



***GmPTF1*, a novel transcription factor gene, is involved in conferring soybean tolerance to phosphate starvation**

X.H. Li, B. Wu, Y.B. Kong and C.Y. Zhang

North China Key Laboratory for Crop Germplasm Resources of Education Ministry,
Department of Plant Genetics and Breeding, Agricultural University of Hebei,
Baoding, Hebei, China

Corresponding author: C.Y. Zhang
E-mail: zhangcaiying@hebau.edu.cn

Genet. Mol. Res. 13 (1): 926-937 (2014)
Received December 13, 2012
Accepted June 16, 2013
Published February 19, 2014
DOI <http://dx.doi.org/10.4238/2014.February.19.3>

ABSTRACT. Phosphorus plays a pivotal role in plant growth and development. In this study, we isolated and characterized *GmPTF1*, a basic helix-loop-helix (bHLH) transcription factor (TF) gene from soybean (*Glycine max*) with tolerance to inorganic phosphate (Pi) starvation. Alignment analysis indicated that GmPTF1 and other reported bHLH TFs share significant similarity in the region of the bHLH domain. As with OsPTF1 and other homologous Pi starvation-related bHLH TFs (His-5, Glu-9, Arg-12, and Arg-13), all recognition motifs for the G-box (CACGTG) were present in the GmPTF1 domain. Prokaryotic expression in *Escherichia coli* strain BL21 (DE3) plyS showed that a novel 40-kDa polypeptide was expressed when cells containing *GmPTF1* were induced. The subcellular localization in cells from onion epidermis and *Arabidopsis* roots demonstrated that the GmPTF1 protein was found in the nucleus. Furthermore, analysis of transcription activity in yeast revealed that full-length GmPTF1 and its N-terminal and C-terminal domains could activate the histidine, adenine, and uracil reporter genes. This suggested that the N-terminal

and C-terminal peptides of *GmPTF1* act as transcriptional activators. When real-time quantitative polymerase chain reaction was performed, the expression of *GmPTF1* under conditions of phosphate starvation was significantly induced in soybean roots of the low-Pi-tolerant variety ZH15. Moreover, the relative level of expression was much higher there than in roots of the sensitive variety NMH from days 7 to 56 of low-Pi stress. These results imply that *GmPTF1* is involved in conferring tolerance to phosphate starvation in soybean.

Key words: *GmPTF1*; Basic helix-loop-helix (bHLH) transcription factor; Phosphate starvation; Low phosphate stress tolerance; Soybean

INTRODUCTION

Phosphorus is one of the most important nutrients for plant growth and metabolism. However, it is also one of the least available of all essential nutrients in the soil (Raghothama, 1999). The available concentration of inorganic phosphate (Pi) is generally below that of many other micronutrients, being present in concentrations of only about 1 to 2 μM in most sandy, alkaline, or acidic soils (Yang and Finnegan, 2010). Even in the most fertile soils, the Pi concentration in a solution rarely exceeds 8 μM (Smith et al., 2003). Therefore, in many countries, this low level of availability, consequently resulting in a deficiency, is the major constraint on plant productivity and focus of efforts to improve crop quality (Raghothama, 1999).

In response to persistently low levels of available Pi in the rhizosphere, plants employ many highly specialized physiological, biochemical, and molecular mechanisms to acquire it from the environment. Those steps might include modifications to root growth and architecture (Williamson et al., 2001; Chevalier et al., 2003; Lambers et al., 2006; Li et al., 2012b), formation of proteoid roots (Watt and Evans, 1999), secretion of organic acids (Shen et al., 2005), induction of acid phosphatases (Baldwin et al., 2001; Miller et al., 2001), or activation of Pi starvation-response genes (Wu et al., 2003; Franco-Zorrilla et al., 2004). This last phenomenon seems to be essential and universal within all organisms (Rubio et al., 2001; Hammond et al., 2004). Recently, many molecular biology laboratories have identified hundreds of genes that were associated with the phosphorus metabolic pathway (Li et al., 2012a). However, only a few of those genes have been used in transgenic breeding programs, and the functions of most related genes remain unclear.

Basic helix-loop-helix (bHLH) proteins are found throughout the eukaryotic kingdom. They constitute the second largest family of transcription factors (TFs) in plants, where they are key regulators of important metabolic, physiological, and developmental processes (Pires and Dolan, 2010). These TFs are believed to confer roles in light and hormone signaling, responses to wounding and drought stress, phosphate and iron homeostases, shoot branching, the formation of fruit and flowers, and the development of microspores, trichomes, stomata, and roots (Carretero-Paulet et al., 2010). For example, Lc from maize was the first plant bHLH protein to be described. This TF cooperates with MYB factors C1 (colorless 1) and PL1 (purple plant 1) to regulate the anthocyanin pathway (Ludwig et al., 1989). To date, numerous bHLH proteins have been reported that could regulate downstream genes through sequence-specific interactions in their promoter regions (Feng et al., 2012). In *Arabidopsis*, transparent testa glabra 1

(TTG1) and specific MYB partners can combine with bHLH proteins transparent testa 8 (TT8), glabra 3 (GL3), or enhancer of glabra 3 (EGL3) to control anthocyanin synthesis: GL3 and/or EGL3 initiates trichomes and root hairs, TT8 or EGL3 controls seed coat mucilage, and TT8 regulates proanthocyanidin synthesis in seed coats (Ramsay and Glover, 2005).

In *Arabidopsis*, bHLH32 negatively regulates a range of Pi starvation-induced processes, including activity by phosphoenolpyruvate carboxylase kinase, which is involved in modifying metabolism so that Pi is spared (Chen et al., 2007). Another key TF, OsPTF1, with a bHLH domain, has been isolated from a cDNA library and confirmed to be responsive to a Pi deficiency. When *OsPTF1* is overexpressed in a low-Pi-sensitive rice variety, plants show enhanced Pi efficiency in both solution cultures and soil experiments (Yi et al., 2005; Yan et al., 2006). Real-time reverse transcription polymerase chain reaction (RT-PCR) has revealed that *ZmPTF1*, which was isolated from maize, is quickly and significantly upregulated in roots under phosphate-starvation conditions. Its overexpression in maize improves root development, enhances biomass production in both hydroponic cultures and sand pots, and promotes the formation of more tassel branches and larger kernels when plants are grown in low-phosphate soil (Li et al., 2011).

To study the response to low-Pi stress in soybean (*Glycine max*), we first isolated a novel bHLH TF gene, *Glycine max* phosphate-starvation-induced transcription factor (*GmPTF1*). We then characterized its expression and activation.

MATERIAL AND METHODS

Plant materials

From 156 soybean varieties, we screened one variety each that was tolerant (ZH15) and sensitive (NMH) to Pi starvation (Li et al., 2010). These were used to clone and analyze *GmPTF1*, a bHLH TF gene.

Seedling culture and treatment in pot with vermiculite

Seeds of ZH15 and NMH were surface washed and sterilized, and then they were treated for germination at 26°C for 2 days. After germination, seedlings were transplanted into tanks with Hoagland nutrient solution. The solution was well aerated and the pH was maintained between 5.8 and 6.0 with daily additions of 1.0 MKOH or HCl. The plants grew for 15 days in a clear plastic awning until the first trifoliolate leaves were fully developed. Then, they were treated with 2 phosphate concentrations (0.002 and 1.0 mM). At 0, 7, 14, 21, 28, 35, 42, 49, and 56 days after treatment, roots, stems, and leaves were harvested separately. All tissue samples were immediately frozen in liquid nitrogen and stored at -80°C for RNA extraction.

Extraction of total RNA and mRNA purification

Soybean total RNA of the ZH15 and NMH tolerant and sensitive varieties, respectively, was extracted from the roots, stems, and leaves that were harvested above using TRNzol Total RNA Reagent (Tiangen, China). The extracted total RNA was purified to obtain mRNA using a PolyAtract mRNA Isolation System according to manufacturer instructions (Promega, USA).

Homologous sequence BLAST and primer design

The databank of soybean expressed sequence tags was searched using the sequence for *OsPTF1* (AY238991) to clone, *in silico*, the homologous gene *GmPTF1*. From those results, we designed 2 special primer sets (F1/R1) with the Primer Premier 5.0 software from the start code region to the terminator region in order to clone the presumed homologous gene in soybean. Four restriction sites - *EcoRI* (GAATTC), *KpnI* (GGTACC), *XbaI* (TCTAGA), and *Sall* (GTCGAC) - were added to the primers, which are listed in Table 1.

Table 1. Primer oligonucleotide sequences and their applications.

Primer	Sequence (5' to 3')	Description	Product (bp)
F1	CAGGAATTCGGTACCCGCTCAATG GACCAAGCTCC	Cloning <i>GmPTF1</i>	1101
R1	CCTGTCGACTCTAGATCAGAGCTCCATCTTCAATTG		
F2	CAGGAATTCGGTACCCGCTCAATGGACCAAGCTCC	GmPTF1 sub-cellular localization	1098
R2	GGTGGTACCGAGCTCCATCTTCAATTGATTGAG		
F3	CAGGAATTCGGTACCCGCTCAATGGACCAAGCTCC	Activation ability analysis (GmPTF1)	1101
R3	CCTGTCGACTGCAGTCAGAGCTCCATCTTCAATTG		
F4	CAGGAATTCGGTACCCGCTCA ATGGACCAAGCTCC	Activation ability analysis (P1)	480
R4	CCTGTCGACTCA GTTGTTTTCAACCGCGGGATC		
F5	CAGGAATTCGGTACCCGCTCA ATTCAGAGGACCGCC	Activation ability analysis (P2)	348
R5	CCTGTCGACTCAGAAATCAATTCTAGGGTTAATG		
F6	CAGGAATTCGGTACCCGCTCA AGCCTCGACAGTTTA	Activation ability analysis (P3)	273
R6	CCTGTCGACTGCAGTCAGAGCTCCATCTTCAATTG		
F7	CTTGGCAGAAAGGGCTAGGAG	qPCR of <i>GmPTF1</i>	225
R7	GCCCCATCAGTAGCCAATAAAC		
F8	ATCTTGACTGAGCGTGGTTATTCC	qPCR of <i>Actin11</i>	126
R8	GCTGGTCTGGCTGTCTCC		

cDNA PCR amplification and acquisition of the gene open reading frame (ORF)

The cDNA of ZH15, which was isolated from plants grown under low-Pi or high-Pi conditions, was obtained with a PrimeScript™ 1st Strand cDNA Synthesis Kit (TaKaRa, China). Following amplification with the 2 designed primers (F1/R1), we identified the ORF of our desired gene. PCR products were separated by electrophoresis on a 1.0% agarose gel. The purified targeted fragments were recovered, cloned into the pGM-T vector, and transformed into *Escherichia coli* Top10 cells according to the manufacturer instruction manual (Tiangen). The selected positive clones were identified by using bacterium liquid PCR and digestion with restriction enzymes *EcoRI* and *Sall*. Afterward, the transformed clones were sequenced by Sangon, China.

Bioinformatic analysis of *GmPTF1*

The ORF of the cloned *GmPTF1* and its encoding protein sequence were analyzed with the following online bioinformatic software: for the conserved domain, <http://www.ncbi.nlm.nih.gov>; trans-membrane region, <http://www.cbs.dtu.dk/services/TMHMM>; signal peptide, subcellular localization, <http://wolfsort.org>; and predictions of nuclear localization signals, <http://cubic.bioc.columbia.edu/predictNLS/>. Multiple-sequence alignment of *GmPTF1* and its encoding protein was done with the DNAMAN software.

Prokaryotic expression of *GmPTF1*

The ORF of *GmPTF1*, with *EcoRI* and *SalI* sites at the 5'- and 3'-ends, was digested from a selected recombinant plasmid, pGM-T-*GmPTF1*. The digested products were then ligated with prokaryotic expression vector pET-28a. The recombinant plasmids identified here, pET-28a-*GmPTF1* and pET-28a, were transformed into *E. coli* BL21 (DE3) plysS. Following the pET system manual (Novagen, Germany), we induced protein expression in the *E. coli* cells with pET-28a-*GmPTF1*, pET-28a, and BL21 (DE3) with isopropyl β -D-1-thiogalactopyranoside (IPTG) for 0, 3, 6, 9, or 18 h under the same culturing conditions. The protein extracts from those induced cells above were analyzed via sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

Subcellular protein localization of GmPTF1

The ORF of *GmPTF1* (TGA codon was omitted) with the *KpnI* site was obtained by amplifying pGM-T-*GmPTF1* with primers F2 and R2. Then, the ORF was digested by *KpnI* and ligated with the eukaryotic expression vector pCamE-*GFP*. The newly identified recombinant plasmids pCamE-*GmPTF1-GFP* and pCamE-*GFP* were transformed into onion epidermis cells with a gene gun and the Biolistic PDS-1000/He Particle Delivery System (Bio-Rad, USA). The transformed onion cells were cultured for 24 h under darkness on MS solid medium at 26°C before being examined with a fluorescence microscope to determine the subcellular location of GmPTF1. Concurrently, recombinant plasmid pCamE-*GmPTF1-GFP* and vector control pCamE-*GFP* were transferred to *Agrobacterium tumefaciens* strain GV3101 to transform *Arabidopsis* ecotype *Columbia*.

Detecting the activation domain of GmPTF1

The Matchmaker yeast (*Saccharomyces cerevisiae*) one-hybrid system was used to detect the activation domain of GmPTF1 according to the manufacturer protocol. Four pairs of primers (F3/R3, F4/R4, F5/R5, and F6/R6) were designed by Primer Premier 5.0. The deduced amino acid sequence of GmPTF1 was divided into 3 fragments: P1 (N-terminal fragment including 160 amino acids that did not contain the bHLH domain; primers F4/R4), P2 (fragment containing amino acids 161 to 276, which includes the bHLH domain; primers F5/R5), and P3 (the final 90 amino acids in the C-terminus; primers F6/R6). Insertion into the pGBKT7 vector resulted in a fusion with the GAL4 DNA binding domain (BD). We performed an α -galactosidase assay according to the CLONTECH protocol, using the pGBKT7 vector as a negative control.

Fluorescent quantitative real-time PCR (qPCR) of *GmPTF1*

RNA was isolated from soybean samples using a TRNzol Total RNA Reagent according to the manufacturer standard protocol (Tiangen). The quality and quantity of each sample were detected by agarose gel electrophoresis. The cDNA was synthesized from the total RNA using a PrimeScript® RT Reagent Kit with gDNA Eraser (TaKaRa) according to manufacturer instructions. *Actin11* served as the internal control for qPCR. Based on the sequences for *GmPTF1* and *Actin11*, 2 pairs of qPCR primers were designed (F7/R7; F8/R8). Real-time

PCR was conducted with SYBR® Premix Ex Taq™ II (TaKaRa) and a Roche Light Cycler 2.0. Reaction conditions included an initial step at 95°C for 30 s and 40 cycles of 95°C for 15 s, 58°C for 20 s, and 72°C for 30 s. Relative fold-changes in expression were calculated by the 2^{-ΔΔCt} method.

RESULTS

PCR amplification of the *GmPTF1* ORF

Using a PrimeScript™ 1st Strand cDNA Synthesis Kit, we obtained the cDNA of ZH15 soybeans when plants were exposed to either low- or high-Pi conditions. The cDNA then was amplified with 2 primers (F1/R1). An amplified band (about 1100 bp) was detected by agarose gel electrophoresis from the low-Pi cDNA but not from the high-Pi cDNA. After this polymorphic band was purified and cloned into the pGM-T vector, positive clones were identified by bacteria liquid PCR and restriction enzyme digestion, and then they were sequenced.

Sequenced results were analyzed through an NCBI search, which revealed that the ORF of the cloned gene was 1101 bp long (Figure 1). Moreover, a deduced peptide containing 366-amino acid residues had a predicted molecular weight of 40 kDa, including a bHLH transcription regulatory domain. We named this gene *GmPTF1* (*Glycine max* phosphate starvation-induced transcription factor 1), which is now designated as GenBank accession No. FJ617239.

```

atggaccaagctccggaggattgcggatccgttcaccacggcgccacgctcgggtcaattcaatcaacgaagaa
M D Q A P G G L R G S V H H G A T L G S I Q F N E E
attcagggaactcatggcgcggcgggaaacgcgagctcgttcaccgcgctgctgaactccgcgcgacgagcg
I Q G L M A P A P E N A S S F T A L L E L P P T Q A
gtggagctccttcaactcgcgggaatctgacagcgtcggaaaccgctaatgccatgacagcgaaccagaacc
V E L L H S P E S D S A R K P P N C H V S A N Q K P
tacctgctcaattcttcggtagcaatttgacctcccttccaagtcgcttaatcgaacgcccgcgaagtctcc
Y L L N S F G S N L T F P S N V A L I E R A A K F S
gttttcgocggcagaactcgcgctccgcggggaggcgcgctgattcagcaggaacgggttcgactttggac
V F A G E N S P P P P G E A R L I P A G T G S T L D
agagtgaagaacgagcccaagagaccgattcgaaccgctgctcagctcagggctgggttcatttcggatcccgc
R V K N E P Q E T D S N P C S S S R L G C I S D P A
gttgaacaacattcagaggaccgccaagaaagagcagagagaagaggcctaaaggatcgtgaagaagcgttaag
V E N N I Q R T A K R K E R E K K A K G S S K K R K
agcgcgcccgatgagacatccggtagcggcagagaagcttccctacgtccatgttcgagttcgtcaggtcaagccact
S A A D E T S G D G E K L P Y V H V R V R R G Q [A]
gatagccatagcttggcagaaaggctaggagagagaagaatcaacgctcgcgatgaagcttttgcagagctgtccca
D S H S L A E R A R R E K I N A R M K L L Q E L V P
ggctgcgacaagaatcagggacggcaatgggttctggatgaaatcaccatgcaatctctacagcgtcaagtg
G C D K I S G T A M V L D E I I N H V Q S L Q R Q V
gagatcttatcaatgaaactggcagcagttaccctagaattgattcagcctcagacagtttatggctactgatggg
E I L S M K L A A V N P R I D F S L D S L L A T D G
gcattctaatggataataacttacctagcatggtagacacctcctcagtgggccagaatcccactcaatggcaacgg
A S L M D N N L P S M V T P L M W P E I P L N G N R
cagcattatcaacagcagtgccgattggatgcatccaccaacctctctgggagagggagaagaagcaaccaatctt
Q H Y Q Q Q W R L D A F H Q P L W E R E E V N H N F
atgactccggaactcacttttgattatgactctggcaaacctcagcttctctgcactcaaatcaattgaagatg
M T P E N S L L S Y D S S A N S A S L H S N Q L K M
gagctctga
E L *

```

Figure 1. Nucleotide sequence of transcription factor *GmPTF1* and primary structure of the GmPTF1 protein. Lowercase letters represent the nucleotide sequence of *GmPTF1* and capital letters represent its deduced amino acid residues. Boxed letters show the bHLH domain of GmPTF1. Gray letters show the nuclear localization signal sequence in GmPTF1.

Characterization of *GmPTF1* and its encoded protein

The analysis of the conserved domain for the encoded GmPTF1 protein showed that the sequence contained a bHLH domain, a DNA-binding region, an E-box/N-box specificity site, and a dimerization interface. These domains are also present in OsPTF1 from rice. Bioinformatic analysis also indicated that GmPTF1 has a potential calmodulin-binding site (amino acid position 210-236) but it has neither a transmembrane region nor a signal peptide. The scanning of GmPTF1 predicted sites for N-glycosylation, N-myristoylation, and the phosphorylations of cAMP- and cGMP-dependent protein kinase, protein kinase C, and casein kinase II. The subcellular analysis showed that the protein was located in the nucleus, with the nuclear localization signal sequence KRKEREKKAKGSSKRRK also being found at amino acids 166 to 182 (Figure 1).

Alignment via the DNAMAN software indicated significant similarities between GmPTF1 and some bHLH proteins in the region of the bHLH domain (Figure 2). Moreover, just as for OsPTF1, PHO4, and other low Pi tolerance-related homologous bHLH TFs, we noted the presence of His-5, Glu-9, Arg-12, and Arg-13 in the basic region of GmPTF1 (Figure 2). These residues constituted the recognition motifs for the G-box (a kind of E-box, CACGTG), leading us to conclude that GmPTF1 is a G-box-binding type transcription regulatory factor. The full-length DNA of *GmPTF1* is 8275 bp long, including 7 exons and 6 introns (Figure 3). The longest intron is about 5365 bp; the shortest is only 116 bp.

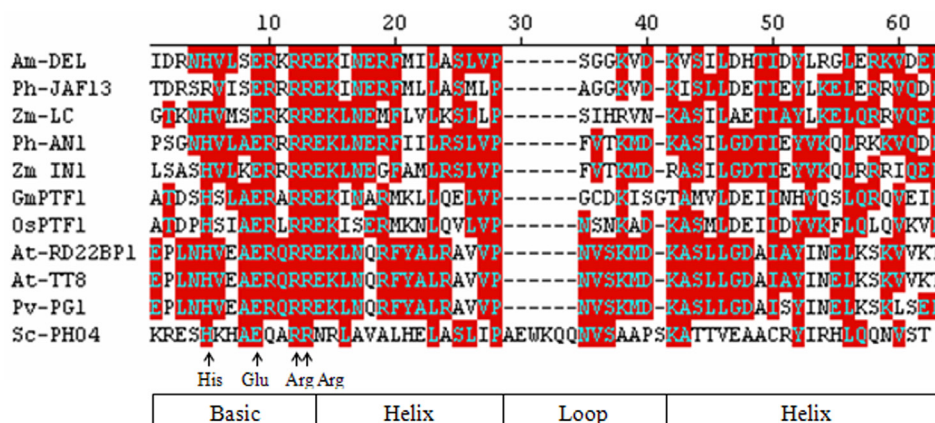


Figure 2. Basic helix-loop-helix domain of GmPTF1 shares highest identity with OsPTF1.

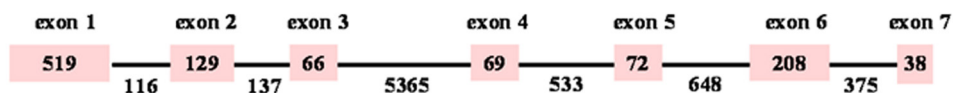


Figure 3. Genome sequence of *GmPTF1* (8275 bp).

Prokaryotic expression of *GmPTF1*

The recombinant plasmid pGM-T-*GmPTF1* was digested by *EcoRI* and *SalI*. After the target band was ligated with pET-28a, the recombinant plasmids pET-28a-*GmPTF1* and pET-28a were transformed into *E. coli* BL21 (DE3) plysS. Positive clones were identified after bacterium liquid PCR and digestions with restriction enzymes. Subsequently, clones that contained the correct sequence for *GmPTF1* were induced.

E. coli BL21 (DE3) plysS cells with pET-28a-*GmPTF1* showed by SDS-PAGE that a novel polypeptide with a molecular mass of about 44 kDa was expressed (Figure 4). Expression products from 3 unattached clones were coincidental. The novel polypeptide contained *GmPTF1* protein (40 kDa) and pET-28a tag fusion protein (4 kDa). However, under the same culturing conditions, this polypeptide could not be expressed in either control (*E. coli* BL21 DE3 plysS cells with pET-28a or untransformed *E. coli* BL21 DE3 plysS cells). This indicated that *GmPTF1* was successfully expressed prokaryotically in *E. coli* BL21 (DE3) plysS.

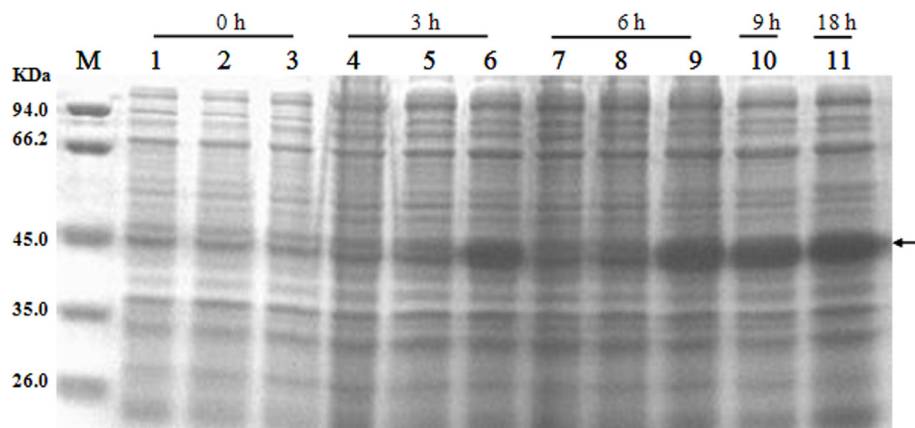


Figure 4. SDS-PAGE of *GmPTF1* protein expression in *Escherichia coli* BL21 (DE3) plysS. Lane M = protein molecular marker; lanes 1, 4, 7 = BL21(DE3) plysS induced 0, 3, 6 h; lanes 2, 5, 8 = BL21(DE3) plysS with pET-28a induced 0, 3, 6 h; lanes 3, 6, 9, 10, 11 = BL21(DE3) plysS with pET-28a-*GmPTF1* induced 0, 3, 6, 9, 18 h.

Subcellular protein localization of *GmPTF1*

Transformed onion epidermis cells and transgenic *Arabidopsis* root cells containing pCamE-*GmPTF1*-GFP and pCamE-GFP were examined with a fluorescence microscope. The translated protein *GmPTF1*-GFP was located in the nuclei of both plant sources (Figure 5) whereas 35S-GFP was expressed constitutively. This result was consistent with our prediction from bioinformatic analysis, suggesting that its function correlated with that of its TF.

Capacity for transcription activation by *GmPTF1*

To analyze its capacity for transcription activation, we expressed *GmPTF1* and its 3 subclones in yeast strain AH109 and fused them to the DNA-BD of the yeast TF *GAL4*. The 4 clones were designated BD-*GmPTF1*, BD-P1, BD-P2, and BD-P3.

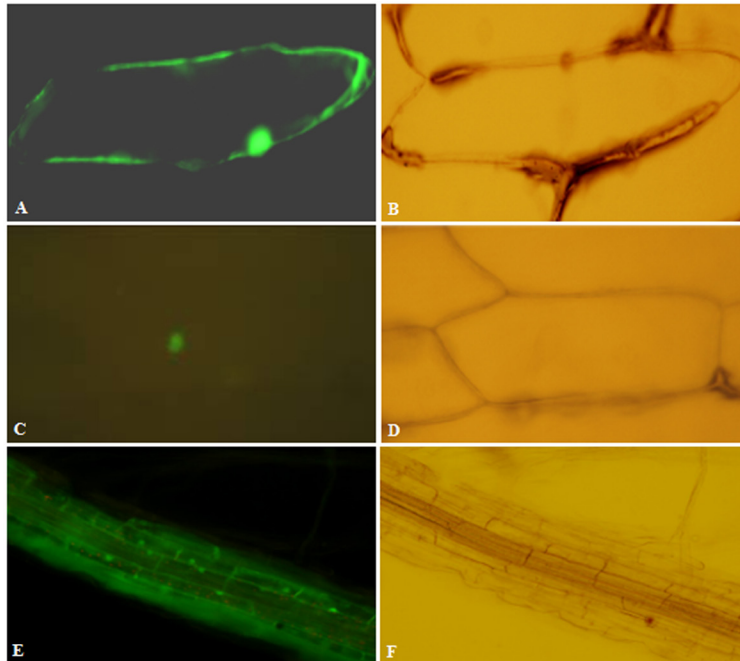


Figure 5. Sub-cellular localization of GmPTF1 in cells from onion epidermis and *Arabidopsis* roots. **A.** Microscopic image of onion cells harboring a control gene 35S::GFP. **B.** The same cells in A with transmitted light. **C.** Microscopic image of onion cells harboring a 35S::GmPTF1::GFP fusion gene. **D.** Same cells in C with transmitted light. **E.** Microscopic image of *Arabidopsis* root cells harboring a 35S::GmPTF1::GFP fusion gene. **F.** Same cells in E with transmitted light.

When tryptophan (Trp) was absent from the growth medium, all constructs grew well, indicating that the 4 vectors had been successfully constructed and introduced separately into the yeast (Figure 6). When the medium lacked histidine (-His), tryptophan (-Trp), uracil (-Ura), and adenine (-Ade), only the yeast clones with BD-GmPTF1, BD-P1, and BD-P3 grew well, indicating that full-length GmPTF1 and its N-terminal and C-terminal domains were able to activate the reporter genes for His, Ade, and Ura (Figure 6). We also tested the activation of a MEL reporter gene with α -galactosidase assays. A blue coloring in the yeast clones indicated that the α -galactosidase gene also could be activated by the TF GmPTF1 and its N-terminal and C-terminal domains. Therefore, results from our reporter gene assays demonstrated that the N-terminal and C-terminal peptides of GmPTF1 act as transcription activation domains.

Expression profile of *GmPTF1* under low-Pi stress

To investigate the expression pattern of *GmPTF1*, we conducted real-time qPCR analysis under low-Pi stress. Transcripts were more abundant in all organs sampled from the tolerant variety ZH15 than from the sensitive variety NMH, especially in the roots, where expression was at least 7-fold greater (Figure 7A). When transcripts were evaluated at various time points during the stress period, we noted that expression was induced by Pi starvation in the roots from day 7 to day 56 and that the relative expression was much higher in ZH15 than

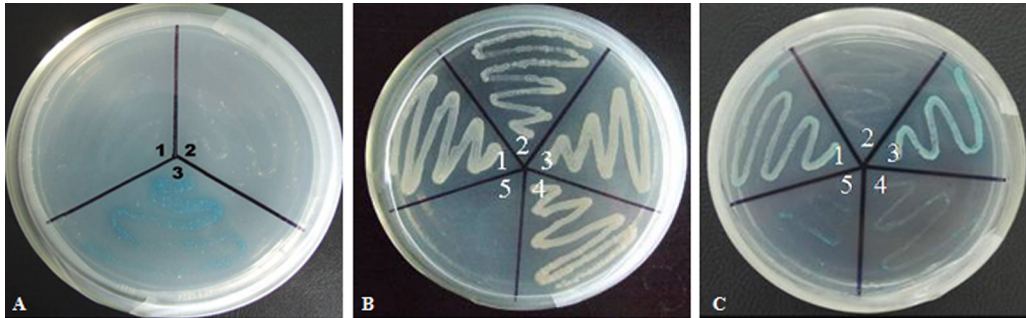


Figure 6. Yeast system used to detect activating domain of GmPTF1. **A.** Transformed yeast strain AH109 with constructs of BD-GmPTF1 grew well in synthetic dextrose growth medium with X- α -gal without Trp, His, Ade, Ura (SD/-Trp/-His/-Ade/-Ura/X- α -gal). 1 = AH109 control; 2 = AH109 with pGBKT7 vector control; 3 = AH109 with pGBKT7-GmPTF1. **B.** Yeast with constructs BD-P1 (N-terminal fragment including 160 amino acids that did not contain the bHLH domain), BD-P2 (fragment containing amino acids 161 to 276, which included the bHLH domain), BD-P3 (the final 90 amino acids in the C-terminus), and BD control (AH109 with pGBKT7) grew well in synthetic dextrose growth medium without Trp. 1 = BD-P1; 2 = BD-P2; 3 = BD-P3; 4 = AH109 with pGBKT7 control; 5 = AH109 control. **C.** Yeast AH109 with constructs BD-P1 and BD-P3 grew well in SD/-Trp/-His/-Ade/-Ura/X- α -gal medium, but yeast with constructs BD-P2 could not grow in SD/-Trp/-His/-Ade/-Ura/X- α -gal medium. 1 = BD-P1; 2 = BD-P2; 3 = BD-P3; 4 = AH109 with pGBKT7 control; 5 = AH109 control.

in NMH (Figure 7B). Therefore, these findings suggested that *GmPTF1* from soybean functions primarily in the roots and that it might be involved in external phosphorus assimilation and the transfer of Pi from the roots to other tissues or organs.

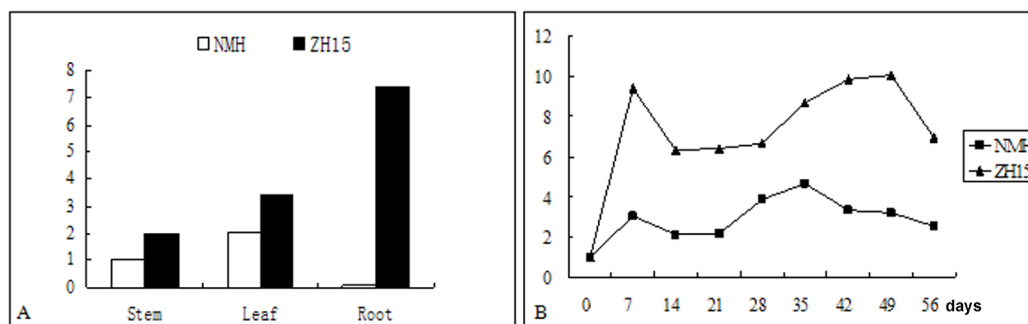


Figure 7. Expression of *GmPTF1* in NMH (sensitive) or ZH15 (tolerant) soybean varieties.

DISCUSSION

In this study, we cloned a novel bHLH TF gene, *GmPTF1*. Its expression was characterized in tolerant and sensitive soybean varieties that were grown under Pi starvation. The deduced protein contained a typical bHLH domain and was closely related to OsPTF1, the first TF that was reported to confer tolerance to Pi starvation in rice. In that report, Yi et al. (2005) evaluated how plants responded to Pi-deficient conditions. When *OsPTF1* was over-

expressed, tillering ability, root and shoot biomasses, and P contents were about 30% higher in the transgenic plants than in the wild-type plants when grown hydroponically, while tiller numbers, panicle weights, and P contents were 20% higher in transgenic plants when they and the wild-type plants were grown in either soil pots or the field. Similarly, our real-time qPCR showed that, under Pi-deficient conditions, *GmPTF1* expression was much higher in tolerant ZH15 than in sensitive NMH from days 7 to 56 of the stress period. This was especially true in the roots, which implied that this gene is involved in the assimilation of external phosphorus from the soil.

TFs bind to specific DNA sequences and regulate gene expression by altering the ability of RNA polymerase to bind to a target promoter sequence. These proteins have vital roles in many plant responses to stress conditions, including Pi deprivation (Yang and Finnegan, 2010). Generally, TFs have at least 2 discrete domains - a DNA-binding domain and an activation or repression domain - that operate together to modulate the rate of transcriptional initiation from target gene promoters (Heim et al., 2003). The bHLH TF family is defined by a bHLH DNA-binding domain that has 2 components: 1) a basic region involved in the binding of DNA cis-regulatory elements with the consensus CANNTG and 2) a hydrophobic HLH region that functions as a homo- or hetero-dimerization domain (Carretero-Paulet et al., 2010). Our NCBI analysis of the deduced protein GmPTF1 revealed a bHLH domain, an E-box/N-box specificity site, and a dimerization interface in the protein sequence, as well as recognition motifs, His-5, Glu-9, Arg-12, and Arg-13, for a G-box. These findings demonstrated that GmPTF1 is a bHLH TF with a recognition sequence for a G-box.

Analysis by Yi et al. (2005) of the capacity for rice OsPTF1 to be activated in yeast showed that only the full-length OsPTF1 and its N-terminal domain can activate the His and Ade reporter genes. Moreover, OsPTF1 that lacks the N-terminal domain cannot activate the reporter gene, thereby indicating that the N-terminal peptide of OsPTF1 acts as the transcription activation domain. However, this was not the case for GmPTF1 in our study. Instead, both the full-length GmPTF1 and its N-terminal domain (the first 160 amino acids) were able to activate the His, Ade, Ura, and α -galactosidase reporter genes in yeast strain AH109. Likewise, the C-terminal domain (the final 90 amino acids) of GmPTF1 activated those reporter genes (Figure 7). Our findings indicate that both the N-terminal and C-terminal peptides of GmPTF1 can serve as the transcription activation domain.

In eukaryotes, TFs (like most proteins) are transcribed in the nucleus but then are translated in the cytoplasm. Many proteins contain nuclear localization signals that direct them to the nucleus. For many TFs, this is in fact a key point in their regulation. We found that the deduced GmPTF1 protein has a nuclear localization signal at amino acids 166 to 182. GFP fluorescence tests of onion epidermis and *Arabidopsis* root cells showed that GmPTF1-GFP was located in the nucleus. This is not only consistent with our prediction of the bioinformatic results but also suggests a possible correlation with its function as a transcription factor.

ACKNOWLEDGMENTS

Research supported by the National Project (#2009ZX08004-004B) from the Ministry of Agriculture of China, the Natural Science Foundation of China (#31071441), and the Natural Science Foundation of Hebei Province (#C2010000749).

REFERENCES

- Baldwin JC, Karthikeyan AS and Raghothama KG (2001). *LEPS2*, a phosphorus starvation-induced novel acid phosphatase from tomato. *Plant Physiol.* 125: 728-737.
- Carretero-Paulet L, Galstyan A, Roig-Villanova I, Martínez-García JF, et al. (2010). Genome-wide classification and evolutionary analysis of the bHLH family of transcription factors in *Arabidopsis*, poplar, rice, moss, and algae. *Plant Physiol.* 153: 1398-1412.
- Chen ZH, Nimmo GA, Jenkins GI and Nimmo HG (2007). BHLH32 modulates several biochemical and morphological processes that respond to Pi starvation in *Arabidopsis*. *Biochem. J.* 405: 191-198.
- Chevalier F, Pata M, Nacry P, Doumas P, et al. (2003). Effects of phosphate availability on the root system architecture: large-scale analysis of the natural variation between *Arabidopsis* accessions. *Plant Cell Environ.* 26: 1839-1850.
- Feng XM, Zhao Q, Zhao LL, Qiao Y, et al. (2012). The cold-induced basic helix-loop-helix transcription factor gene *MdC1bHLH1* encodes an ICE-like protein in apple. *BMC Plant Biol.* 12: 22-36.
- Franco-Zorrilla JM, Gonzalez E, Bustos R, Linhares F, et al. (2004). The transcriptional control of plant responses to phosphate limitation. *J. Exp. Bot.* 55: 285-293.
- Hammond JP, Broadley MR and White PJ (2004). Genetic responses to phosphorus deficiency. *Ann. Bot.* 94: 323-332.
- Heim MA, Jakoby M, Werber M, Martin C, et al. (2003). The basic helix-loop-helix transcription factor family in plants: a genome-wide study of protein structure and functional diversity. *Mol. Biol. Evol.* 20: 735-747.
- Lambers H, Shane MW, Cramer MD, Pearse SJ, et al. (2006). Root structure and functioning for efficient acquisition of phosphorus: Matching morphological and physiological traits. *Ann. Bot.* 98: 693-713.
- Li XH, Gai JY, Chang WS and Zhang CY (2010). Identification of phosphorus starvation tolerant soybean (*Glycine max*) germplasms. *Front. Agric. China* 4: 272-279.
- Li XH, Chang WS and Zhang CY (2012a). Research progress of candidate genes for improving plant phosphorus-efficiency. *J. Plant Genet. Resour.* 13: 83-97.
- Li Z, Gao Q, Liu Y, He C, et al. (2011). Overexpression of transcription factor *ZmPTF1* improves low phosphate tolerance of maize by regulating carbon metabolism and root growth. *Planta* 233: 1129-1143.
- Li Z, Xu C, Li K, Yan S, et al. (2012b). Phosphate starvation of maize inhibits lateral root formation and alters gene expression in the lateral root primordium zone. *BMC Plant Biol.* 12: 89.
- Ludwig SR, Habera LF, Dellaporta SL and Wessler SR (1989). *Lc*, a member of the maize R gene family responsible for tissue-specific anthocyanin production, encodes a protein similar to transcriptional activators and contains the myc-homology region. *Proc. Natl. Acad. Sci. U. S. A.* 86: 7092-7096.
- Miller SS, Liu J, Allan DL, Menzhuber CJ, et al. (2001). Molecular control of acid phosphatase secretion into the rhizosphere of proteoid roots from phosphorus-stressed white lupin. *Plant Physiol.* 127: 594-606.
- Pires N and Dolan L (2010). Early evolution of bHLH proteins in plants. *Plant Signal. Behav.* 5: 911-912.
- Raghothama KG (1999). Phosphate acquisition. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 50: 665-693.
- Ramsay NA and Glover BJ (2005). MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. *Trends Plant Sci.* 10: 63-70.
- Rubio V, Linhares F, Solano R, Martin AC, et al. (2001). A conserved MYB transcription factor involved in phosphate starvation signaling both in vascular plants and in unicellular algae. *Genes Dev.* 15: 2122-2133.
- Shen J, Li H, Neumann G and Zhang F (2005). Nutrient uptake, cluster root formation and exudation of protons and citrate in *Lupinus albus* affected by localized supply of phosphorus in a split-root system. *Plant Sci.* 168: 837-845.
- Smith FW, Mudge SR, Rae AL and Glassop D (2003). Phosphate transport in plants. *Plant Soil* 248: 71-83.
- Watt M and Evans JR (1999). Linking development and determinacy with organic acid efflux from proteoid roots of white lupin grown with low phosphorus and ambient or elevated atmospheric CO₂ concentration. *Plant Physiol.* 120: 705-716.
- Williamson LC, Ribrioux SP, Fitter AH and Leyser HM (2001). Phosphate availability regulates root system architecture in *Arabidopsis*. *Plant Physiol.* 126: 875-882.
- Wu P, Ma L, Hou X, Wang M, et al. (2003). Phosphate starvation triggers distinct alterations of genome expression in *Arabidopsis* roots and leaves. *Plant Physiol.* 132: 1260-1271.
- Yan XL, Wu P, Ling HQ, Xu GH, et al. (2006). Plant nutrionics in China: an overview. *Ann. Bot.* 98: 473-482.
- Yang XJ and Finnegan PM (2010). Regulation of phosphate starvation responses in higher plants. *Ann. Bot.* 105: 513-526.
- Yi K, Wu Z, Zhou J, Du L, et al. (2005). OsPTF1, a novel transcription factor involved in tolerance to phosphate starvation in rice. *Plant Physiol.* 138: 2087-2096.