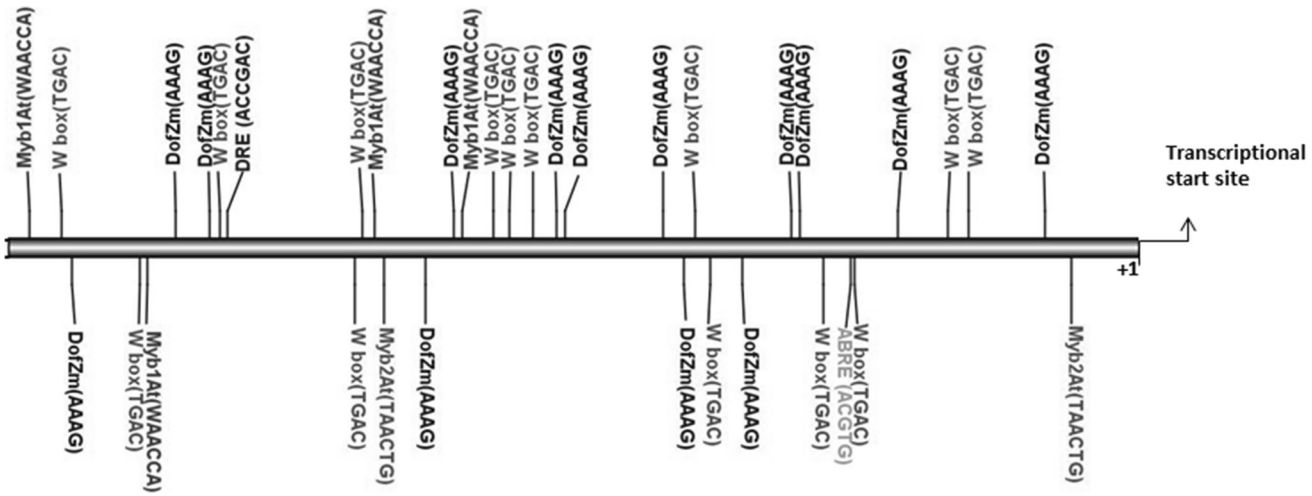

Fig. 1 Position of different stress-responsive *cis*-elements on the 1109 bp *Os*RGLP2 promoter region

conserved WRKY domain and zinc finger-like motif at its C terminus ($C–X_5–C–X_{23}–H–X_1–H$) that places it in group IIA of the WRKY family (Supplementary Fig. 1). Based on the putative *cis*-element binding sites, sequences of *At*MYB1 and *Zm*DOF1 (predicted by PLACE, Consite and JASPAR) were retrieved from the NCBI database. Homologous MYB and DOF sequences in rice genome were identified by BLAST against *At*MYB1 and *Zm*DOF1 query sequences. The result showed *At*MYB1 to have 42% sequence similarity (all in the DNA-binding domain region) with *Os01g0850400* which is similar to partial *Os*MYB1 protein sequence (accession no. CAE00856.1). Multiple sequence alignment with MYB proteins from different plants showed that it contains two conserved MYB DNA-binding domains at the N-terminus and belongs to R2R3-type subfamily of MYB DNA proteins (Supplementary Fig. 2).

The protein sequence of *Zm*DOF1 when subjected to BLAST search revealed 54% amino acid sequence similarity with putative rice protein. Protein functional analysis by InterProScan showed identity with an InterPro accession number IPR003851 Znf_DOF, thereby confirming its identity to DOF-like proteins. This protein sequence was found in the Database of Rice Transcription Factors (DRTF) under locus ID *Os08g38220* named as *Os*DOF18/MNB1A. It encodes a unique DNA-binding domain of 55 amino acids encompassing a C2–C2 zinc finger (Supplementary Fig. 3). To investigate the association of *Os*WRKY71, *Os*MYB1 and *Os*DOF18 transcripts with stress responses, an examination of global expression was performed using the Genevestigator (Supplementary Table 3). The transcript level of all genes was significantly up-regulated under different abiotic stresses to which the *Os*RGLP2 promoter is responsive including salt and drought.

Binding of recombinant GST-fused *Os*WRKY71, *Os*MYB1 and *Os*DOF18 DNA-binding domains with the *Os*RGLP2 promoter region

The possible physical interaction of *in silico* identified transcription factors with *Os*RGLP2 promoter was explored by producing GST-fused recombinant DNA-binding domains in *E. coli* BL21 and purified as GST/WRKY71 (167–271), GST/DOF18 (1–148) and GST/MYB1 (88–222) (Supplementary Fig. 4). To test the hypothesis on the regulation of the *Os*RGLP2 promoter by *Os*WRKY71, *Os*MYB1 and *Os*DOF18, electrophoretic mobility shift assays (EMSAs) were performed to assess the potential binding of these proteins to the *Os*RGLP2 promoter region. WRKY proteins are generally considered to bind to the consensus W-box sequence C/TTGACC/T (Eulgem and Somssich 2007). The *Os*RGLP2 promoter contains four true W-boxes; two true W-boxes TTGACT at positions −800 and +491, and two true W-boxes CTGACC at positions −129 and −340, although a number of TGAC core sequences are present. As a first step toward characterization of *Os*WRKY71 binding sites in the *Os*RGLP2 promoter, three double-stranded 30 bp labeled promoter fragments containing either true W-box or core W-box sequences were assayed for their ability to bind to a purified GST/WRKY71 alone or in the presence of wild-type and mutant DNA fragments. There was only one true W-box and one TGAC cores in the WRKY71-Probe1 (WRKY71-P1) (Fig. 2a). No binding signal was detected in negative controls, while *Os*WRKY71 bound to the wild-type WRKY71-P1. Binding specificity was demonstrated by the competition experiment with 200-fold excess of unlabeled fragment which outcompeted the binding to the probe, suggesting that the *Os*RGLP2 promoter fragment specifically interacted with WRKY71. No shift was

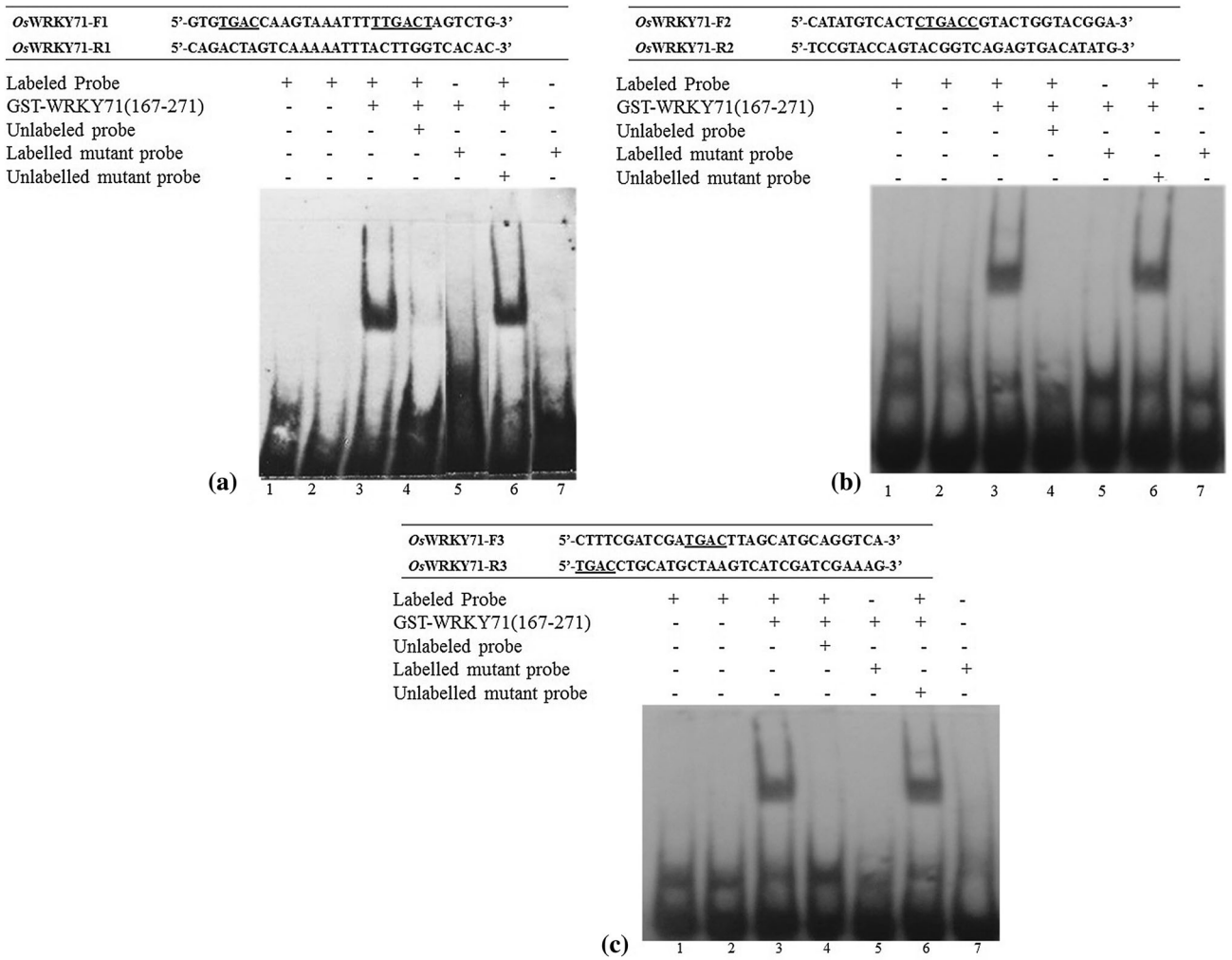

Fig. 2 In vitro DNA-binding ability of *Os*WRKY71 with DIG-labeled *Os*RGLP2 fragments containing core and true W-boxes. EMSA of recombinant *Os*WRKY71 with fragment 1 containing one core and one true W-box (**a**), with fragment 2 containing one true W-box (**b**), and with fragment 3 containing one core W-box on both strands (**c**). In gels **a, b, c** *lanes* are in the same order. *Lane 1* *Os*WRKY71 oligos negative control, *lane 2* GST-negative control, *lane 3* 1 µg GST–*Os*WRKY71 with *Os*WRKY71 oligos, *lane 4* with 200 molar excess of unlabeled *Os*WRKY71-2 oligos, *lane 5* with mutant labeled oligos, *lane 6* 200-fold molar excess of unlabeled mutant oligos, *lane 7* mutant labeled oligos negative control

observed when the TGAC sequences (core W-box) were mutated to AAAA. However, a competitor with both elements mutated (mTGACs) could not abolish the binding of the protein to the probe. Similar results were observed for WRKY71-P2 and WRKY71-P3 probes that had only one true W-box or one W-box core sequence, respectively (Fig. 2b, c). No shift was observed in lanes containing the labeled probe without protein and labeled probe with GST in case of WRKY71-P2. GST–*Os*WRKY71 bound directly to WRKY71-P2 when incubated with the labeled probe. To confirm binding, W-box sequence (TGAC) was changed to AAAA and EMSA was performed using labeled mutated probe. This mutated W-box containing EMSA probe failed to bind with GST–*Os*WRKY71. There was no shift in the negative control of the labeled mutant probe. For competition, 200-fold molar excess of unlabeled cold WRKY71-P2 was added to the EMSA reaction before addition of the labeled probe. Intriguingly, cold EMSA probe competes well with the radiolabeled probe. When excess of mutated cold EMSA probe was added to the reaction, it failed to compete with the labeled W-box containing a fragment of the *Os*RGLP2 promoter (Fig. 2b). With WRKY71-P3, there was no shift in both negative controls, i.e., labeled probe without protein (GST and WRKY71) and labeled probe with GST. GST–*Os*WRKY71 bound to the core sequence and a shift was observed, but this binding was weak as compared to true W-box sequence. The specificity of shift was further established through competition and mutation assays. The shifted band was totally competed out by 200-fold molar excess of unlabeled cold probe. It was observed that there was no shift when the core sequence (TGAC) was mutated to AAAA. Excess of the unlabeled

mutant probe also failed to abolish the WRKY–DNA complex with labeled *Os*WRKY71-P3. Taken together, the data suggested that *Os*WRKY71 can recognize the *Os*RGLP2 promoter in vitro, binding specifically to TGAC containing *cis*-elements.

To investigate the function of the several AAAG motifs in the *Os*RGLP2 5′-flanking region, DIG-labeled DNA Probes 1 and 2 representing two and one AAAG motifs, respectively, were used in EMSA. When DOF18-P1 was incubated with the GST/*Os*DOF18 protein, a shifted complex was observed that was competed out when the probe was incubated with 200 molar excess of the corresponding unlabeled oligonucleotides (Fig. 3a). As expected, this binding was not produced when the control GST protein was used in the assay. The same result was obtained when DOF18-P2 containing a single DOF-box was used (Fig. 3b), but the shift was faint in comparison to a probe containing two DOF boxes. The specificity of the interaction was also confirmed using variants of these probes. The binding of the GST/*Os*DOF18 to the DOF-box motif in the *Os*RGLP2 gene promoter was abolished when the core AAAG sequence was mutated to AGAC. As shown in Fig. 3b, the mutant version of the DOF probe was not bound by the GST/*Os*DOF18 protein or was not able to compete the binding of the corresponding wild-type probe at 200 molar excess of mutant DOF18 probe. These results entail greater DOF18 binding to a complex of two AAAG motifs than a single motif (Fig. 3c), implying an additive binding effect due to the presence of a greater number of DOF elements. The presence of multiple DOF18 binding sites in the *Os*RGLP2 promoter is thus suggestive of the greater DOF18 binding at multiple sites to enhance its effect.

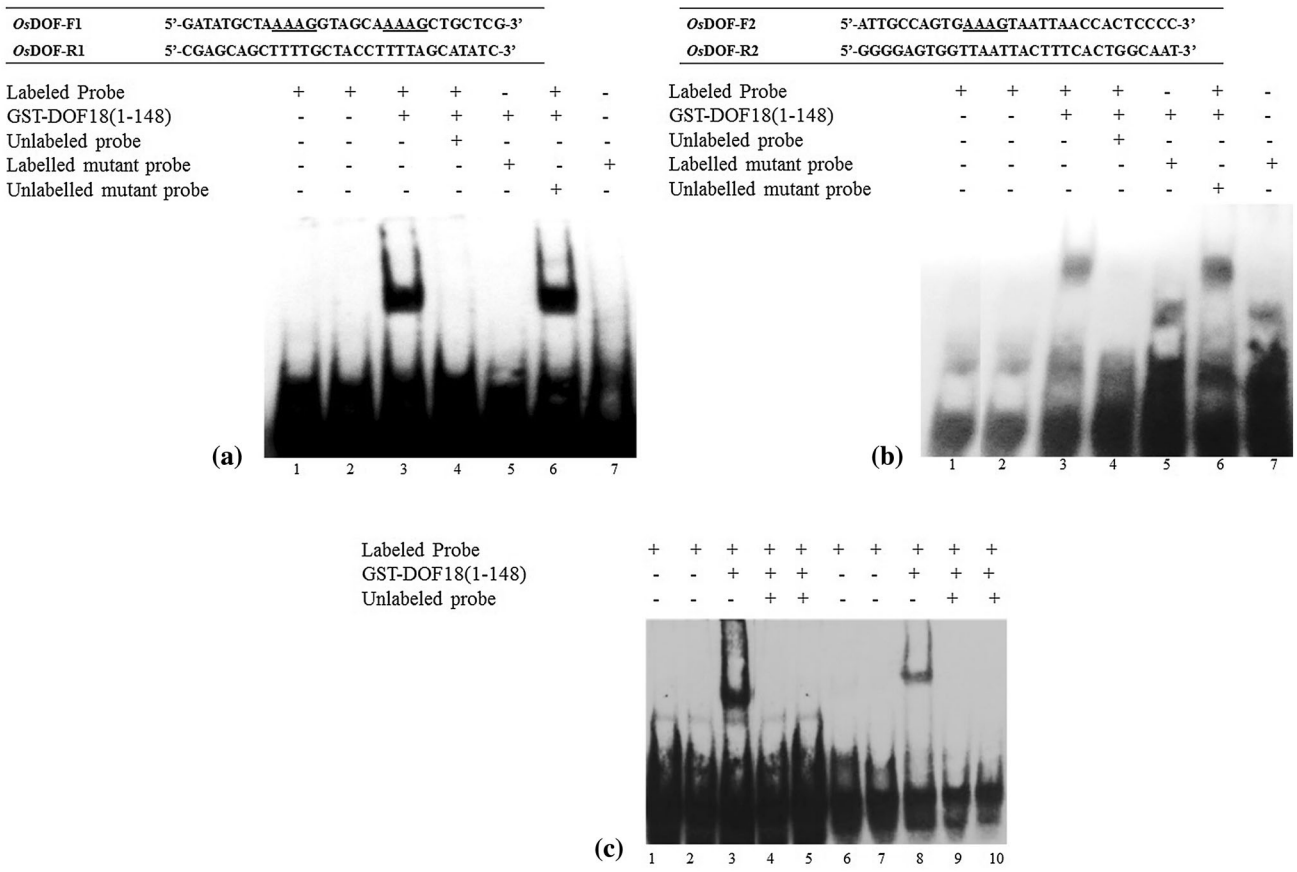

Fig. 3 In vitro DNA-binding ability of *Os*DOF18 with DIG-labeled *Os*RGLP2 fragment containing two or one AAAG motif. **a**, **b** EMSA of recombinant *Os*DOF18 with *Os*RGLP2 promoter fragment containing two and one AAAG motifs, respectively. In both **a, b,** *lanes* are in the same order. *Lane 1 Os*DOF18 oligos negative control, *lane 2* GST-negative control, *lane 3* 1 µg GST–*Os*DOF18 with *Os*DOF18 oligos, *lane 4* with 200 molar excess of unlabeled *Os*DOF18 oligos, *lane 5* with mutant labeled oligos, *lane 6* 200-fold molar excess of unlabeled mutant oligos, *lane 7* mutant labeled oligos negative control. **c** DOF18-P1 show more specificity for *Os*DOF18 as compared to DOF18-P2. *Lane 1 Os*DOF18-P1 oligos negative control, *lane 2* GST-negative control, *lane 3* 1 µg GST–*Os*DOF18 with *Os*DOF18-P1, *lane 4* with 200 molar excess of unlabeled *Os*DOF18-P1 oligos, *lane 5* with 400 molar excess of unlabeled *Os*DOF18-P1 oligos, *lane 6 Os*DOF18-P2 oligos negative control, *lane 7* GST-negative control, *lane 8* 1 µg GST–*Os*DOF18 with *Os*DOF18-P2, *lane 9* with 200 molar excess of unlabeled *Os*DOF18-P2 oligos, *lane 10* with 400 molar excess of unlabeled *Os*DOF18-P2 oligos

MYB transcription factors are reported to recognize specific binding motifs, A/TAACCA and C/TAACG/TG (Abe et al. 2003). In the *Os*RGLP2 promoter, MYB-binding motifs (5′-TAACCA-3′ and 5′-AAACCA-3′) were also found at positions −137 and −446. As shown in Fig. 4, the GST/*Os*MYB1 fusion protein was able to bind to the DIG-labeled MYB1-P1 and MYB1-P2. The unlabeled probes competed with DIG-labeled probe, suggesting that the signal from the probe represents the specific binding ability of the *Os*MYB1 with the *Os*RGLP2 promoter fragments. When the MYB-box (WAACCA) was mutated to TCATGA, no shift was observed. In addition, probes with a mutation in the MYB-binding sequence failed to compete with the wild-type form of the probe. These results demonstrated that *Os*MYB1 was able to directly and specifically bind to the MYB binding site of the *Os*RGLP2 promoter in vitro.

The interactions of rice nuclear protein extract with all probes were also investigated by EMSAs. It was observed that shifts appeared as multiple bands in case of WRKY71-P2 and P3 (Fig. 5). The specificity of binding was also demonstrated by competitive EMSA. When WRKY71-P2 and WRKY71-P3 sequences were scanned with PLACE, Consite, JASPAR and PlantCare, these promoter regions were also found to contain binding sites for proteins other than WRKY71 (data not shown). The appearance of multiple bands may be a consequence of binding of different proteins to the employed probes as a result of binding different proteins (in addition to the three tested in this study) or transcription factor with and without modulatory proteins.

Discussion

Transcriptional regulation of plant genes plays a vital role in the activation of inducible plant defense response. Many studies have been conducted on the expression of GLP genes during development and stress conditions (Li et al. 2010), but still little is known about the mechanism of their regulation. *Os*RGLP2 promoter showed strong expression of GUS gene in tobacco plants during wounding/mechanical stress, salt and dehydration (Mahmood et al. 2007). The publication of the genome sequences for rice plant and the development of recent bioinformatics tools like Genevestigator (http://genevestigator.com), TRANSFAC (http://www.biobase-international.com/product/transcription-factor-binding-sites), have hastened the studies on the regulation of expression of genes of interest. The *cis*-elements play a vital part in the transcriptional regulation of gene expression, controlling complicated abiotic stresses to increase the tolerance of plants under fluctuation environments. In our report, different bioinformatics tools were used to identify the putative stress-related *cis*-acting regulatory elements located in the *Os*RGLP2 promoter region. Significant variations were observed in the occurrence of stress-related *cis*- acting elements in the promoter including DOF, MYB, WRKY, ABRE, DREB and GT-1 elements. The *Os*RGLP2 promoter contains more than ten binding sites for WRKY and DOF proteins and three sites for MYB proteins, which indicates that these genes might be strongly associated with functions under different stresses. However, to validate the computational prediction of transcription factors binding sites within the promoter, it

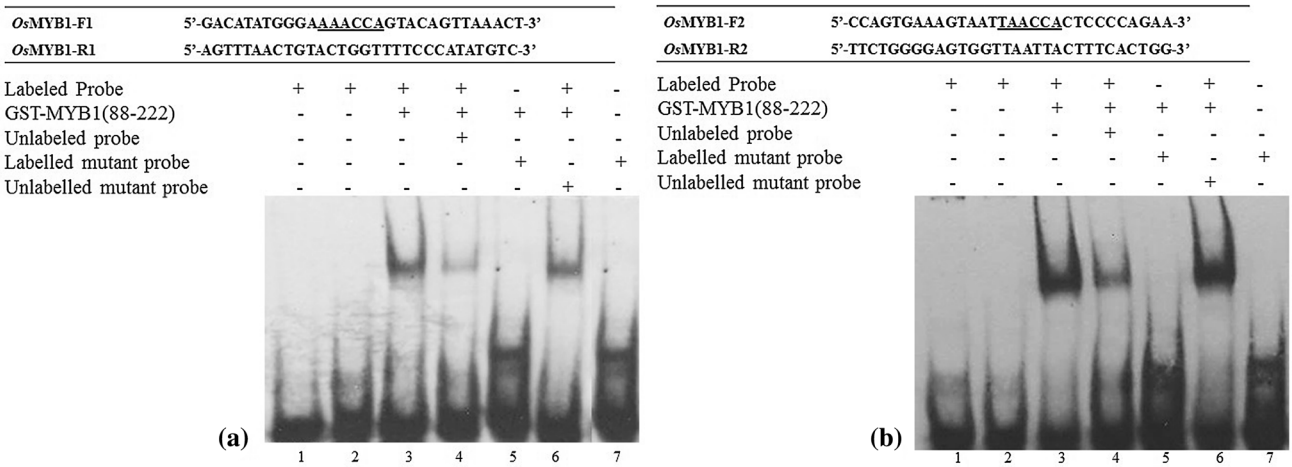

Fig. 4 In vitro DNA-binding ability of *Os*MYB1 with DIG-labeled *Os*RGLP2 fragments containing AAACCA or TAACCA MYB binding site, respectively. **a**, **b** EMSA of recombinant *Os*MYB1 with a fragment containing AAACCA and TAACCA motifs, respectively. Both in **a**, **b** *lanes* are in the same order. *Lane 1* *Os*MYB1 oligos negative control, *lane 2* GST-negative control, *lane 3* 1 µg GST–*Os*MYB1 with *Os*MYB1 oligos, *lane 4* with 200 molar excess of unlabeled *Os*MYB1 oligos, *lane 5* with mutant labeled oligos, *lane 6* 200 fold molar excess of unlabeled mutant oligos, *lane 7* mutant labeled oligos negative control

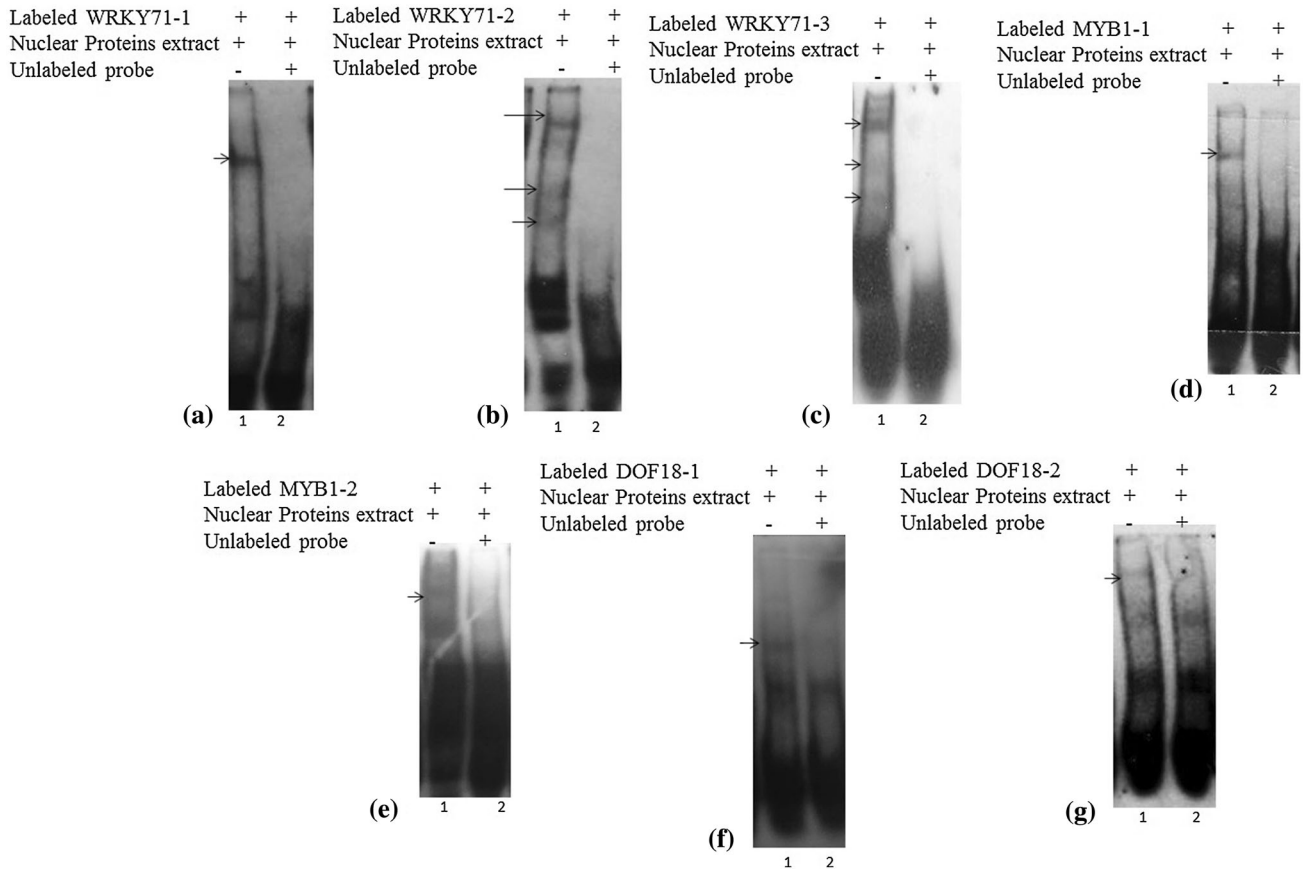

Fig. 5 In vitro DNA-binding ability of nuclear protein extract with *Os*RGLP2 promoter fragments. In each gel, *lane 1* 7 µg nuclear protein extract with labeled probes, and *lane 2* with 200 molar excess of respective unlabeled probe. The *arrows* represent the specific shift due to binding of protein to *cis*-acting elements on probe regions. The specific shift faded away when an excess of unlabeled probes were incubated with nuclear protein extract

is essential to experimentally investigate the interaction between putative *cis*-elements and corresponding binding factors. To examine the expression of *Os*WRKY71, *Os*MYB1 and *Os*DOF18 under various abiotic stress conditions in rice, we took advantage of the available microarray from Genevestigator. Analyses of microarray data demonstrated that the expression of *Os*WRKY71, *Os*MYB1 and *Os*DOF18 genes is regulated by various abiotic stress conditions.

W-boxes with typical TGAC core motif forming binding sites for WRKY proteins are reported to be involved in plant development and abiotic and biotic stresses (Eulgem and Somssich 2007; Himmelbach et al. 2010). AAAG motif is a core site for binding of DOF proteins, a type of zinc finger regulatory protein that plays diverse roles in plant gene expression such as in developmental stages (Konishi and Yanagisawa 2007) and different stresses (Kang et al. 2003). MYB transcription factors bind to the WAACCA motif and respond to various environmental stimuli (Yu et al. 2012) and plant metabolism (Goicoechea et al. 2005). The high copy number of W-boxes, WAACCA and AAAG motifs in the upstream regulatory region of *OsRGLP2* may describe them as stress-responsive proteins and these *cis*-regulatory elements may participate in the regulation of its expression to wounding, drought, salt and cold stresses.

DNA-binding domains recognize the regulatory motif sequence and binds with it through short motif, mostly an α-helix or a β-sheet. These short motifs make contact with the major groove of double-stranded DNA. Usually, the DNA–protein contact spread out across 5 bp, with a relative high affinity and sequence specificity (Ciolkowski et al. 2008). It is observed that neighboring nucleotides of *cis*-elements contribute in determining what type of transcription factors will be recruited (Ciolkowski et al. 2008). Moreover, the small differences that exist between DNA-binding domains could play a role in fine-tuning of their function (Bergholtz et al. 2001). *Os*WRKY71 was predicted as an interaction partner of W-box in the *Os*RGLP2 promoter. Zhang et al. (2004) functionally characterized *Os*WRKY71 and found it to be overexpressed in the aleurone layer/cells of rice and to act as a negative

regulator of GA signaling. *Os*WRKY71 interacts with the GA-inducible Amy32b promoter and represses the GAMYB-mediated GA signaling in aleurone cells (Xie et al. 2006). Liu et al. (2007) observed that *Os*WRKY71 expression was up-regulated in response to wounding and pathogen infection in rice. *Os*MYB1 and *Os*DOF18 sequences retrieved from *O. sativa* genome, followed by a prediction of their secondary structures, suggested that they might interact with DNA. MYB and DOF proteins are involved in various physiological processes, including abiotic stress response (Yu et al. 2012), but the role of *Os*MYB1 and *Os*DOF18 in biotic and abiotic stresses still need to be addressed.

Scientists use diverse methods to study the DNA–protein interaction: for instance, yeast one-hybrid, phage display technique and EMSA. In this study, DIG-labeled EMSA was used to scan the potential transcriptional regulatory factors for *Os*RGLP2 promoter in vitro. The W-boxes are present in the promoter regions of various plant defense providing genes, including WRKY itself. Both computational and functional studies have found clusters of W-boxes in the promoters of stress-inducible genes: for instance, occurrence of W-boxes in the promoters of PR-1, a marker gene for SAR (Maleck et al. 2000). Multiple W-boxes have a synergistic effect on transcription (Mare et al. 2004). There are four putative W-boxes on *Os*RGLP2 promoter: two (TTGACT) W-boxes on the positive strand and two (CTGACC) on the negative strand. In addition, many TGAC core sequences were also found. The present study demonstrated that *Os*WRKY71 interacts with a W-box core sequence (TGAC) as well as extended W-box (C/TTGACC/T) and controls the expression of the *Os*RGLP2 gene. Adjacent sequences of core W-box may aid in/act synergistically to enhance specificity and affinity (Mare et al. 2004). Co-expression and up-regulation of WRKY and GLPs were observed in rice upon infection with *Fusarium fujikuroi* (Matic et al. 2016).

In the *Os*RGLP2 promoter, MYB-binding motifs (5′-TAACCA-3′ and 5′-AAACCA-3′) were found. The cloned DNA-binding domain of *Os*MYB1 was able to bind both these sequences. Binding was confirmed by competitor EMSA and EMSA with mutant oligonucleotides. Nectarin-I proteins belong to the GLP family and its promoter analysis revealed the presence of consensus MYB DNA binding site (Carter and Thornburg 2003). Binding of *Nt*MYB305 with Nectarin-I promoter was evaluated by EMSA and it was found that R2R3 MYB DNA-binding domain containing *Nt*MYB305 is able to interact with the Nectarin-I promoter (Liu et al. 2009). Similarly, to check the binding of *Os*DOF18 with AAAG motifs located in the *Os*RGLP2 promoter, EMSA was carried out with probes containing one or two AAAG motifs. The specificity of interaction was confirmed by adding an excess of unlabeled probes and mutant probes. Shifts were also observed with all the probes when EMSA was carried out with nuclear proteins extracted from rice. The shifted bands were specific as confirmed by competition with an excess of unlabeled probes. We therefore conclude that the *Os*MYB1 and *Os*DOF18 transcription factors can directly interact with the *Os*RGLP2 promoter and transactivate *Os*RGLP2 expression.

In conclusion, the present study highlights the *in silico* identification of stress-responsive *cis*-regulatory elements in the promoter of the *Os*RGLP2 gene and their corresponding binding proteins. The DNA-binding domains of the identified proteins *Os*WRKY71, *Os*MYB1 and *Os*DOF18 were able to interact with the *Os*RGLP2 promoter as confirmed by EMSA. *Os*WRKY71 is characterized for its role in biotic stresses, but no work has been done on its role in response to abiotic stresses. *Os*MYB1 and *Os*DOF18 are the proteins yet to be characterized in rice. The availability of more information regarding these proteins in rice may further enhance our understanding about gene expression and its manipulation during biotic as well as abiotic stresses.

Author contribution statement FD performed most of the laboratory work and participated in all other activities. TS participated in laboratory experiments. TM was involved in planning the study. COS participated in laboratory work and bioinformatics analysis. KS was involved in planning the experiments and provided laboratory facilities in Denmark. SMSN planned and supervised all the work, provided laboratory facilities and also handled the manuscript as corresponding author. All authors read and approved the final manuscript.

Acknowledgements This work was supported by the grant of Higher Education Commission (HEC) of Pakistan through an Indigenous Fellowship [Pin: 117-3978-BM7-083 (50018550)] and International Research Support Initiative Program Fellowships to Ms. Farah Deeba (Pin: IRSIP25BMS08) and HEC Funded Research Project NRPU1909 to S.M. Saqlan Naqvi.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Abe H, Urao T, Ito T, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2003) *Arabidopsis AtMYC2* (bHLH) and *AtMYB2* (MYB) function as transcriptional activators in abscisic acid signaling. Plant Cell 15:63–78

Bergholtz S, Andersen TO, Andersson KB, Borrebæk J, Lüscher B, Gabrielsen OS (2001) The highly conserved DNA-binding

domains of A-, B- and c-Myb differ with respect to DNA-binding, phosphorylation and redox properties. Nucleic Acids Res 29(17):3546–3556

Bohnert HJ, Nelson DE, Jensen RG (1995) Adaptations to environmental stresses. Plant Cell 7:1099–1111

Carter C, Thornburg RW (2003) The nectary-specific pattern of gene expression is regulated by multiple promoter elements in the tobacco nectarin I promoter. Plant Mol Biol 51:451–457

Cheng X, Huang X, Liu S, Tang M, Hu W, Pan S (2014) Characterization of germin-like protein with polyphenol oxidase activity from *Satsuma mandarine*. Biochem Biophys Res Commun 449:313–318

Ciolkowski I, Wanke D, Birkenbihl R, Somssich I (2008) Studies on DNA-binding selectivity of WRKY transcription factors lend structural clues into WRKY-domain function. Plant Mol Biol 68:81–92

De Freitas CD, Da Cruz WT, Silva MZ, Vasconcelos IM, Moreno FB, Moreira RA, Monteiro-Moreira ACO, Alencar LMR, Sousa JS, Rocha BAM, Ramos MV (2016) Proteomic analysis and purification of an unusual germin-like protein with proteolytic activity in the latex of *Thevetia peruviana.* Planta 243:11–14

Dunwell JM, Gibbings JG, Mahmood T, Naqvi SMS (2008) Germin and germin-like proteins: evolution, structure, and function. Crit Rev Plant Sci 27(5):342–375

Escobar C, Aristizábal F, Navas A, del Campo FF, Fenoll C (2001) Isolation of active DNA-binding nuclear protein from tomato galls induced by root-knot nematodes. Plant Mol Biol Rep 19:375a–375h

Eulgem T, Somssich IE (2007) Networks of WRKY transcription factors in defense signaling. Curr Opin Plant Biol 10:366–371

Goicoechea M, Lacombe E, Legay S, Mihaljevic S, Rech P, Jauneau A, Lapierre C, Pollet B, Verhaegen D, Chaubet-Gigot N, Grima-Pettenati J (2005) *Eg*MYB2, a new transcriptional activator from eucalyptus xylem, regulates secondary cell wall formation and lignin biosynthesis. Plant J 43:553–567

Himmelbach A, Liu L, Zierold U, Altschmied L, Maucher H, Beier F, Muller D, Hensel G, Heise A, Schutzendubel A, Kumlehnand J, Schweizera P (2010) Promoters of the barley germin-like GER4 gene cluster enable strong transgene expression in response to pathogen attack. Am Soc Plant Biol 22:937–952

Hussain S (2015) Effect of 5′ deletions on efficiency and specificity of *Os*RGLP2 promoter. PhD thesis, PMAS-Arid Agriculture University, Rawalpindi

Jiang Y, Yang B, Harris NS, Deyholos MK (2007) Comparative proteomic analysis of NaCl stress-responsive proteins in *Arabidopsis* roots. J Exp Bot 58:3591–3607

Kang HG, Foley RC, Onate-Sanchez L, Lin C, Singh KB (2003) Target genes for OBP3, a DOF transcription factor, include novel basic helix-loop-helix domain proteins inducible by salicylic acid. Plant J 35:362–372

Konishi M, Yanagisawa S (2007) Sequential activation of two DOF transcription factor gene promoters during vascular development in *Arabidopsis thaliana*. Plant Physiol Biochem 45:623–629

Li HY, Jiang J, Wang S, Liu FF (2010) Expression analysis of *Th*GLP, a new germin-like protein gene, in *Tamarix hispida*. J For Res 21:323–330

Liu X, Bai X, Wang X, Chan C (2007) *Os*WRKY71, a rice transcription factor, is involved in rice defense response. J Plant Physiol 164:969–979

Liu G, Ren G, Guirgis A, Thornburg RW (2009) The MYB305 transcription factor regulates expression of nectarin genes in the ornamental tobacco floral nectary. Plant Cell 21(9):2672–2687

Mahmood T, Hyder MZ, Naqvi SMS (2007) Cloning and sequence analysis of germin-like protein gene 2 promoter from *Oryza sativa* L. ssp. Indica. DNA Seq 18:26–32

Maleck K, Levine A, Eulgem T, Morgan A, Schmid J, Lawton KA, Dangl JL, Dietrich RA (2000) The transcriptome of *Arabidopsis thaliana* during systemic acquired resistance. Nat Genet 26:403–410

Mare C, Mazzucotelli E, Crosatti C, Francia E, Stanca AM, Cattivelli L (2004) *Hv*-WRKY38: a new transcription factor involved in cold and drought-response in barley. Plant Mol Biol 55:399–416

Matić S, Bagnaresi P, Biselli C, Carneiro GA, Siciliano I, Valé G et al (2016) Comparative transcriptome profiling of resistant and susceptible rice genotypes in response to the seedborne pathogen *Fusarium fujikuroi*. BMC Genom 17(1):608

Rodriguez-Lopez M, Baroja-Fernandez E, Zandueta-Criado A, Moreno-Bruna B, Munoz FJ, Akazawa T, Pozueta-Romero J (2001) Two isoforms of a nucleotide-sugar pyrophosphatase/phosphodiesterase from barley leaves (*Hordeum vulgare* L.) are distinct oligomers of *Hv*GLP1, a germin-like protein. FEBS Lett 490:44–48

Sassaki FT, Bravo JP, Gonzalez ER, Maia IG (2014) Expression pattern and promoter analysis of a *Eucalyptus grandis* germin-like gene. Plant Mol Biol Rep 33:12–21

Segarra CI, Casalongúe CA, Pinedo ML, Ronchi VP, Conde RD (2003) A germin-like protein of wheat leaf apoplast inhibits serine proteases. J Exp Bot 54:1335–1341

Wang T, Chen X, Zhu F, Li H, Li L, Yang Q, Chi X, Yu S, Liang X (2013) Characterization of peanut germin-like proteins, AhGLPs in plant development and defense. PLoS ONE 8:61722

Xie Z, Zhang ZL, Zou X, Yang G, Komatsu S, Shen QJ (2006) Interactions of two abscisic-acid induced WRKY genes in repressing gibberellin signaling in aleurone cells. Plant J 46(2):231–242

Yasmin T, Mumtaz A, Mahmood T, Hyder MZ, Naqvi SMS (2015) A germin-like protein gene of rice increased superoxide dismutase activity in transformed tobacco. Biol Plant. doi:10.1007/s10535-015-0524-z

Yu L, Chen H, Guan Q, Ma X, Zheng X, Zou C, Li Q (2012) *At*MYB2 transcription factor can interact with the CMO promoter and regulate its downstream gene expression. Biotechnol Lett 34:1749–1755

Zhang Z-L, Xie Z, Zou X, Casaretto J, T-hD Ho, Shen QJ (2004) A rice WRKY gene encodes a transcriptional repressor of the gibberellin signaling pathway in aleurone cells. Plant Physiol 134:1500–1513