2630–2644 *Nucleic Acids Research, 2009, Vol. 37, No. 8* *Published online 6 March 2009*
doi:10.1093/nar/gkp126

Tissue-specific expression of the *PNZIP* promoter is mediated by combinatorial interaction of different *cis*-elements and a novel transcriptional factor

Yu-Tao Yang[1,2], Yan-Li Yu[1], Guo-Dong Yang[1], Jie-Dao Zhang[1] and Cheng-Chao Zheng[1,*]

[1]State Key Laboratory of Crop Biology, College of Life Sciences, Shandong Agricultural University, Taian, Shandong 271018 and [2]Beijing Institute for neuroscience, Capital Medical University, Beijing 100069, P.R. China

Received October 30, 2008; Revised January 21, 2009; Accepted February 16, 2009

ABSTRACT

Recent studies demonstrated that *PNZIP* and its homologs encode a special cyclase and play an important role in chlorophyll biosynthesis in higher plants. To investigate the molecular mechanism governing the *PNZIP* gene, the *PNZIP* promoter was isolated and analyzed. Deletion analysis indicated that G-box is an important element in the regulation of the reporter gene expression. Further mutation assay demonstrated that G-box and GATA CT elements are necessary and sufficient for the high and tissue-specific expression of the *GUS* gene. Using yeast one-hybrid screening, we have isolated a novel tobacco bZIP protein, NtbZIP, which can specifically recognize the G-box of the *PNZIP* promoter. The NtbZIP protein shares a limited amino acid homology to *Arabidopsis* ABI5 and AtAREB1 and very low homology to other bZIP proteins. Northern blot analysis showed that the *NtbZIP* gene is not induced by exogenous ABA and is expressed in different tobacco organs. Cotransformation assays showed that the NtbZIP protein could activate the transcription of the *GUS* gene driven by the *PNZIP* promoter. Transgenic tobaccos analysis demonstrated that constitutively expressing antisense *NtbZIP* gene resulted in a lower NTZIP synthesis and reduced chlorophyll levels. We suggest that *NTZIP* is a target gene of *NtbZIP*, which is involved in the regulation of chlorophyll biosynthesis.

INTRODUCTION

Photosynthesis is the most important source of energy on the earth, and chlorophyll molecules play a central role in harvesting light energy channeled for photosynthesis. Chlorophyll, as a component of chloroplast, is bound to proteins of the photosynthetic membranes to harvest sunlight (1).

Chlorophyll production starts with the condensation of eight molecules of δ-aminolevulinic acid (ALA) to uroporphyrinogen III, the first cyclic tetrapyrrole. Then, uroporphyrinogen III is converted to protoporphyrin IX, which is the branch point between hemes and chlorophylls. Insertion of Mg ions into protoporphyrin IX by Mg chelatase results in the formation of Mg-protoporphyrin IX (MgP); then MgP is converted to MgP monomethyl ester (MgPMME) by a methyl transferase (2,3). MgPMME is next converted to divinyl protochlorophyllide (Pchlide) harboring the fifth ring characteristic of all chlorophylls by the cyclase reaction. In angiosperms, Pchlide is converted to chlorophyllide *a* by NADPH-Pchlide oxidoreductase and then a polyisoprene tail is added to finish chlorophyll *a* production through the light-dependent pathway.

Since chlorophyll is the principal pigment that traps light energy, the biosynthesis of chlorophyll has presented a number of challenging topics in the field of plant molecular biology (4). To date, many studies have explored the mechanism of biosynthesis of chlorophyll both in higher plants and in photosynthetic organisms. Many genes involved in biosynthesis of chlorophyll, including *xantha-f*, *xanfha-g*, *xanfha-h*, *bchl*, *bchH* and *Chl I* (5–8), have been isolated and characterized from photosynthetic organisms such as algae, bacteria and higher plants.

Chlorophyll biosynthesis has been extensively studied by genetic methods, and nearly all the enzymes have been identified at the molecular level in higher plants. One of the least understood enzymatic steps is the formation of the isocyclic ring, which is catalyzed by the Mg-protoporphyrin IX monomethyl ester (MgPMME) cyclase that is involved in the conversion of MgPMME to protochlorophyllide (Pchlide). We previously isolated and characterized a novel gene from the short-day plant *Pharbitis nil* that encodes a protein with a leucine zipper

*To whom correspondence should be addressed. Tel: +86 538 8242894; Fax: +86 538 8226399; Email: cczheng@sdau.edu.cn

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

motif, designated *PNZIP* (*pharbitis nil* leucine zipper). *PNZIP* is a single copy gene that is expressed specifically in photosynthetically active mesophyll cells but not in other nonphotosynthetic tissues such as epidermis, trischomes and vascular tissues (9). Two years later, two *PNZIP* homologs, *Crd1* (copper response defect gene) from a green alga and *AcsF* (aerobic cyclase system Fe-containing subunit gene) from a photosynthetic bacterium were characterized by mutant analysis (10,11). Crd1 was identified as a putative diiron enzyme required for photosystem I accumulation in copper deficiency (10). AcsF, in purple bacteria *Rubrivivax gelatinosus*, has been proven to be involved in the aerobic oxidative cyclization of Mg-Protoporphyrin IX mono-methyester, one of the intermediates in the synthesis of bacterio-chlorophyll (11). These results suggest that this class of genes might be involved in biosynthesis of photosystem components in bacteria and algae. Recently, three *PNZIP* homologs have been isolated and identified from higher plants. *CHL27* encodes a protein that is required for the synthesis of protochlorophyllide in *Arabidopsis* and was proven to be a candidate subunit of the aerobic cyclase in chlorophyll biosynthesis (12). *NTZIP* from tobacco was characterized by the antisense RNA strategy, and the results indicated that the *NTZIP* gene plays a vital role during chlorophyll biosynthesis in tobacco (13). *Xantha-l* from barley encodes a membrane subunit of the aerobic Mg-protoporphyrin IX monomethyl ester cyclase involved in chlorophyll biosynthesis (14). Taken together, these results indicate that *PNZIP* and its homologs encode a special cyclase that is involved in the conversion of MgPMME to protochlorophyllide (Pchlide) and play an important role in chlorophyll biosynthesis in higher plants.

To our knowledge, there have been no reports about the regulatory mechanism for the genes encoding cyclases in chlorophyll biosynthesis. Therefore, corresponding research on the regulation of this class of genes is necessary to further elucidate the regulatory mechanism. Tobacco, an important model plant, has been widely used as a heterologous system to study promoters and genes from other plants that may be more difficult to work with (15–17). In this study, we isolated the *PNZIP* promoter and used tobacco to analyze the cis-element and the transcription factors required for the high and tissue-specific expression of the *PNZIP* gene. We demonstrated that G-box and GATACT elements in the *PNZIP* promoter are sufficient to control the expression of the *GUS* reporter gene. Moreover, we have isolated a novel bZIP transcription factor, NtbZIP, interacting with the G-box of the *PNZIP* promoter, and observed that it can transactivate the reporter gene expression in transgenic tobaccos. Importantly, we demonstrated that *NTZIP* is a target gene of *NtbZIP*, suggesting its special function in regulation of chlorophyll biosynthesis.

MATERIALS AND METHODS

Plant materials and growth conditions

Tobacco (*Nicotiana tabacum* L. cv. NC89) seedlings grown in a growth chamber at 25°C with a 16-h light/8-h dark cycle (450 μmol photons $m^{-2} s^{-1}$) were used in this experiment. All plants were harvested at similar developmental stages.

Isolation of the 5′-upstream sequences of the *PNZIP* gene

The 5′-upstream region of the *PNZIP* gene was obtained using the Universal Genome Walker kit (CLONTECH, Palo Alto, CA, USA). First, separate aliquots of genomic DNA were digested with three blunt-end restriction enzymes (EcoRV, DraI and ScaI), and ligated to genome walker adaptors. Primary PCR was performed using adaptor primer 1 (AP1) and a *PNZIP* cDNA specific primer (5′-CTGCGACGTGGTGGCCCTCGAC AT-3′). The second PCR was performed using adaptor primer 2 (AP2) and the same *PNZIP* cDNA specific primer. The amplified PCR products were examined on an agarose gel, and subcloned into the pGEMT-T easy vector, and then sequenced. The plasmid construct harboring the *PNZIP* gene promoter region was designated as PN-1.

Identification of the *PNZIP* gene transcription start site

To identify the *PNZIP* gene transcription start site, 5′-RLM-RACE, based on RNA ligase-mediated and oligo-capping rapid amplification of cDNA, was carried out by using the GeneRacer™ kit (Invitrogen, Carlsbad, CA, USA). This kit ensures amplification of only full-length mRNA by elimination of truncated molecules from the amplification process. Briefly, total RNA was treated with calf intestinal phosphatase (CIP) and then full-length mRNAs was decapped with tobacco acid pyrophosphatase (TAP). An RNA oligonucleotide was liagated to the full-length decapped mRNAs. Then, ligated mRNAs was reverse transcribed with Oligo-dT primer and SuperScript™ II RT (Invitrogen). Finally, the 5′-cDNA end was amplified by nested PCR with two *PNZIP* gene specific primers (RA1:5′-CGAAGTGGGT CTGGTTGTAGT-3′, RA2: 5′-CTGGAGAACTCTCAA CTCGGAGT-3′). The amplified PCR products were cloned into the pGEMT-T easy vector and sequenced.

Tobacco transformation

The different expression vector plasmids were transferred into *Agrobacterium tumefaciens* LBA4404 by electroporation and then transformed into tobacco by the leaf-disc method described by Horsch *et al.* (18). Transformed plants were selected on MS medium containing 100 μg/ml kanamycin and 250 μg/ml carbenicillin. After regeneration (3 to 4 weeks), shoots were transferred to root-inducing medium for 2 to 3 weeks and then transferred to a greenhouse to generate T_0 plants. T_1 plants were obtained by *in vitro* sowing surface-sterilized seedlings of the inbred T_0 plants on MS medium containing 100 mg/l kanamycin to select transformed resistant plants.

Construction of reporter plasmids

All plasmids for expression assays were constructed on a pBI121 vector (Clontech, Palo Alto, CA, USA) as the backbone, with the replacement of the CaMV 35S

Table 1. Sequence of oligonucleotides used to create deletion constructs

Name	Sequence
P1	5′-AAGCTTACATGGGGATGAGGCGG-3′
P2	5′-AAGCTTAGTCAAGTTAATTAGGT-3′
P3	5′-AAGCTTCTATACACTCCACAGAC-3′
P4	5′-AAGCTTCAGCATTGGTGTTCCTT-3′
P5	5′-AAGCTTCAATCAAGCTGGCCTGTC-3′
P6	5′-AAGCTTCAGACCAATATTTAATCCCAT-3′
P7	5′-GGATCCGGGTAGAGTGTACTGT-3′
D1	5′-CCATGGTGACAGGCCAGCTTGATTAC-3′
D2	5′-CCATGGCTCAGCAATCTTAAATGT-3′
D3	5′-CCATGGAAATTGAATTTCACTATGT-3′
D4	5′-CCATGGCTAATTAACTTGACTAAGGT-3′

promoter sequence. A series of 5′-deletion mutations of the *PNZIP* promoter were generated by PCR amplification from plasmid PN-1 using different forward primers and a single downstream primer. These forward primers were P1, P2, P3, P4, P5 and P6, whereas the downstream primer was designated as P7. The full-length and five deleted derivatives were cloned into the pGEM-T easy vector and sequenced; then different length fragments cut by HindIII and BamHI were inserted into the pBI121 vector, replacing the CaMV 35S promoter. The resultant plasmids were named Q1, Q2, Q3, Q4, Q5 and Q6, according to the position at the 5′-end, respectively. The primers used are shown in Table 1.

To construct plasmids for 3′-deletion mutations, the regions from –1415 to –115, –1415 to –375, –1415 to –570 and –1415 to –1090 of the *PNZIP* promoter were amplified by PCR from plasmid Q1 using different downstream primers (D1, D2, D3 and D4) and a single upstream primer (P1). The corresponding PCR products were cloned into the pGEM-T easy vector and sequenced, and then the different length fragments were cloned into the HindIII/NcoI sites of the Q6 construct, respectively. The resultant plasmids were named N1, N2, N3 and N4, respectively. The primers used are shown in Table 1.

In order to construct chimeric promoters containing different deleted fragments of the *PNZIP* promoter and 90-bp CaMV 35S promoter, we first inserted CaMV 35S promoter into the HindIII/BamHI site of the pUC118 vector. Next, the regions from –1415 to –115, –1415 to –375, –1415 to –570 and –1415 to –1090 of the *PNZIP* promoter were cloned into the HindIII/EcoRV sites of the recombinant pUC118 vector following the method described above, respectively. Then, different recombinant plasmids were digested by HindIII and BamHI and the released fragments were cloned into the corresponding site of the pBI-101. Corresponding plasmids were designated as 90-N1, 90-N2, 90-N3 and 90-N4, respectively.

To construct chimeric promoters containing different copies of GAAATA element and the region from –133 to +1 of the *PNZIP* promoter, we first synthesized 4, 3, 2 and 1 copies of GAAATA element which contains a potential HindIII site, respectively. The sense and antisense oligonucleotide sequences were as follow: 5′-AGCT T [GGAAATAA]nA-3′, 5′-AGCTT[TTATTTCC]nA-3′ (n = 1–4). Corresponding oligonucleotides were annealed and cloned into the HindIII site of the Q5 vector. The normal orientation clones were verified by PCR and sequencing, and the corresponding vectors were named as E1, E2, E3 and E4, respectively.

Construction of expression plasmids

To construct plasmids for expressing the NtbZIP protein, the *NtbZIP* cDNA fragment was prepared by PCR from tobacco cDNA using primers YF1 and YF2. The PCR products were cut by EcoRI and SalI, and corresponding products were cloned into the EcoRI and SalI sites of the pET-30a to generate the pET-NtbZIP construct. The primer sequences were as follows: YF1, 5′-GAATTCAT GAACTTCAAGAACTTTGC-3′ (EcoRI was introduced into the 5′-end) and YF2, 5′-GTCGACCTACCAAGGTC CTGTTAGTGT-3′ (SalI was introduced into the 5′-end). To construct the effector plasmids used in the transient transactivation experiment, the *NtbZIP* cDNA was excised by BamHI and SalI from pET-NtbZIP, and then the DNA fragment containing the *NtbZIP* coding regions was cloned into a modified pBI121 vector, and the corresponding clone was named 35S-NtbZIP. In order to construct expression vector harboring the antisense *NtbZIP* gene, we inserted partially nonconserved *NtbZIP* cDNA region into the modified pBI121 vector in reverse orientation under control of the CaMV35S promoter and the nopaline synthase 3′-termination sequences. All of the above constructs were sequenced to avoid error by PCR.

Site-directed mutagenesis analysis

Base mutations were carried out using the QuikChange site-directed mutagenesis kit (Stratagene, La Jolla, CA, USA). The Q5-U-M construct (containing the mutated G-box) and the Q5-D-M construct (containing the mutated GATACT element) were obtained according to the manufacturer's manual. The primers used were as follows: Q5-U-M, 5′-CTGGCCTGTCACAGTATG CTATGTATCAGACCAATATTTAATCC-3′, GGATT AAATATTGGTCTGATACATAGCATACTGTGACA GGCCAG; Q5-D-M, 5′-GGATTTCTTTGGATGAGA TAACATTCCATCACTTTCATCCAATT-3′, 5′-AATT GGATGAAAGTGATGGAATGTTATCTCATCCAAA GAAATCC-3′. The mutant sites are underlined.

GUS activity assays and histochemical staining

One hundred to 200 mg of plant materials were ground in extraction buffer consisting of 50 mM sodium phosphate, pH 7.0, 10 mM 1, 2-diaminocyclohexane-N, N, N, N-tetraacetic acid, 0.1% Triton X-100, 10 mM 2-mercaptoethanol and 0.1% sodium lauryl sarcosine. After centrifugation for 10 min (12 000 *g*) at 4°C, the supernatant was collected. Protein concentrations were determined by the method of Bradford (19) and fluorometric GUS assays were performed as previously described (20).

GUS activity was histochemically detected according to Jefferson *et al.* (20). Hand-cut sections or the whole tissues were incubated in a solution of 1 mM X-gluc in 50 mM sodium phosphate (pH 7.0) for 1–12 h at 37°C.

Green tissues of tobacco seedlings were cleared of chlorophyll by incubation in 70% ethanol. The samples were observed and photographed with a microscope.

RT-PCR analysis of the *GUS* gene expression

Total RNA was extracted from by using the RNase Plant Mini Kit (Qiagen, Hilden, Germany) and was treated with RNase free DNase-I to remove genomic DNA. Then, the first-strand cDNA was synthesized with SuperScript^{TM} II RT (Invitrogen). The transcription products were amplified by PCR using the *GUS* gene primers (5′-GCAACT GGACAAGGCACT-3′, 5′-GCGTCGCAGAACATTAC A-3′) and histone gene primers (5′-GATTTTGTAGTTC AAGATTA-3′, 5′-AATAGAATAACTCCATAAAG-3′). The tobacco histone gene was used as a standard control in the RT-PCR reactions. The PCR experiment has been repeated at least three times.

Electrophoretic mobility shift assays (EMSAs)

The preparation for nuclear extract from tobacco leaves was carried out as previously described (21). The digoxigenin (DIG) gel shift (Roche, Mannheim, Germany) was used for protein–DNA binding assays. Synthetic oligonucleotides (only the sense strand is shown) containing four copies of the GAAATA element and four copies of the G-box of the *PNZIP* promoter for EMSA were as follows: 5′-[GGAAATAA] $_4$-3′, 5′-[GCCACGTGTC]$_4$-3′. The oligonucleotides were annealed, labeled and used in the gel-shift reactions according to the manufacturer instructions. To confirm the specific of protein–DNA complexes, the mutated tetrameric G-box (5′-[GCCGTTATTC]$_4$-3′) and GAAATA element (5′-[GTACACAA]$_4$ -3′) were used as nonspecific oligonucleotides, respectively. The mutant sites are underlined. Reactions were incubated at room temperature for 15 min and the resulting protein–DNA complexes were electrophoresed in 6% native polyacrylamide gel. After electrophoresis, the gel was transferred to a nylon membrane by electro-blotting. Nylon membranes were rinsed briefly in washing buffer, and incubated for 30 min in anti-Digoxigenin-AP (1:10 000) for 30 min. Then the membranes were equilibrated and were placed on hybridization bag and CSPD working solution was applied. Finally, the membranes were exposed to X-ray film for 40 min.

Yeast one-hybrid screening

Yeast one-hybrid screening was performed to isolate genes encoding proteins associated with the G-box of the *PNZIP* promoter. To construct the bait plasmid, three copies of the G-box of the *PNZIP* promoter were synthesized and inserted into the EcoRI/MluI site of the pHIS$_2$ vector, and the corresponding construct was named G-pHIS$_2$. Next, the appropriate concentration of 3-AT was determined. Tobacco cDNA libraries were prepared from tobacco leaves using BD SMART^{TM} cDNA synthesis. We cotransformed yeast strain Y187 with double-strand cDNA, pGADT7-Rec2 and G-pHIS$_2$, and restreaked the yeast cells on SD/-His-Leu-Trp + 15 mM 3-AT plates. The yeast screening procedure was performed according to the manufacturer's protocol (CLONTECH Matchmaker one-hybrid system).

Southern blot analysis

DNA was extracted from leaf tissue using the procedure described by Zheng *et al*. (9). Ten micrograms of genomic DNA was digested with HindIII, EcoRV and EcoRI, then separated on a 1% agarose gel, and blotted onto a Nytran membrane. The *NTbZIP* cDNA fragment was labeled by the random priming method and used as a probe. Hybridization was performed at 65°C in a solution of 0.05 M sodium phosphate (pH 7.0), 5× SSC, 5× Denhart's solution, 0.2 mg/ml sheared denatured salmon testes DNA (type III, Sigma) and 0.2% SDS. After the blot had been washed three times with 0.1× SSC, 0.1% SDS at 55, 60 and 65°C, autoradiography was performed at –80°C using a Kodak X-ray film with one intensifying screen for 2 days.

Northern blot analysis

RNA extraction and RNA gel-blot hybridization were carried out as previously described by Zheng *et al*. (9). Total RNA (20 μg/lane) was separated on 0.8% formaldehyde agarose gel and blotted onto a Nytran membrane. Blots were hybridized as described above and washed three times for 20 min at 55, 60 and 65°C with 0.2× SSC and 0.1% SDS solution, then exposed to X-ray film for 3 days.

6× His-tagged protein expression and purification

To obtain the NtbZIP protein, we transformed pET-NtbZIP into *Escherichia coli* BL21 (DE3), and recombinant proteins were expressed and purified by nickel–-nitrilotriacetic acid agarose (Ni–NTA) affinity chromatography according to the protocol (Qiagen, Hilden, Germany). Then recombinant proteins were dialyzed in 20 mM HEPES-KOH, pH 7.9, 1 mM $MgCl_2$, 50 mM KCl, 1 mM DTT, 20% glycerol and 0.02% NP-40 overnight. The recombinant proteins were concentrated and stored at –70°C.

Transactivation experiment with tobacco leaves

Transient expression analysis in tobacco leaves was performed by a minor modification of the method described by Wu *et al*. (22). To determine the effect of different copies of the GAAATA element on the activity of Q5, Plasmid 35S-LUC, which contains the luciferase reporter gene driven by the constitutive CaMV 35S promoter, was used as the internal control to normalize GUS activities of the reporter construct. One microgram of different reporter plasmids and 30 ng of internal control plasmid were co-bombarded into tobacco leaves, respectively. To detect whether NtbZIP can regulate the activity of Q5, three types of DNA constructs were used in the transient expression experiments: reporter, effector and internal control. Plasmid Q5 and Q5-U-M were used as the reporter construct; plasmid 35S-LUC was used as an internal control; plasmid 35S-NtbZIP was used as an effector; 35S-GFP or 35S-DREB1 was used as effector control. One microgram

of reporter plasmid, 1 μg of the effector plasmid or effector control and 30 ng of internal control plasmid were co-bombarded into tobacco leaves. After particle bombardment, the tobacco leaves were incubated in MS medium for 24 h, and then the total proteins were collected. LUC assays were performed using the Luciferase Assay System (Promega, Madison, WI, USA). All the GUS values were normalized to the corresponding LUC values.

Subcellular localization analysis of the NtbZIP-GFP protein

Plasmid NtbZIP-GFP was constructed to investigate the subcellular localization of NtbZIP in onion epidermal cells. A full-length *NtbZIP* ORF without the termination codon was prepared by PCR amplification using the tobacco cDNA as a template. Two *NtbZIP* gene-specific oligonucleotide primers (5′-GGATCCATgAACTTCAA gAACTTTgC-3′ and 5′-GGATCCCTACCAAGGTCCT GTTAGT-3′), both containing BamHI sites, were synthesized to amplify the *NtbZIP* gene ORF. The PCR products were digested with BamHI, and inserted into the BamHI site of the pBI121-GFP, so as to be fused in-frame to the N-terminal end of the GFP coding sequence. The PCR products described above were verified by sequencing. The plasmid NtbZIP-GFP and the control plasmid GFP were introduced separately into the onion epidermis cells by particle bombardment. The transformed cells were cultured on the MS medium at 28°C for 16 h and observed under a confocal microscope.

Determination of chlorophyll content

The chlorophyll *a* and chlorophyll *b* of tobacco leaves at the same developmental stage were extracted with 80% (V/V) acetone and measured by the method as previously described (13). Absorbance was recorded at 664 and 647 nm using a Shimadzu UV-1601 spectrophotometer (Shimadzu, Tokyo, Japan).

Determination of MgPMME and Pchlide content

Plant material (100–-200 mg) was weighed and ground in liquid nitrogen, and then the powdered sample were suspended with acetone/H_2O/37% NH_3 (80:20:1, v/v/v) under dim light at 4°C. After centrifugation for 15 min at 20 000 *g* (4°C), the supernatant was extracted three times with hexane and were kept in dark at 4°C. Then, MgPPE and Pchlide were isolated according to the methods as previously described (23,24). The pigment contents were identified and quantified by their absorption and fluorescence spectrometer (Shimadzu, Tokyo, Japan) compared with standards (Frontier Science, Logan, UT). The excitation (Ex) and emission (Em) for MgPPE and Pchlide were: MgPPE, Ex 418 nm and Em 595 nm; Pchlide, Ex 440 nm and Em 633 nm. Fluorescence intensity of the extract was normalized to the total fresh weight for each sample.

RESULTS

Sequence analysis of the *PNZIP* promoter

To characterize the regulatory mechanisms controlling transcription of the *PNZIP* gene, we isolated its 5′-upstream region. Figure 1 shows the sequence of the 5′-upstream region of the *PNZIP* gene, which extends into the 5′-untranslated region (GenBank accession no. AF373414). In order to identify the *PNZIP* gene transcription start site, 5′-RLM-RACE was performed. The amplified fragments (about 300 bp) were then cloned into the pGEM-T vector, and 15 independent clones were sequenced to determine the 5′-end of the products. Sequence analysis showed the *PNZIP* gene has only one transcription start site, and the transcription begins at an adenine residue located 122 bp upstream of the translational start site, which conforms to the principle of the transcription start site in plant genes (25). The putative TATA-box and CCAAT-box were located at position –30 and position –96, respectively, from the transcriptional start site. In the upstream region, a perfect palindromic G-box, the target-binding site of plant basic leucine zipper proteins (26), was found at –110 to –100 from the transcriptional start site. In addition, an AT-1-box-like sequence, Box-II (27,28) and a 3-AF1 binding site (29) were also observed in the *PNZIP* promoter. Also, two GATACT elements were observed to be located at positions –278 and –54, and four GAAATA elements were found at positions –1007, –943, –446 and –184 in the promoter sequence, respectively, suggesting that the two kinds of putative motifs might regulate gene expression.

The *PNZIP* promoter can specifically drive reporter gene expression in photosynthetic tissues

The spatial expression pattern of the *PNZIP* promoter was examined in transgenic tobacco plants. The full-length promoter sequence was fused with the *GUS* gene and introduced into tobacco via *Agrobacterium*-mediated transformation. The primary transformants were first verified by PCR analysis before being transferred into a greenhouse. Histochemical analysis of T_1 generation tobacco seedlings harboring the full-length promoter showed that blue staining was exclusively located in mesophyll cells, but no GUS staining was detected in roots (Figure 2A and B), which was in agreement with the expression patterns of *PNZIP* mRNA in *Pharbitis nil* seedlings (9). In addition, we also examined the GUS activity and *GUS* mRNA accumulation in different transgenic tobacco organs. The results showed that there were the highest GUS activity and mRNA level in the leaf and very low activity and transcripts in the root, fruit and flower, further confirming that the *PNZIP* promoter is tissue-specific (Figure 2C).

Identification of several positive elements in the *PNZIP* promoter by deletion analysis

To determine the specific region that is involved in *PNZIP* gene expression, a series of 5′-deletions in the *PNZIP* promoter was constructed (Figure 3A) and the corresponding

```
–1415  ACATGGGGAT GAGGCAGGGT ATGAGTGCTT TGCTTTTTTT TTTTGGTTAA AATTTTTTCC TCCAGGTTGA
                                                          Box-II
–1345  GATCCTCTCT CAGATTCGAA TTGAGTAATT GCTATGCATT ATACCTACAC TCAGCCTTTC GATGAGTTGT
–1275  CTCGGGGGTT TTGCAGTTTG ATACATGGAT TGATGGAGTA TTAAGATGTT TGTTTATGGA GATAAAAGTG
–1205  TTGAGCATTC ATATTTGTTG GATTTGGCTC ATTAGTAATG TGAAGGAGAG GACACAACTT TGTTGAAAAT
–1135  ACAACATTGT TATATATGGC ACCAAACAAT TTCTAGTATA GGGATAACCT TAGTCAAGTT AATTAGGTTG
–1065  TTGATTTAAT AACTATAAGG AATGATCAAT TGACCTTTAT GGTTTGAGCC GAGAAATACT CTTGAAGATA
                                                             GAAATA element
–995   CAAGATCAAC TACAGAAAAT TACTTGTGTC AGGCAAAATA ACTCTTTGAAA TATTATTCTA TTATTGTACC
                                        3-AF1 binding site   GAAATA element
–925   CCCAATTATT TCATTATAAA TCCCATTAAT TCTCATTATT TTATCAGTAT TCTACCATCA AACTTGACTC
–855   TCCATTATTT CACCATAAAT AGCCATTGTC CACCCTCAAA TAGTCATTAA TATAATCTAC CATGAATGTC
–785   TTGCCATGAA TACTATGTCC TCTACCACTA TAAAAAGACT CTACAACCAA CAAGGGAGGA GACCAAGCTC
–715   TATAGCTCTA CTTCAAGCTA CTCAAGTTCG AGAATAGTAT TTCTACAAAT TCTATACACT CCACAGACTC
–645   TATGAATTAT TCCTAGCTTG AGTGCAACAA CGTACATGCA TCTCCACACT TCAAAACATA GTGAAATTCA
–575   ATTTGATTCA TCTTAAAAAT GCATGTAATT TTTGTTTTTA CTTCAATTTT CACATTAAAC CCCTTGTTTC
–505   TTGTCCTATA CATATGACTC  TAGAACTAAC  ATGCGCAATA AGAGAATTAT GTGGGAAATA  AATTGTAATT
                                                             GAAATA element
–435   CCGTGAGGAA AGAATAAAGG TGAATGTTAC  AAATTTAAGA TTGCTGAGCT  AAACTGCATC AGCATTGGTG
                  GAAATA element
–365   TTCCTTCCAT TTACTCCACG ATTATCTTAT CAGCTTTGGA TTGGTAAAGA ATCTGCAACT CCAGATTCTA
–295   CCCAAGTTAA AGATACTATA CTACTACTGA  ATATTGTCAC TGTGCAATGC TATAAAGATT  TTTATTATAT
                  GATACT element                                          AT-1 like element
–225   TGTGAAGAAG ATTAAGTATA GATTGTGTAC CAATAACATT GTGAAACGTA CCATGAACAA CATCAGCCAC
–155   AAAATACACA  AAATGACCAT GTAATCAAGC  TGGCCTGTCA CAGTATGCCA CGTGTCAGAC  CAATATTTAA
                                                          G-box          CCAAT-box
–85    TCCCATGGAT TTCTTTGGAT GAGATGATAC TCCATCACTT TCATCCAATT ATATATCCTC TCCAGCACCC
                                  GATACT element              TATA-box
–15    ATAGCTTCAC AGTACACTCT ACCCAGAAAA AAAAATGGCA GCAGAAATGG CATTGGTAAG GCCCATATCG
+56    AAGTTCGGCG CCACCGCCAC TCCGCGGCTG AGCGGCAGGC GGAAACTGGC ACCCTTAAGC GTGAGAATGT
+126   CGAGCGCCAC  CACGTCGCAG GCGGCGGCGG
```

Figure 1. Nucleotide sequence of the 5′-region of the *PNZIP* gene. The *PNZIP* gene upstream region containing the promoter and short stretch of transcribed region is shown. Nucleotides are numbered on the left, with the transcriptional start site designated as +1. The arrowhead indicates the transcriptional start site determined by 5′-RLM-RACE. The putative TATA-box (TATATA), CAAT-box (CAAT) and the translation start sites are underlined. The G-box element (5′-CACGTG-3′), box-II (5′-GGTTAA-3′), 3-AF1-binding site (5′-AAATAACTCTT-3′), AT-1-like element (5′-ATTTTTATT-3′), GAAATA element and GATACT element are boxed.

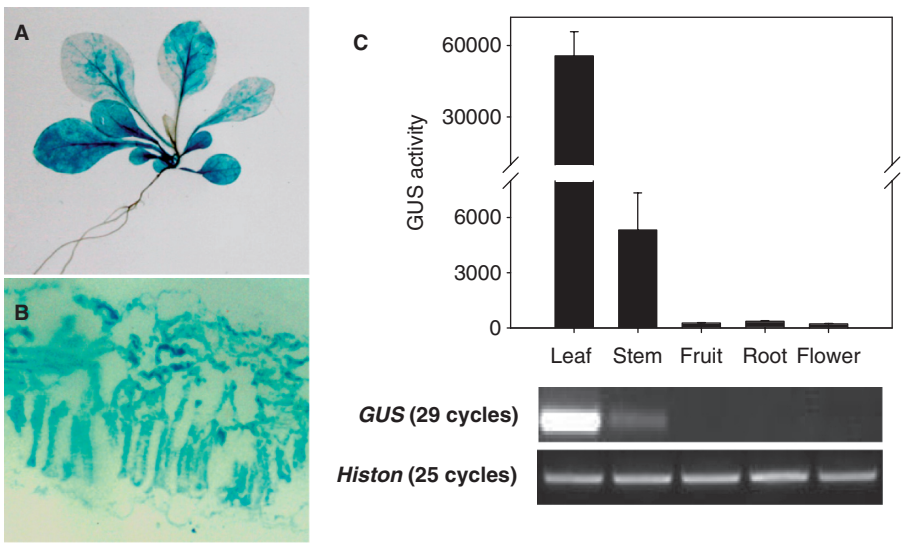

Figure 2. The activity of *PNZIP* promoter specifically expressed in photosynthetic tissue. (**A**) The tobacco seedling harboring the full-length *PNZIP* promoter was stained with GUS staining solution overnight and decolorized with 70% ethanol. (**B**) The hand-cut cross-sections of young leaves were stained with X-gluc overnight, decolorized with 70% ethanol and observed using an anatomy microscope. (**C**) GUS activities and RT-PCR analysis of the *gusA* expression in different organs of the transgenic tobaccos carrying *PNZIP::gusA*. The RT-PCR conditions were verified by a linear amplification of a histone-specific band. The picture represents one of three independent experiments that gave similar results.

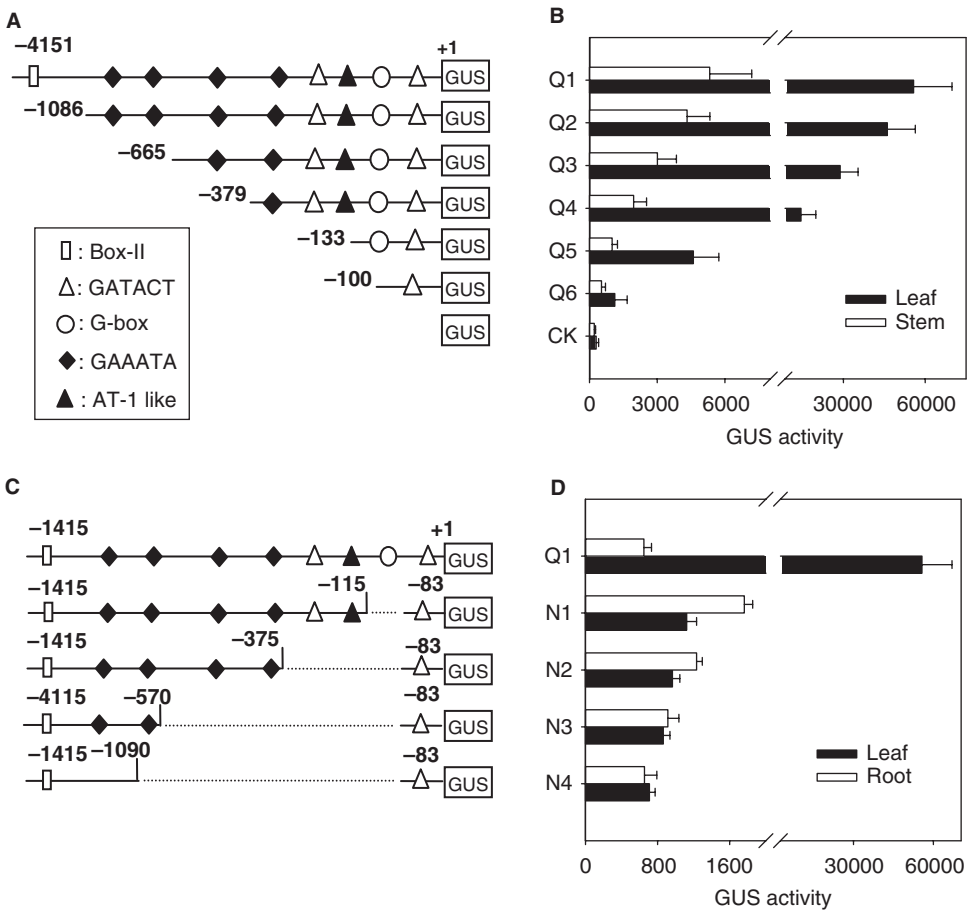

Figure 3. Deletions analysis of the *PNZIP* gene upstream region. (**A**) The different 5′-deletion constructs of the *PNZIP* promoter. (**B**) GUS activities from leaves and stems of transgenic tobacco plants for 5′-deletion construct are shown. (**C**) The different 3′- deletion constructs of the *PNZIP* promoter. (**D**) GUS activities from leaves and roots of transgenic tobacco plants for 3′-deletion construct are shown. GUS activities were measured by enzymatic conversion of 4-methylumbelliferone, which was quantified with a spectrofluorimeter. The activity is expressed as pmol 4-methylumbelliferone $min^{-1}mg^{-1}$ protein. Values are means of at least 10 independent samples, with error bars representing SE ($n \geq 10$).

transgenic tobaccos were regenerated. As shown in Figure 3B, deletion of the region between –1415 and –1086 with Box-II resulted in a 17.2% decrease of GUS activity in leaves and an 18.7% decrease in stems compared with the full-length promoter construct (Q1). Deletion of the region between –1086 and –665 containing two GAAATA elements and a 3-AF1 binding site gave rise to a 37.6% decrease of GUS activity in leaves and a 30.3% decrease in stems compared with construct Q2. However, deletion of the region between –665 and –379 harboring two GAAATA elements resulted in a 49.9% decrease of GUS activity in leaves and a 35.1% decrease in stems compared to those of construct Q3. Removal of the region between –379 and –133, which contained a GA AATA element, an AT-1 like box and a GATACT element, gave rise to a 68.1% decrease of GUS activity in leaves and 48.2% in stems compared with construct Q4. When we deleted the promoter to the position –100, GUS activities in the leaf and stem were nearly undetected, suggesting that the region between –133 and –100 carrying a G-box (5′-CACGTG-3′) is essential for high expression of the *GUS* reporter gene. Taking these results together, we suggest that the combination of the GAAATA element, AT-1 like box, GATACT element and G-box may play an important role in the regulation of the high-level expression of the *PNZIP* gene.

To further confirm these motifs play crucial roles in the full-length *PNZIP* promoter, a series of internal deletion constructs was produced (Figure 3C). In the absence of the region between –115 and –83, the activity of GUS activity in transgenic leaves decreased greatly and the expression pattern of the reporter gene was changed. A progressive decline of GUS activity in leaves was observed in constructs N2, N3 and N4, respectively, implying that some positive regulatory element is likely to be located in the region between –1415 and –115, which is in agreement with the 5′-deletion results (Figure 3D). These results strongly suggest that the G-box is an important element that is responsible for the high and specific expression of the *PNZIP* gene, and the other elements display positive characteristics in the regulation of the *PNZIP* gene.

The upstream sequence of the *PNZIP* promoter can activate the 35S minimal promoter

To further test whether the region between –1415 and –115 can activate a heterologous promoter, we fused several

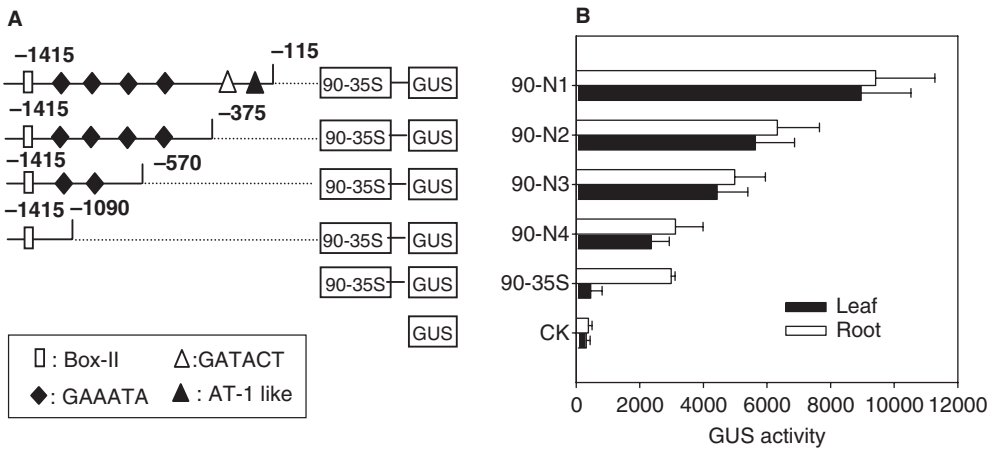

Figure 4. Activation of the 90-35S minimal promoter by the different 3′-deleted fragments of the *PNZIP* gene upstream region. (**A**) The fusion constructs of 90-35S minimal promoter and different 3′-deleted *PNZIP* upstream regions. (**B**) GUS activities from leaves and roots of transgenic tobacco plants for each construct are shown. GUS activities were measured by enzymatic conversion of 4-methylumbelliferone, which was quantified with a spectrofluorimeter. The activity is expressed as pmol 4-methylumbelliferone $min^{-1}mg^{-1}$ protein. Values are means of at least 10 independent samples, with error bars representing SE ($n > 10$). The dotted lines indicate the regions deleted in *PNZIP* promoter sequences.

3′-deleted fragments of the *PNZIP* promoter to minimal CaMV 35S promoter (–90 from the transcriptional start site, designated as 90-35S) (Figure 4A). The 90-35S promoter, containing a normal TATA-box and an ASF1 transcription factor-binding site (30), displayed a very low level of promoter activity in transgenic tobacco leaves. The transformants harboring the N1 construct, containing the sequence between –1415 and –115 of the *PNZIP* promoter, exhibited high GUS activity, not only in leaves but also in roots. This suggested that some elements in this region might interact with the as-1 sequence in the 90-35S promoter and then change its expression patterns. With the increase of the length of the promoter internal deletion, a progressive decline in the GUS activities both in roots and in leaves was observed. When we removed one AT-1 like and one GAAATA element in 90-N2, the GUS activity decreased by 37.1% in leaves and 43.5% in roots compared with 90-N1. Further deletion of two GAAATA elements resulted in a 21.4% decrease of GUS activity in leaves and a 21.3% decrease in roots compared with 90-N2. Finally, when the last two GAAATA elements were deleted, the GUS activity decreased by 40% in leaves and 37% in roots compared with 90-N3 (Figure 4B). These results indicate that the GA AATA element may be a positive element in the *PNZIP* promoter.

The GAAATA element is a positive element in the *PNZIP* promoter

To examine the presence of nuclear factors such as DNA-binding proteins bound to the GAAATA element specifically, we conducted EMSA using nuclear extracts from tobacco leaves. Synthetic oligonucleotides containing four copies of the GAAATA element were prepared and labeled with ^{32}P by polynucleotide kinase, then incubated with nuclear extract prepared from tobacco leaves. As shown in Figure 5A, one major shifted band was formed that was competed out by 200 and 500 molar excess of unlabeled GAAATA element but not with 500 molar excess of unlabeled mutated GAAATA element, suggesting that the band is due to the binding of a sequence-specific protein, which is consistent with the functional identification of the sequence as a regulatory element.

In order to further explore whether the GAAATA element is a positive element in the *PNZIP* promoter, the oligomerizations of the GAAATA element were fused directly to the Q5 construct in a single copy or in tandem repeat arrays, respectively (Figure 5B). As shown in Figure 5C, oligomerizations of the GAAATA element (one to four tandem repeats) could all enhance the GUS activity of the Q5 construct in transient expression assays with tobacco leaves, and the four copies in the tandem repeat of the GAAATA element yielded higher GUS activity than the other repeats. Together, these results indicated that the GAAATA element is a positive element in the *PNZIP* promoter.

G-box and GATACT elements are necessary for the high expression of the PNZIP gene

As shown in Figure 3A, construct Q5, containing the fragment from –133 to +1 of the *PNZIP* promoter, was still able to drive the specific expression of the reporter gene, indicating that some *cis*-elements in this region may play important roles. Sequence analysis revealed that in addition to the CAAT-box and TATA-box, a G-box and a GA TACT element that is very similar to the I-Box (5′-GATA AG-3′) (31) exist in this region (Figure 6A). To investigate whether the G-box and GATACT elements are involved in the expression of the *PNZIP* gene, we carried out mutation analysis. When the GATACT element was mutated in the region between –133 and +1, the GUS activity in roots was increased by more than 2-fold, whereas the GUS activity in leaves was reduced about 1.4-fold. When we mutated the G-box in the region between –133 and +1 of the *PNZIP* promoter, the GUS activities in leaves dropped markedly, whereas the GUS activities in roots

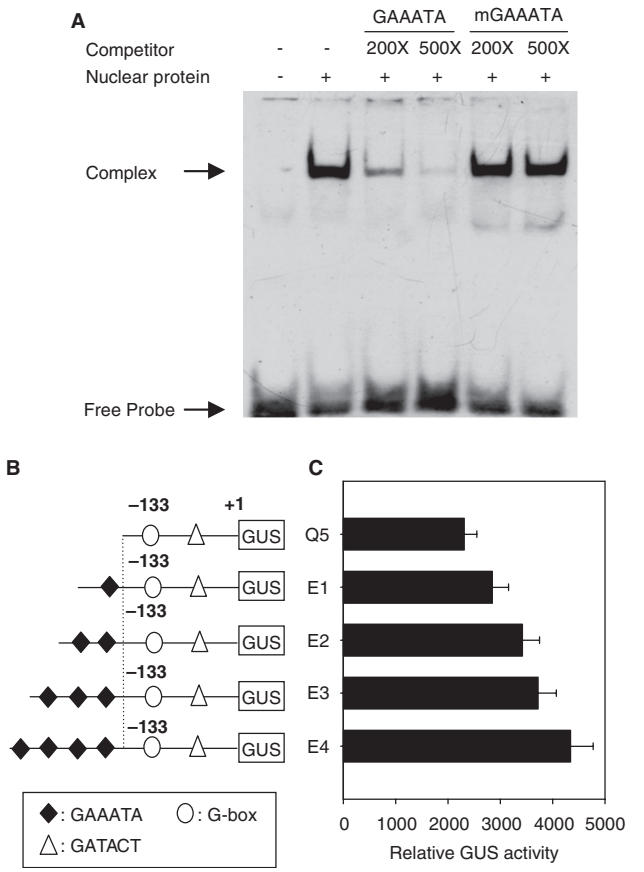

Figure 5. The GAAATA element is a positive element in the *PNZIP* promoter. (**A**) Interaction of nuclear proteins from tobacco leaves with the GAAATA element. Lane 1, control reaction without nuclear extract; lane 2, reactions containing 10 µg of nuclear extract; lanes 3 and 4, reactions containing 10 µg of nuclear extract, plus 200- and 500-fold excess of the unlabeled synthetic oligonucleotides, respectively; lanes 5 and 6, reactions containing 10 µg of nuclear extract, plus 200- and 500-fold excess of the unlabeled mutated GAAATA element (mGAAATA), respectively. The mGAAATA was used as a nonspecific competitor. (**B**) The fusion constructs of the *PNZIP* Q5 region and different copies of the GAAATA element. (**C**) The transient GUS activities in tobacco leaves following particle bombardment were measured by enzymatic conversion of 4-methylumbelliferone, which was quantified with a spectrofluorimeter. The activity is expressed as pmol 4-methylumbelliferone $min^{-1}mg^{-1}$ protein. All GUS values were normalized to the LUC values. Values are the means of at least six independently bombarded samples, with error bars representing SE ($n \geq 6$).

increased significantly (Figure 6B). These results demonstrated that the G-box and GATACT elements are necessary and sufficient to confer the high expression of the *PNZIP* gene.

Isolation of cDNAs encoding DNA-binding proteins that interact with the *PNZIP* promoter

It has been proposed that the G-box element serves as a binding site for the transcriptional activator, and multiple proteins with distinct G-box DNA-binding specificities have been identified in several plant species (32,33). In order to isolate new transcription factors that can interact

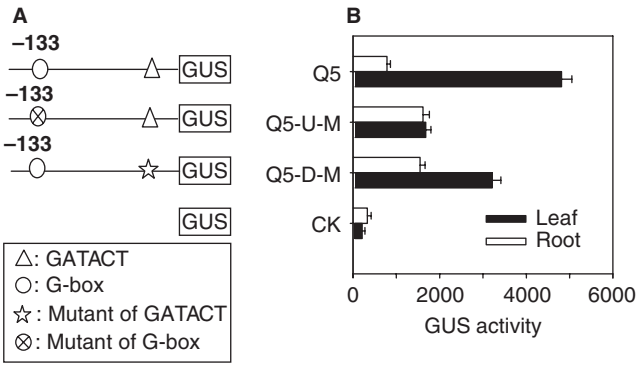

Figure 6. G-box and GATACT elements are necessary for the high expression of the GUS reporter gene. (**A**) The constructs of wild-type and mutated *PNZIP* Q5. (**B**) GUS activities from leaves and roots of transgenic tobacco plants for each construct were measured by enzymatic conversion of 4-methylumbelliferone, which was quantified with a spectrofluorimeter. The activity is expressed as pmol 4-methylumbelliferone $min^{-1}mg^{-1}$ protein. Values are means of at least 10 independent samples, with error bars representing SE ($n \geq 10$).

with the G-box of the *PNZIP* promoter, we carried out yeast one-hybrid screening of a tobacco leaf cDNA library using three tandem copies of G-box in the *PNZIP* promoter as the target binding sequence. 220 positive colonies from selective-medium plates (SD/-His-Leu-Trp + 15 mM 3-AT) were obtained. When the concentration of 3-AT was increased from 15 to 80 mM, one clone growing normally on SD/-His-Leu-Trp + 60 mM 3-AT plates was isolated. Sequence analysis indicated that this cDNA encodes a protein with a bZIP DNA-binding domain. Using a BD SMART™ RACE kit, the full-length cDNA sequence of the bZIP protein was obtained, and the gene was designated as *NtbZIP*.

The *NtbZIP* cDNA contains a single ORF of 400 amino acid residues and encodes a putative protein with a predicted molecular mass of 44 kDa, which has been registered in GenBank with the accession number DQ073639. Amino acid sequence analysis indicated that NtbZIP contains a canonical bZIP motif including three heptad repeats of leucine in the zipper domain and a 23 amino acid sequence rich in basic amino acids of arginine and lysine immediately downstream. Comparison of the amino acid sequences of NtbZIP with homologs from other plant species revealed that four regions containing putative phosphorylation sites are highly conserved (Figure 7A). Moreover, Southern blot analysis suggested that there are two copies of *NtbZIP* in the tobacco genome (Figure 7B).

NtbZIP protein is exclusively localized to the nucleus and expressed in all plant organs

To confirm the possibility that NtbZIP protein acts as a nuclear factor, we examined the subcellular localization of the NtbZIP protein. The ORF of NtbZIP was fused to the upstream region of the green fluorescent protein gene, which acted as a fluorescent marker. 35S-GFP or 35S-NtbZIP-GFP plasmids were introduced into the onion epidermal cells by particle bombardment, and the GFP fluorescence was visualized using fluorescence microscopy.

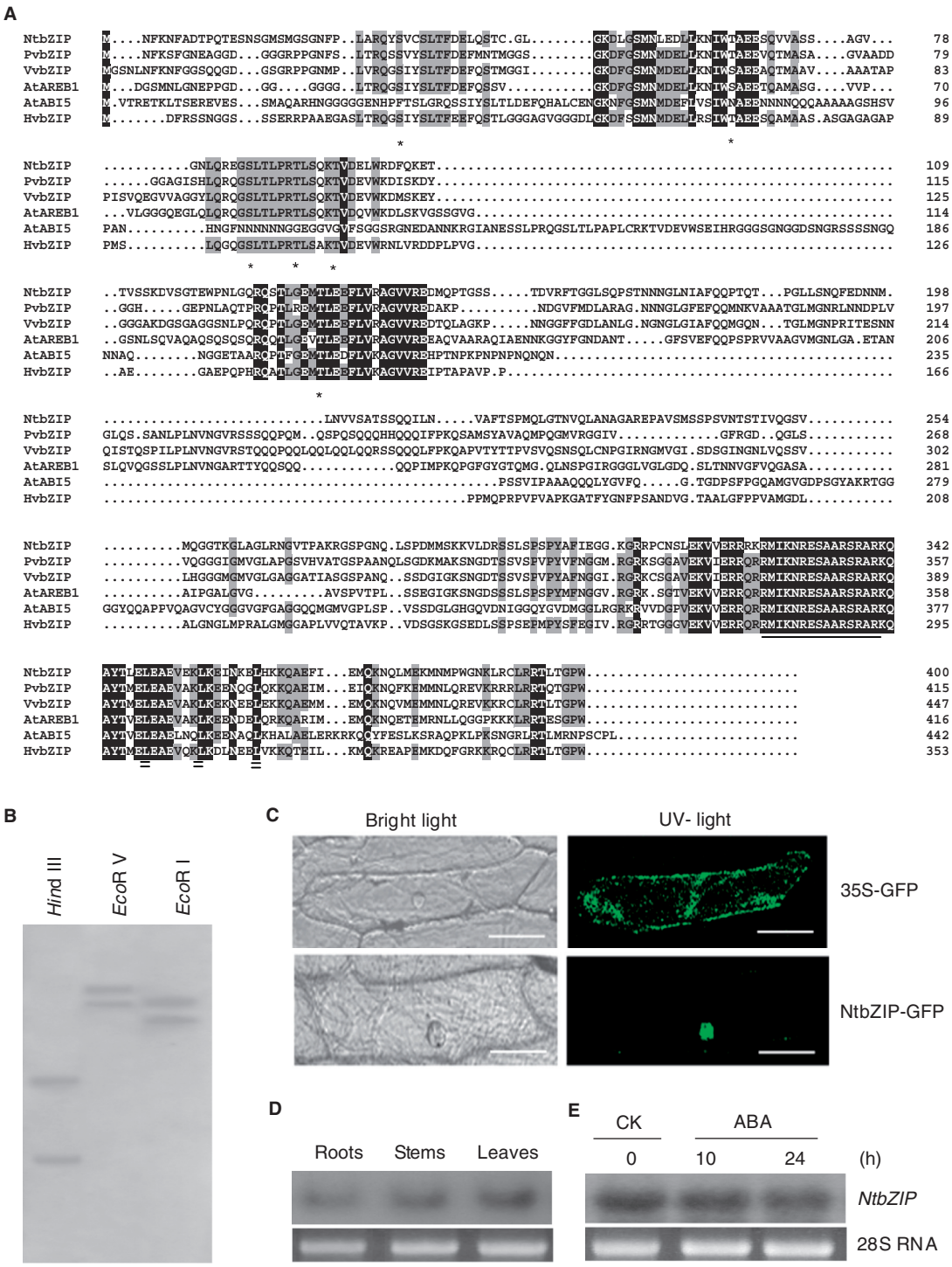

Figure 7. Characterization of NtbZIP. (**A**) Comparison of the amino acids sequences of NtbZIP with five bZIP proteins in other plants. The basic region is underlined, and the leucine residues in bZIP are double underlined. Asterisks denote potential phosphorylation sites, and sources of these bZIPs are as follows: NtbZIP, *Nicotiana tabacum*; PvbZIP, *Phaseolus vulgaris*; VvbZIP, *Vitis vinifera*; HvbZIP, *Hordeum vulgare subsp.Vulgare*; AtAREB1, *Arabidopsis thaliana*; AtABI5, *Arabidopsis thaliana*. (**B**) Southern blot analysis of tobacco genomic DNA. Ten micrograms of genomic DNA was digested with different enzymes: EcoRI, HindIII and EcoRV. The DNA gel blot was hybridized with ^{32}P-labeled *NtbZIP* cDNA probe. (**C**) Subcellular localization of the NtbZIP protein. The fusion construct for NtbZIP-GFP and the GFP control plasmid were introduced into onion epidermis cells by biolistic bombardment transformation. The transformed cells were cultured on the MS medium at 28°C for 24 h and observed under a microscope. Bar, 50 μm. (**D**) *NtbZIP* mRNA accumulation in different organs. (**E**) *NtbZIP* mRNA levels in leaves of tobacco seedlings treated with ABA for indicated hours. About 20 μg of total RNA was analyzed by RNA gel blotting. The blot was hybridized with 3′-noncoding regions of *NtbZIP* cDNA. The ethidium bromide-stained 28S RNA is shown as a loading control.

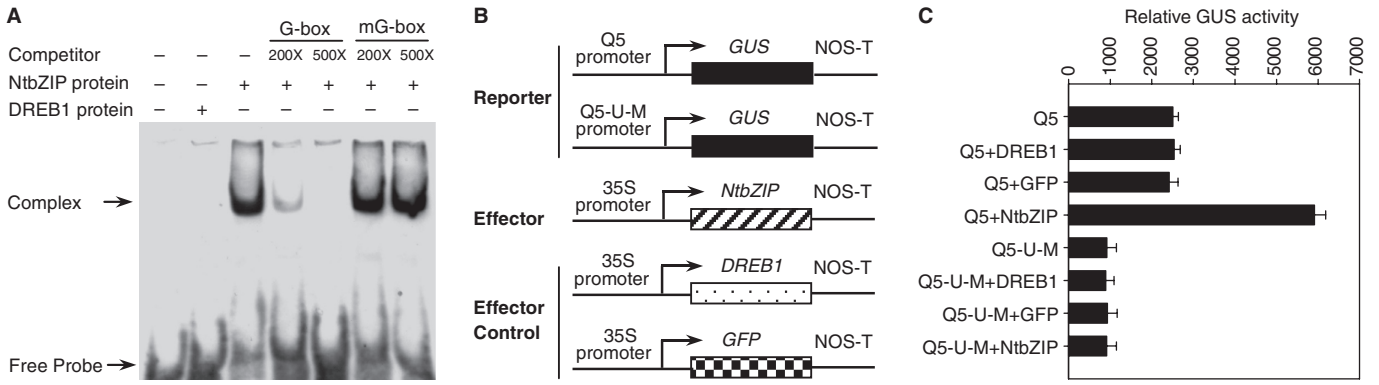

Figure 8. NtbZIP protein can regulate the activity of the *NTZIP* promoter. (**A**) Interaction of NtbZIP protein with the G-box of the *PNZIP* promoter. Lane 1, control reaction without NtbZIP protein; lane 2, control reaction with *in vitro* expressed *Arabidopsis* DREB1 protein; lane 3, reaction with 100 ng of NtbZIP protein; lanes 4 and 5, reactions with 100 ng of NtbZIP protein, plus 200- and 500-fold excess of the unlabelled synthetic oligonucleotide, respectively; lanes 6 and 7, reactions with 100 ng of NtbZIP protein, plus 200- and 500-fold excess of the unlabeled mutated G-box (mG-box), respectively. The mG-box was used as a nonspecific competitor. (**B**) Schemes of the reporter and effector constructs used in the transient expression assays. (**C**) The reporter construct Q5 or Q5-U-M was co-bombarded into tobacco leaves along with effector construct (35S-NtbZIP) or effector controls. GUS activities were measured by enzymatic conversion of 4-methylumbelliferone, which was quantified with a spectrofluorimeter. The activity is expressed as pmol 4-methylumbelliferone $min^{-1} mg^{-1}$ protein. All GUS values were normalized to the LUC values. Values are the means of at least six independently bombarded samples, with error bars representing SE ($n \geq 6$).

In the control, GFP fluorescence was distributed throughout the cells, while GFP-NtbZIP fusion proteins were localized exclusively in the nuclei (Figure 7C), indicating that NtbZIP is a nuclear protein.

In order to determine the expression pattern of *NtbZIP* in tobacco seedlings, northern blot hybridization analysis was performed. The results showed that *NtbZIP* transcripts were detected in the leaf, stem and root (Figure 7D). Based on previous studies that reported that several genes for plant bZIP protein are inducible by ABA (34,35), we then detected the expression pattern of *NtbZIP* in tobacco seedlings treated with 100 μM ABA for 48 h. The results indicated that the levels of *NtbZIP* mRNA were not changed after application of exogenous ABA (Figure 7E), suggesting that NtbZIP may represent a novel subclass of plant bZIP protein that is not regulated by ABA.

NtbZIP protein can transactivate the *PNZIP* promoter in tobacco leaves

To investigate the mechanism underlying the transcriptional regulation of the *PNZIP* promoter activity by NtbZIP, we determined the NtbZIP DNA-binding characteristics by EMSA by using the G-box of the *PNZIP* promoter. Recombinant NtbZIP protein was expressed in *E. coli* as a His-tag fusion protein and purified by affinity chromatography. As shown in Figure 8A, the Arabidopsis DREB1 (36) protein alone did not show any binding activity, whereas a strong mobility DNA–protein complex was formed with the recombinant NtbZIP protein. This complex was efficiently competed out with 200 and 500 molar excess of unlabeled G-box but not with 500 molar excess of mutated G-box. Taken together, these results suggest that NtbZIP specifically interacts with G-box of the *PNZIP* promoter *in vitro*.

To determine whether NtbZIP protein could transactivate the reporter gene transcription in plant cells, different effector constructs along with the Q5 or Q5-U-M construct were co-bombarded into tobacco leaves, respectively (Figure 8B). The results showed that expression of the *NtbZIP* cDNA in tobacco leaves could obviously activated Q5 activity, whereas the *GFP* and Arabidopsis *DREB1* could not. Besides, NtbZIP did not affect Q5-U-M activity, indicating that NtbZIP protein could act as a transcription factor to regulate the *PNZIP* promoter activity (Figure 8C).

NtbZIP protein does regulate the expression of the *NTZIP* gene

Our previous study demonstrated that *Pharbitis PNZIP* shares high similarity of amino acid sequence and expression pattern with tobacco *NTZIP* (13). And based on the fact that the NtbZIP protein specifically regulates the activity of the *PNZIP* promoter, we speculate that *NtbZIP* may regulate the expression of *NTZIP* in tobacco plants. To test this, construct of the CaMV35S promoter and *NtbZIP* coding region in normal orientation and construct of the CaMV35S promoter and partially nonconserved *NtbZIP* coding region in reverse orientation were regenerated and transformed into tobaccos. At least 30 transgenic lines harboring the sense *NtbZIP* cDNA were obtained and were phenotypically similar to wild-type plants. However, only three lines carrying the antisense *NtbZIP* cDNA were obtained because many transgenic buds died during growth and development, and the seedlings displayed yellow leaves and low growth rates compared with the wild-type tobacco plants. Northern blots showed that *NTZIP* mRNA accumulation was reduced in transgenic lines carrying antisense *NtbZIP* and increased in the lines harboring the sense *NtbZIP* compared with wild-type plants (Figure 9). Moreover, the level of chlorophyll in the leaves of the antisense plants was half that of wild-type controls or sense plants, whereas the ratio of Chl *a*/*b* did not change much (Table 2), indicating that a chlorophyll deficiency occurred in the antisense lines.

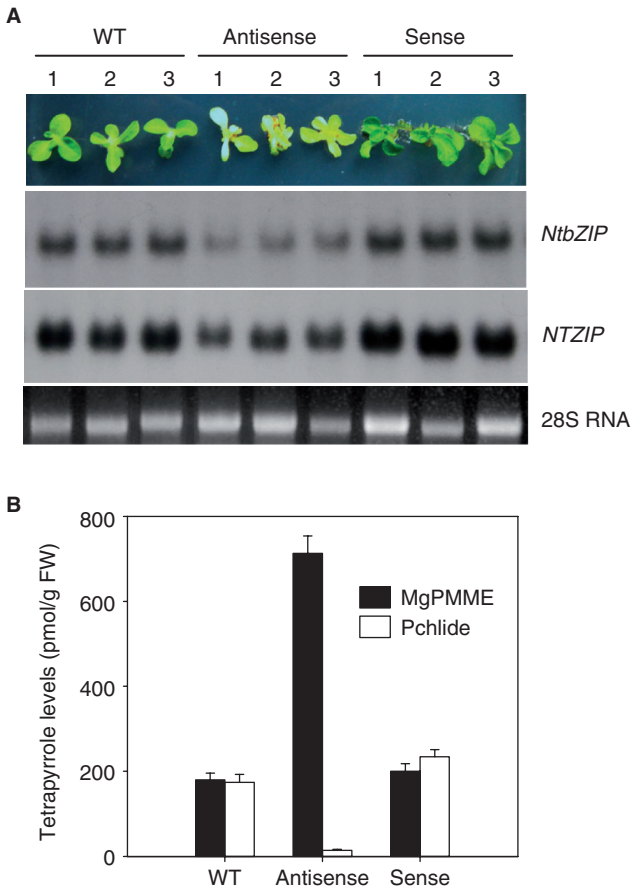

Figure 9. Phenotype and *NTZIP* mRNA accumulation and tetrapyrrole level in transgenic tobacco lines carrying sense and antisense NtbZIP. (**A**) Phenotype and *NTZIP* mRNA accumulation. Total RNA was extracted from tobacco seedlings. About 20 μg of total RNA was analyzed by RNA gel blotting. The blot was hybridized with 3′-noncoding regions of NtbZIP cDNA and NTZIP cDNA as probes. The ethidium bromide-stained 28S RNA is shown as a loading control. (**B**) MgPMME and Pchlide levels. MgPMME and Pchlide were extracted from leaves and quantified by spectrofluorometry. The values are the means of three independent experiments, with error bars representing SE ($n \geq 3$). Pigment contents of the extract were normalized to the total fresh weight (g FW) for each sample and the tetrapyrrole level is expressed as pmol/g FW. WT: wild-type tobaccos; antisense: independent transgenic tobacco lines harboring the antisense *NtbZIP*; sense: independent transgenic tobacco lines harboring the sense *NtbZIP*.

Table 2. Changes in leaf chlorophyll content and Chl *a*/*b* ratio in wild-type, transgenic sense and antisense tobaccos

Name	Chl (mg g^{-1} FW)	Chl a/b
Wild-type plants	2.01 ± 0.13 (100)	3.13 ± 0.26
Sense lines	2.05 ± 0.12 (102)	3.11 ± 0.27
Antisense lines	0.79 ± 0.03 (39)	2.89 ± 0.18

Values indicated with different letters were significantly different at $P = 0.05$. Each value is an average of five measurements. Numbers in parenthesis are percentages.

To determine whether the transition from MgPMME to Pchlide was aberrant in antisense plants, we measured the content of MgPMME and Phclide from wild-type, sense and antisense transgenic plants, respectively. The results showed that antisense plants could accumulate higher level of MgPMME compared with the sense and wild-type plants (Figure 9B). Taken together, all these results suggest that NtbZIP plays an important role in regulation of the biosynthesis of chlorophyll.

DISCUSSION

We have previously shown that *PNZIP* is a photosynthetically specific expression gene, and recent biochemical and genetic data have revealed that *PNZIP* and its homologs encode a special cyclase that plays an important role in chlorophyll biosynthesis in higher plants (9,12–14). Identification of *cis*-regulatory elements and their binding proteins constitutes an important part of understanding gene function and regulation. To our knowledge, however, there have been no reports about the regulatory mechanism of these genes to date. In this study, we employed tobacco, an efficient plant transformation system, to analyzed the *cis*-elements and *trans*-factors involved in *PNZIP* gene expression. Our results revealed that the presence of the G-box and GATACT element was necessary and sufficient for the specific expression of the *GUS* reporter gene (Figure 6). In addition, we have also identified a positive element, GAAATA, as well as a novel NtbZIP transcription factor (Figure 5). Transgenic tobacco analysis showed that overexpression of the sense *NtbZIP* cDNA results in an obvious increase of the *NTZIP* expression, whereas suppression of *NtbZIP* leads to a decline of the *NTZIP* expression, which further causes yellow leaves and lower survival rates of tobacco plants (Figure 9), suggesting its important role in plant growth and development via specifically regulating chlorophyll biosynthesis.

The G-box element is a highly conserved DNA sequence that has been identified in the 5′ -upstream region of plant genes exhibiting regulation by a variety of environmental signals and physiological cues (37), and a family of plant basic leucine zipper proteins has been identified that interacts with G-box elements (26). A perfect palindromic G-box element, 5′-CACGTG-3′, was found in the 5′ upstream region of the *PNZIP* promoter (Figure 1). We therefore presumed that the G-box element is a target for transcription factor controlling the high and specific expression of the *PNZIP* gene. Stable expression assays with 5′- and 3′-deletion constructs, however, indicated that the G-box is not a single dominant regulatory element responsible for the specific expression of the of the *PNZIP* gene and the downstream sequence is also necessary for high activity of the *PNZIP* promoter, suggesting that the G-box requires additional regulatory sequences for its functions.

The GATACT element, located in the downstream of the G-box and sharing the same core nucleotide with the GATA motif (Figure 1), is similar to the I-box (5′-GATAAG-3′) (31), and we thus designated it as an I-box-like element. The I-box element is less well characterized than the G-box and seems to be involved in light-regulated and circadian clock-related gene expression of photosynthesis genes (38,39). However, GATA motifs are found in

some light-regulated gene promoters (38,40). The GATA motif in the *Lhcb* promoter has been reported to be involved in activating transcription in green tissue (39). In tomato, the GATA motif has been shown to be an activating *cis*-element in the *RBCS* promoter (40). Our further deletion analysis indicated that the GATACT element is crucial to the specific expression of the *PNZIP* promoter, and G-box and GATACT elements act cooperatively to confer the specific expression of the *GUS* reporter gene (Figure 8). Therefore, combinatorial interaction of several elements plays an important role in the regulation of the *PNZIP* gene.

Our deletion analysis showed that several distinct elements contribute to the high expression of the *PNZIP* promoter. The presence of multiple positive elements is a common feature of many eukaryotic promoters (41). The Box-II element, 5′-GGTTAA-3′, has been found in promoters of genes involved in photosynthesis (28), flavonoid biosynthesis (42) and photomorphogenesis (43). Deletion of the Box-II element within the region between –1415 and –1086 of the *PNZIP* promoter resulted in a 17% reduction of GUS activity in leaves (Figure 3). One common feature in all Box-II elements is a core sequence of 4 or 5 nt, which consists of a 'T' or 'A' preceded by one or two 'G' nucleotides at the 5′ end and is currently defined as 5′-G-Pu-(T/A)-TA-(T/A)-3′ (44). It has been proposed that the high degeneracy of the Box-II element partly explains its diverse functions. In the bean chalcone synthase *CHS15* promoter, it acts as a negative element (42), but in the rice phytochrome *PHYA*, it acts as an activating element (45). Thus, the Box-II element can exert positive or negative effects depending upon its position and association with other promoter elements. Interestingly, we also identified a novel positive element (GAAATA) in the *PNZIP* promoter. Although the GAA ATA element is similar to the Box-II element, the key nucleotides are quite different. Competition experiments in the EMSA clearly showed that the GAAATA element could be specifically bound to the nuclear extracts of the tobacco. Gain-of-function experiments revealed that the GAAATA element is a positive element (Figure 5), implying that it may be a protein-binding site for regulating the transcription activity of the *PNZIP* gene.

The bZIP DNA-binding proteins have been intensively studied because they play roles in the regulation of several biological processes such as seed-storage gene expression (46,47), photomorphogenesis (48), leaf development (49), flower development (50), abscisic acid (ABA) response (51) and gibberellin biosynthesis (52). Plant bZIP proteins share a highly conserved basic domain and bind to target sequences usually containing the ACGT element (26,53). To date, many plant bZIP proteins have been characterized, such as GBF1-4, CPRF1-3, HBP-1a, TRAB1, ABI5, AREBs, GA2.2 and PPI1 (53–59). In the present study, we isolated NtbZIP, which binds to the G-box of the *PNZIP* promoter and only shares a limited amino acid homology to ABI5 and ABER1. Interestingly, ABI5 and ABER1 are induced by exogenous ABA and mediate ABA-regulated gene expression in seeds and vegetative tissues in both monocots and dicots (60,61). However, northern blot analysis showed that *NtbZIP* is not induced by exogenous ABA, suggesting its special function compared to ABI5 and ABER1 (Figure 7E). In addition, transgenic tobaccos analysis demonstrated that constitutively expressing antisense *NtbZIP* gene resulted in a lower *NTZIP* mRNA accumulation and higher MgPMME level (Figures 9 and 10). We thus suggest that NtbZIP, as a novel bZIP transcription factor, might regulate chlorophyll biosynthesis by controlling the NTZIP gene expression.

In conclusion, our findings suggest that a combination of several regulatory elements determines the high and tissue-specific expression of the *PNZIP* gene, and also show that the NtbZIP protein can regulate chlorophyll biosynthesis to affect plant growth and development by controlling the *NTZIP* gene expression. Future studies will focus on further identification of the binding protein that specifically binds to the GAAATA element, which might provide new insights into the mechanisms for governing *PNZIP* gene expression.

FUNDING

The National Basic Research Program (grant no. 2006CB1001006), the Program for Changjiang Scholars and Innovative Research Team in University (grant no. IRT0635); and the National Natural Science Foundation (grant no. 30570144) in China. Funding for open access charge: National Natural Science Foundation (grant no. 30570144 to C.C.Z.).

Conflict of interest statement. None declared.

REFERENCES

1. Wettstein,D.Von., Gough,S. and Kannangara,C.G. (1995) Chlorophyll Biosynthesis. *Plant Cell*, **7**, 1039–1057.
2. Liu,H.H. and Zheng,C.C. (2008) MPEC: an important gene during the chlorophyll biosynthesis pathway in photosynthetic organisms. *Photosynthetica*, **46**, 321–328.
3. Block,M.A., Tewari,A.K., Albrieux,C., Maréchal,E. and Joyard,J. (2002) The plant S-adenosyl-L-methionine:Mg-protoporphyrin IX methyltransferase is located in both envelope and thylakoid chloroplast membranes. *Eur. J. Biochem.*, **269**, 240–248.
4. Xiong,J. and Baner,C.E. (2002) Complex evolution of photosynthesis. *Rev. Plant Physiol. Plant Mol. Biol.*, **53**, 503–521.
5. Wettsteln,D., Kahn,A., Nielsen,O.F. and Gough,S.P. (1974) Genetic regulation of chlorophyll synthesis analyzed with mutants of barley. *Science*, **184**, 800–802.
6. Koncz,C., Mayerhofer,R., Koncz-Kalman,Z., Nawrath,C., Relss,B., Redel,G.P. and Schell,J. (1990) Isolation of a gene encoding a nove1 chloroplast protein by T-DNA tagging in *Arabidopsis thaliana*. *EMBO J.*, **9**, 1137–1146.
7. Hudson,A., Carpenter,R., Doyle,S. and Coen,E.S. (1993) Olive: a key gene required for chlorophyll biosynthesis in *Antirrhinum majus*. *EMBO J.*, **12**, 3711–3719.
8. Rissler,H.M., Collakova,E., DellaPenna,D., Whelan,J. and Pogson,B.J. (2002) Chlorophyll biosynthesis. Expression of a second *Chl I* gene of magnesium chelatase in *Arabidopsis* supports only limited chlorophyll synthesis. *Plant Physiol.*, **128**, 770–779.
9. Zheng,C.C., Porat,R. and O'Neill,S.D. (1998) PNZIP is a novel mesophyll specific cDNA that is regulated by phytochrome and a circadian rhythm and encodes a protein with a leucine zipper motif. *Plant Physiol.*, **116**, 27–35.
10. Moseley,J., Quinn,J., Eriksson,M. and Merchant,S. (2000) The Crd1 gene encodes a putative di-iron enzyme required for photosystem I accumulation in copper deficiency and hypoxia in chlamydomonas reinhardtii. *EMBO J.*, **19**, 2139–2151.
11. Pinta,V., Picaud,M., Reiss-Husson,F. and Astier,C. (2002) Rubrivivax gelatinosus AcsF (previously orf358) codes for a

conserved, putative binuclear iron cluster containing protein involved in aerobic oxidative cyclization of Mg-protoporphyrin IX monomethylester. *J. Bacteriol.*, **184**, 746–753.
12. Tottey,S., Block,M.A., Allen,M., Westergren,T., Albrieux,C., Scheller,H.V., Merchant,S. and Jensen,P.E. (2003) Arabidopsis CHL27, located in both envelope and thylakoid membranes, is required for the synthesis of protochlorophyllide. *Proc. Natl Acad. Sci. USA*, **100**, 16119–16124.
13. Liu,N., Yang,Y.T., Liu,H.H., Yang,G.D., Zhang,N.H. and Zheng,C.C. (2004) *NTZIP* antisense plants show reduced chlorophyll levels. *Plant Physiol. Biochem.*, **42**, 321–327.
14. Rzeznicka,K., Walker,C.J., Westergren,T., Kannangara,C.G., Wettstein,D.V., Merchant,S., Gough,S.P. and Hansson,M. (2005) *Xantha-l* encodes a membrane subunit of the aerobic Mg-protoporphyrin IX monomethyl ester cyclase involved in chlorophyll biosynthesis. *Proc. Natl Acad. Sci. USA*, **102**, 5886–5891.
15. Gowik,U., Burscheidt,J., Akyildiz,M., Schlue,U., Koczor,M., Streubel,M. and Westhoff,P. (2004) *Cis*-regulatory elements for mesophyll-specific gene expression in the C_4 plant *Flaveria trinervia*, the promoter of the C_4 phosphoenolpyruvate carboxylase gene. *Plant Cell*, **16**, 1077–1090.
16. Qi,X., Zhang,Y. and Chai,T. (2007) Characterization of a novel plant promoter specifically induced by heavy metal and identification of the promoter regions conferring heavy metal responsiveness. *Plant Physiol.*, **143**, 50–59.
17. Hoyerová,K., Perry,P., Hand,P., Lankova,M., Kocabek,T., May,S., Kottova,J., Paces,J., Napier,R. and Zazimalova,E. (2008) Functional characterization of PaLAX1, a putative auxin permease, in heterologous plant systems. *Plant Physiol.*, **146**, 1128–1141.
18. Horsch,R., Fry,J.E., Hoffmann,N., Eichholtz,D., Rogers,S. and Fraley,R.T. (1985) A simple and general method for transferring genes into plants. *Science*, **227**, 1229–1231.
19. Bradford,M.M. (1976) A rapid sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.*, **72**, 248–254.
20. Jefferson,R.A., Kavanagh,T.A. and Bevan,M.W. (1987) GUS fusions: beta-glucouronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.*, **6**, 3901–3907.
21. Foster,R., Gasch,A., Kay,S. and Chua,N.-H. (1992) *Methods in Arabidopsis Research*. World Scientific Publishing Co. Pte. Ltd., Singapore, pp. 378–392.
22. Wu,K., Tian,L., Malik,K., Brown,D. and Miki,B. (2000) Functional analysis of HD2 histone deacetyase homologues in Arabidopsis thaliana. *Plant J.*, **22**, 19–27.
23. Gough,S.P., Petersen,B.O. and Duus,J.O. (2000) Anaerobic chlorophyll isocyclic ring formation in *Rhodobacter capsulatus* requires a cobalamin cofactor. *Proc. Natl Acad. Sci. USA*, **97**, 6908–6913.
24. Walker,C.J., Castelfranco,P.A. and Whyte,B.J. (1991) Synthesis of divinyl protochlorophyllide: enzymological properties of the Mg-protoporphyrin IX monomethyl ester oxidative cyclase system. *Biochem. J.*, **276**, 691–697.
25. Joshi,C.P. (1987) An inspection of the domain between putative TATA box and translation start site in 79 plant genes. *Nucleic Acids Res.*, **15**, 6643–6653.
26. Foster,R, Izawa,T. and Chua,N.-H. (1994) Plant bZIP proteins gather at ACGT elements. *FASEB J.*, **8**, 192–200.
27. Datta,N. and Cashmore,A.R. (1989) Binding of a pea nuclear protein to promoters of certain photoregulated genes is modulated by phosphorylation. *Plant Cell*, **1**, 1069–1077.
28. Schindler,U. and Cashmore,A.R. (1990) Photoregulated gene expression may involve ubiquitous DNA binding proteins. *EMBO J.*, **9**, 3415–3427.
29. Lam,E., Kano-Murakami,Y., Gilmartin,P., Niner,B. and Chua,N.-H. (1990) ASF-2: A factor that binds to the Cauliflower Mosaic Virus 35S promoter and a conserved GATA motif in Cab promoters. *Plant Cell*, **2**, 857–866.
30. Fang,R.X., Nagy,F., Sivasubramaniam,S. and Chua,N.-H. (1989) Multiple cis regulatory elements for maximal expression of cauliflower mosaic virus 35S promoter in transgenic plants. *Plant Cell*, **1**, 141–150.
31. Meier,I., Callan,K.L., Fleming,A.J. and Gruissem,W. (1995) Organ-specific differential regulation of a promoter subfamily for the ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit genes in tomato. *Plant Physiol.*, **17**, 1105–1118.
32. Giuliano,G., Pichersky,E., Malik,V.S., Timko,M.P., Scolnik,P.A. and Cashmore,A.R. (1988) An evolutionary conserved protein binding sequence upstream of a plant light-regulated gene. *Proc. Natl Acad. Sci. USA*, **85**, 7089–7093.
33. Staiger,D., Kaulen,H. and Schell,J. (1989) A CACGTG motif of the Antirrhinum majus chalcone synthase promoter is recognized by an evolutionary conserved nuclear protein. *Proc. Natl Acad. Sci. USA*, **86**, 6930–6934.
34. Bonetta,D. and McCourt,P. (1998) Genetic analysis of ABA signal transduction pathways. *Trends Plant Sci.*, **3**, 231–235.
35. Grill,E. and Himmelbach,A. (1998) ABA signal transduction. *Curr. Opin. Plant Biol.*, **1**, 412–418.
36. Liu,Q., Kasuga,M., Sakuma,Y., Abe,H., Miura,S., Yamaguchi-Shinozaki,K. and Shinozaki,K. (1998) Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in Arabidopsis. *Plant Cell*, **10**, 1391–1406.
37. Mary,E., Randy,F.W. and Chua,N.-H. (1992) Sequences flanking the hexameric G-Box core CACGTG affect the specificity of protein binding. *Plant Cell*, **4**, 485–496.
38. Terzaghi,W.B. and Cashmore,A.R. (1995) Light-regulated transcription. *Annu. Plant Mol. Biol.*, **46**, 445–474.
39. Anderson,S.L. and Kay,S.A. (1995) Functional dissection of circadian clock and phytochrome related transcription of the Arabidposis CAB2 gene. *Proc. Natl Acad. Sci. USA*, **92**, 1500–1504.
40. Manzara,T., Carrasco,P. and Gruissem,W. (1991) Developmental and organ-specific changes in promoter DNA–protein interactions in the tomato rbcS gene family. *Plant Cell*, **3**, 1305–1316.
41. Bolouri,H. and Davidson,E.H. (2002) Modeling DNA sequences-based cis-regulatory gene networks. *Dev. Biol.*, **246**, 2–13.
42. Lawton,M.A., Dean,S.M., Dron,M., Kooter,J.M., Kragh,K.M., Harrison,M.J., Tanguay,L., Dixon,R.A. and Lamb,C.J. (1991) *Plant Mol. Biol.*, **16**, 235–249.
43. Kay,S.A., Keith,B., Shinozaki,K., Chye,M.L. and Chua,N.H. (1989) The rice phytochrome gene: structure, autoregulated expression, and binding of GT-1 to a conserved site in the 5′ upstream region. *Plant Cell*, **1**, 351–360.
44. Zhou,D.X. (1999) Regulatory mechanism of plant gene transcription by GT- elements and GT-factors. *Trends Plant SCI.*, **4**, 210–214.
45. Dehesh,K., Hung,H., Tepperman,J.M. and Quail,P.H. (1992) GT-2: a transcription factor with twin autonomous DNA-binding domains of closely related but different target sequence specificity. *EMBO J.*, **11**, 4131–4144.
46. Albani,D., Hammond-Kosack,M.C., Smith,C., Conlan,S., Colot,V., Holdsworth,M. and Bevan,M.W. (1997) The wheat transcriptional activator SPA: A seed-specific bZIP protein that recognizes the GCN4-like motif in the bifactorial endosperm box of prolamin genes. *Plant Cell*, **9**, 171–184.
47. Chern,M.S., Eiben,H.G. and Bustos,M.M. (1996) The developmentally regulated bZIP factor ROM1 modulates transcription from lectin and storage protein genes in bean embryos. *Plant J.*, **10**, 135–148.
48. Oyama,T., Shimura,Y. and Okada,K. (1997) The *Arabidopsis HY5* gene encodes a bZIP protein that regulates stimulus-induced development of root and hypocotyl. *Genes Dev.*, **11**, 2983–2995.
49. Yin,Y., Zhu,Q., Dai,S., Lamb,C. and Beachy,R.N. (1997) RF2a, a bZIP transcriptional activator of the phloem-specific rice tungro bacilliform virus promoter, functions in vascular development. *EMBO J.*, **16**, 5247–5259.
50. Chuang,C.F., Running,M.P., Williams,R.W. and Meyerowitz,E.M. (1999) The *PERIANTHIA* gene encodes a bZIP protein involved in the determination of floral organ number in *Arabidopsis thaliana*. *Genes Dev.*, **13**, 334–344.
51. Nakagawa,H., Ohmiya,K. and Hattori,T. (1996) A rice bZIP protein, designated OSBZ8, is rapidly induced by abscisic acid. *Plant J.*, **9**, 217–227.
52. Fukazawa,J., Sakai,T., Ishida,S., Yamaguchi,I., Kamiya,Y. and Takahashi,Y. (2000) Repression of shoot growth, a bZIP transcriptional activator, regulates cell elongation by controlling the level of gibberellins. *Plant Cell*, **12**, 901–915.

53. Menkens,A.E. and Cashmore,A. R. (1994) Isolation and characterization of a fourth *Arabidopsis thaliana* G-box-binding factor, which has similarities to F_os oncoprotein. *Proc. Natl Acad. Sci. USA*, **91**, 2522–2526.
54. Weisshaar,B., Armstrong,G.A., Block,A., Silva,O.C. and Hahlbrock,K. (1991) Light-inducible and constitutively expressed DNA-binding proteins recognizing a plant promotor element with functional relevance in light responsiveness. *EMBO J.*, **10**, 1777–1786.
55. Schindler,U., Menkens,A.E., Beckmann,H., Ecker,J.R. and Cashmore,A. R. (1992) Heterodimerization between light-regulated and ubiquitously expressed *Arabidopsis* GBF bZIP proteins. *EMBO J.*, **11**, 1261–1273.
56. Hobo,T., Kowyama,Y. and Hattori,T. (1999) A bZIP factor, TRAB1, interacts with VP1 and mediates abscisic acid-induced transcription. *Proc. Natl Acad. Sci. USA*, **96**, 15348–15353.
57. Finkelstein,R. and Lynch,T. (2000) The Arabidopsis abscisic acid response gene *ABI5* encodes a basic leucine zipper transcription factor. *Plant Cell*, **12**, 599–609.
58. Uno,Y., Furuhata,T., Abe,H., Yoshida,R., Shinozaki,K. and Yamaguchi-Shinozaki,K. (2000) Arabidopsis basic leucine zipper transcription factors involved in an abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions. *Proc. Natl Acad. Sci. USA*, **97**, 11632–11637.
59. Niggeweg,R., Thurow,C., Kegler,C. and Gatz,C. (2000) Tobacco transcription factor TGA2.2 is the main component of as-1-binding factor ASF-1 and is involved in salicylic acid- and auxin-inducible expression of as-1-containing target promoters. *J. Biol. Chem.*, **275**, 19897–19905.
60. Finkelstein,R. and Lynch,T. (2000) . The Arabidopsis abscisic acid response gene *ABI5* encodes a basic leucine zipper transcription factor. *Plant Cell*, **12**, 599–609.
61. Fujita,Y., Fujita,M., Satoh,R., Maruyama,K., Parvez,M.M., Seki,M., Hiratsu,K., Ohme-Takagi,M., Shinozaki,K. and Yamaguchi-Shinozaki,K. (2005) AREB1 is a transcription activator of novel ABRE-dependent ABA signaling that enhances drought stress tolerance in Arabidopsis. *Plant Cell*, **17**, 3470–3488.