Isolation and Characterization of the Trichome-specific *AtTSG1* Promoter from *Arabidopsis thaliana*

**Shiming Ni · Lingjun Meng · Ji Zhao ·
Xuechen Wang · Jia Chen**

Abstract Trichomes are specialized epidermal cells that are commonly present on leaves, stems, and sepals of plants. They are suggested to provide a first line of defense against invading pests and pathogens. By differential display reverse transcription (DDRT)-polymerase chain reaction (PCR) and reverse Northern analysis, an *Arabidopsis thaliana* gene, *AtTSG1*, has been identified from epidermal cells of leaves, and found to be specifically expressed in the epidermis. Following BLAST analysis, it is found that this gene encodes a putative translational activator protein, At1G64790, with similarity to *HsGCN1*. A 1.3-kb promoter sequence of the *AtTSG1* gene has been cloned. This promoter is shown to direct the specific expression of the reporter *uidA* gene, β-glucuronidase (GUS), in trichomes of leaves and stems of *Arabidopsis* plants. Promoter deletion analysis has revealed that the region from ×300 bp to ×1 bp is sufficient to direct trichome-specific expression, and that a novel *cis*-acting element to direct trichome-specific expression is involved in the region from ×250 bp to ×200 bp. The *AtTSG1* promoter sequence from ×100 to ×1 bp likely contains an initiator (Inr) sequence as a core promoter element to mediate the same function as that of a TATA element. As this *AtTSG1* promoter is associated with trichomes, it may provide an efficient bioengineering element for enhancing pest and pathogen resistance in transgenic plants.

Keywords Trichome-specific promoter · *cis*-acting element · Deletion analysis · β-glucuronidase

S. Ni · L. Meng · J. Zhao · X. Wang · J. Chen (✉)
State Key Laboratory of Plant Physiology and Biochemistry, College of Biological Science,
China Agricultural University, Beijing 100094, China
e-mail: chenja@public.bta.net.cn

Introduction

Trichomes are specialized epidermal cells that are distributed along above-ground epidermal tissues of most plants (Johnson 1975). During plant development, each trichome undergoes cell morphogenesis leading to a specialized architecture. Trichomes may be composed of one or more cells and may be unbranched or branched. They frequently function as a first line of defense against insect attack, either by spatial hindrance, i.e., physical entrapment, or by secreting either toxic or behavior modifying chemicals (Duffey 1986; Wagner 1991; Payne et al. 1999). This trichome-based defense system plays an important role in the development of sustainable pest control strategies (Lewis et al. 1997). Glandular trichomes often secrete diverse natural products that are economically important (Lange and Croteau 1999).

Wild-type *Arabidopsis* trichomes are unicellular, typically have three branches, and lack glands. However, it is enough to provide physical protection from attack by predators because the production of trichomes is correlated with reduced damage from insect herbivores in natural populations (Maarit et al. 2007). Trichomes of *Arabidopsis thaliana* are a model system for studying cell development, differentiation, and cell cycle regulation. To exploit this model system with ultimate spatial resolution, gene expression profiling was recently investigated at a single-cell level (Kryvych et al. 2007; Lieckfeldt et al. 2007). A few transcription factor families have been found to be involved in the trichome initiation and morphogenesis, including MYB proteins (Oppenheimer et al. 1991; Larkin et al. 1993; Trudi et al. 2003), HD-ZIP protein (Hulskamp et al. 1994; Rerie et al. 1994), and bHLH proteins (Larkin et al. 1994; Szymanski et al. 1998; Schnittger et al. 1999; Payne et al. 1999).

It would be advantageous to bioengineer the trichome system as plant trichomes are not essential structures for plant growth and development (Wang et al. 2002). Trichome specific bioengineering would restrict the expression of foreign genes to trichomes, and foreign proteins in trichomes might have very high-level product accumulation.

Tobacco CYP71D16 is a trichome-specific P450 protein that catalyses the hydroxylation of cembratrieneol to form cembratrienediol (Wang et al. 2001). The tobacco *CYP71D16* gene promoter was shown to direct the specific expression of the β-glucuronidase (GUS) reporter gene in glandular trichomes of *Nicotiana tabacum* at all developmental stages. Promoter deletion analysis revealed that the region from −223 bp to +111 bp is sufficient to direct trichome-specific expression (Wang et al. 2002). The promoters of the cotton lipid transfer protein genes *LTP3* and *LTP6* were also found to direct GUS expression in trichomes, but not in non-trichome cells of leaves in transgenic tobacco seedlings (Hsu et al. 1999; Liu et al. 2000). A cotton fiber gene, *RD22*-like1 (*RDL1*), encodes a protein with a homeodomain-binding L1 box and an MYB-binding motif; and its promoter confers trichome-specific expression in *Arabidopsis* (Wang et al. 2004).

To analyze the *cis*-acting elements involved in the trichome-specific expression, we cloned the promoter of the *Arabidopsis* gene (At1G64790). The promoter and β-glucuronidase (GUS) reporter fusion vectors were constructed and transformed into *Arabidopsis* plant. The GUS reporter gene driven by the promoter was specifically expressed at significant levels in trichomes on the leaves and stems.

Materials and Methods

Plant Materials

Seeds of *Arabidopsis thaliana* (Columbia) were vernalized at 4°C for 3 days and grew at 22°C with a photoperiod of 12 h. Seedlings were moved to the soil 7-day post germination and grew with a photoperiod of 16 h. To collect epidermal and mesophyll cells from 4-week-old seedlings leaves, epidermal cell layer were torn from leaves by the nipper, then epidermal and mesophyll cell layers were brushed to remove contamination. After checking with microscopy, epidermal and mesophyll cells were frozen quickly in liquid nitrogen, and stored at −80°C for DNA and RNA extraction.

RNA Extraction and Differential Display RT-PCR

Total RNA was extracted using a TRIzol reagent kit (Tianwei, Beijng, China). RNA was treated with RNase-free DNase I for 15 min at 37°C. Differential display reverse transcription (DDRT) was then performed as described (Liang and Pardee 1992). Briefly, 0.2 μg of RNA and 50 μM anchor primer (T11A, T11C or T11G) was reverse transcribed in a 20-μl reaction mixture at 42°C for 60 min with Moloney murine leukemia virus (M-MLV) reverse transcriptase (Promega). PCR was performed using 1/10 v of cDNA, 2 μL of deoxyribonucleotide triphosphate (dNTP), 1 U of Taq DNA Polymerase, 2.5 μM of anchor primer, and 0.5 μM of an arbitrary upstream primer (30 s at 94°C, 2 min at 40°C, and 30 s at 72°C for 40 cycles). Upstream random primers were U1: 5′AGACCAGACT3′, U2: 5′CTGACT GATC3′, U3: 5′CTAGCTTCGT3′, U4: 5′GCACGAACCG3′, U5: 5′CATGCGC GAG3′, and U6: 5′GGATGCCTCG3′. In total, DDRT-PCR was performed using 18 primer combinations. DDRT-PCR products were separated on 6% sequencing denaturing gels and visualized by silver staining (Gardiner et al. 1986). After excision from the gel, partial cDNAs were re-amplified by PCR in the same conditions and cloned using a pGEM®-T Easy Cloning Kit (Promega). The cloned fragments were sequenced with M13 forward and reverse primers.

Reverse Northern Analysis

Forty complementary DNA (cDNA) fragments obtained by DDRT-PCR were amplified by PCR in standard conditions using the universal direct and reverse primers. PCR products were dotted onto two nylon membrane Hybond-N+ (Amersham- Pharmacia). The complex cDNA probes were synthesized at 42°C for 60 min using 5 μg of DNase-treated RNA in a solution with 30 μCi of α-^{32}P-dCTP, 200 U of M-MLV reverse transcriptase, 20 U of RNasin, and 10mM dNTP without dCTP. Prehybridization was performed at 65°C for 6 h in 1%BSA, 0.25 M Na_2HPO_4 (pH7.2), 1 mM ethylenediaminetetraacetic acid (EDTA), and 7% sodium dodecyl sulfate (SDS); then blots were hybridized with complex cDNA probes at 65°C for 16 h. After washing twice at 65°C for 15 min in 2× SSC/1% SDS, twice at 65°C for 15 min in 2× SSC/0.5% SDS, and once at 65°C for 30 min in 0.1× SSC/0.1% SDS, the membranes were exposed to an x-ray film (Kodak, Little Chalfont, Buckinghamshire, UK) using two intensifying screens at −80°C.

RT-PCR Analysis

Total RNA was extracted from 100 μg of frozen tissue using a TRIzol reagent kit following the manufacturer's instructions. The first strand cDNA was synthesized using oligo (dT) primer and M-MLV reverse transcriptase in the presence of 1 unit RNase inhibitor. To carry out RT-PCR, a pair of gene-specific primers was designed: 5′- ATG TTG AAA GAT ATG GAG CCG-3′ (forward) and 5′- TGG TGA TGC CAA AGA AGA AAT-3′ (reverse). The *Actin2* gene was used as an internal control for RT-PCR (Volkov et al. 2003). The PCR conditions were as follows: 94°C for 5 min and 35 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 1 min, with a final extension of 7 min at 72°C. PCR products were separated on a 1% agarose gel containing ethidium bromide and photographed.

Promoter Cloning

Based on the nucleotide sequence of genomic DNA, two specific primers, promotor1 (*GAA TTC* AGT GGA CCG CAA ACA CGT ATG, the *Eco*RI site underlined) and promotor2 (*CTG CAG* TTC TTT CCA CGG AGA TTG ACC, the *Pst*I site underlined), were designed for amplification of the 1.3-kb promoter fragment. The PCR product was cloned into the pGEM®-T vector (Promega), and M13 forward and reverse primers were used to sequence the cloned fragment.

Construction of Deleted Promoter-GUS Fusion Vectors

The 1.3-kb promoter and a series of 5′ progressive promoter deletions were generated using PCR. The positions of the forward primers for deletion constructs were: P-1100 (from −1100 bp to −1075 bp), P-920 (from −920 bp to −895 bp), P-710 (from −710 bp to −685 bp), P-510 (from −510 bp to −485 bp), P-300 (from −300 bp to −275 bp), P-250 (from −250 bp to −225 bp), P-200 (from −200 bp to −175 bp), P-150 (from −150 bp to −125 bp), and P-100 (from −100 bp to −75 bp). An *Eco*RI restriction enzyme site was added at the 5′ end of each of the nine forward primers. A *Pst*I restriction site was added at the 3′ end of the reverse primer (from −25 bp to −1 bp). Each of the nine promoter fragments was double digested with *Eco*RI and *Pst*I, and ligated into the *Eco*RI and *Pst*I sites of the vector pCAMBIA1381, immediately upstream of the ATG start codon for β-glucuronidase. Each deletion was verified by sequencing. Constructs were introduced into *Agrobacterium* and used to transform *Arabidopsis*.

Transgenic Plants

The pCAMBIA1381 vectors containing the promoter–GUS fusions were transferred into *Agrobacterium* strain GV3101 for transformation of *Arabidopsis*. Transformation of *Arabidopsis* (Colombia ecotype) was performed as described previously (Valvekens et al. 1988; Benfey et al. 1989). Primary transgenic plants of *Arabidopsis* were grown at 22°C under a 16-h light/8-h dark cycle, with a photon flux density of 2,000 lx. Mature seeds were collected for transgenic selection under 25 mg/l hygromycin. Putative transgenic plants were analyzed for integration of the intact

promoter-GUS chimeric gene into the genomic DNA by PCR using a border sequence from the multicloning site in pCAMBIA1381 and a sequence from the GUS coding region as primers and genomic DNA as a template.

Histochemical and Fluorometric GUS Assay

Leaf discs and stem, root, and flower petal pieces of the transgenic plant were stained in a 0.1% X-Gluc solution according to the method of Jefferson (1987). After incubation at 37°C for a period of 3 h to overnight, tissues were destained in 50 to 80% ethanol until chlorophyll pigments were completely removed.

The leaves from greenhouse-grown transgenic plants were used for fluorometric GUS assays. The leaves were ground into a fine powder with liquid nitrogen, and then vortexed with GUS extraction buffer containing 50 mM sodium phosphate (pH 7.0), 10 mM β-mercaptoethanol, 10 mM Na_2-EDTA (pH 8.0), 0.1% Triton X-100, and 1% sarcosyl. The extracts were centrifuged for 15 min in a microcentrifuge at 4°C, and the supernatants were collected for GUS assays according to Jefferson (1987) using 4-MUG as a substrate. The amount of 4-MU production was determined using a luminescence spectrometer, and MUG activity was expressed as fluorescence unit h^{-1} mg^{-1} protein.

Results

Identification and Confirmation of the *Arabidopsis* Epidermis-specific Expressed Transcripts

To identify epidermis-specific expression genes, differential display reverse transcription (DDRT) analyses were conducted using RNAs extracted from epidermal and mesophyll cells. Eighteen combinations of primers were tested, and 40 differential bands were found. An example of DDRT-PCR results was shown in Fig. 1.These cDNA fragments were further verified by reverse Northern blot analysis and three of differential expressed cDNA fragments (named as C19, C110, and C112) were selected for further investigation. After cloning into pGEM®-T vector, sequence analysis showed that cDNA fragments C19, C110, and C112, were long: 280, 206, and 296 bp, respectively.

A Blastn search of *Arabidopsis* databases revealed that the cDNA fragment C19 was identical to an *Arabidopsis* gene (At1G64790) known as *AtTSG1* (*Arabidopsis thaliana trichome-specific-gene 1*). The *AtTSG1* codes for a putative translational activator family protein of 2441 amino acids with a calculated molecular mass of 265.34 kDa and p*I* 6.61, similar to HsGCN1, and involved in general control of amino acid synthesis. A Blastp search indicated that the AtTSG1 protein contains two putative conserved domains known as TIP120 and DUF813. The TBP (TATA-binding protein) -interacting protein 120 (TIP120) is reported to be a unique global transcription factor that can interact with TBP and stimulating all classes of eukaryotic transcription. An *A. thaliana* protein of unknown function (DUF813) consists of several uncharacterized proteins from *A. thaliana*.

Fig. 1 An example of DDRT-PCR results. **A**. Total RNA isolated from epidermis in *Arabidopsis*. **B**. Total RNA isolated from mesophyll in *Arabidopsis*. Arrows indicate differentially expressed cDNA fragment


A   B
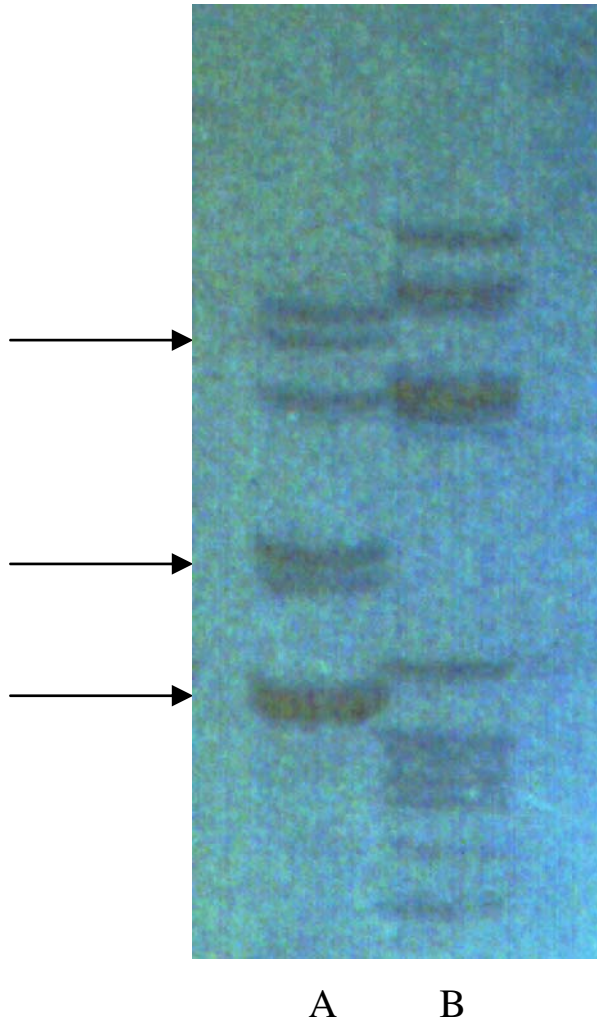
RT-PCR Analysis of the Transcript Levels in Various Tissues

Semi-quantitative RT-PCR experiments were performed on various *Arabidopsis* wild-type tissues. The *AtTSG1* transcript was detected in stems, leaves, epidermis, and the amount in epidermis was more than in leaves (Fig. 2). At the same time, little or no gene expression was detected in roots, flowers, and mesophyll (Fig. 2). In contrast, *actin* transcript levels were similar in these tested tissues.

Isolation of the *AtTSG1* Promoter

Based on the nucleotide sequence of the *Arabidopsis* At1G64790 locus, two specific primers were used to clone 1.3 kb promoter sequence (Fig. 3). A search for cis-regulatory elements in the promoter revealed that three MYB-like recognition sites (−33
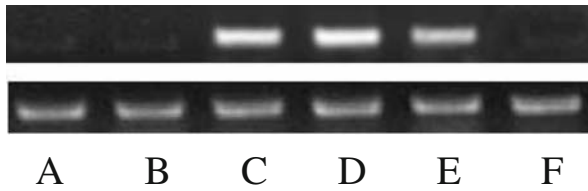
Fig. 2 RT-PCR analysis of *AtTSG1* expression (upper panel). mRNA were extracted from different *Arabidopsis thaliana* tissues: **A**. Flowers; **B**. Roots; **C**. Stems; **D**. Epidermis; **E**. Leaves; **F**. Mesophyll. Relative levels of *AtTSG1* transcript were shown in upper panel and *actin* transcript was detected as an internal control in lower panel

to −39, ACCTACC; −125 to −131, AACTAAC; −199 to −206, AACCAAAC, Fig. 3), one ABRE-like binding site motif (−847 to −854, GACACGTA, Fig. 3) and one DRE-like promoter motif (−993 to −1001, TTGTCGGCC, Fig. 3) were present, but the 'TATA' box and 'CAAT' box were not found on the upstream of the first codon ATG.

Expression Patterns of the *AtTSG1* Promoter–GUS Fusion

The 1301-bp genomic fragment containing the promoter region and the 5′ untranslated region of *Arabidopsis AtTSG1*gene was fused to the GUS coding

```
-1301  AGTGGACCGCAAACACGTATGCATGATTGGTTTAAAAATTATTTTGTGTTTTTTTTTGTAAAAA
-1236  ATTAATATGCATGTGTTATTAAATGGATTTAAGCCGTTTTAAGTGATCTTAAGAAAGTGTATTTT
-1171  TAGTGTATTTTCTAATATCTAACTATACAGACATAAATCTATAACTTTTAACTCAATAATATACA
-1106  TATTTATAACCGTTTATATGTTGTTTGATGTGATTTAAGTGAGTGTTGTTGTCTAATTTTTGTTT
-1041  AGAATTTATCATTTTTAACACACAATTTTTTATTTATTTT**TTGTCGGCC**ATTTTGATTGATAAAG
-976   CCCAAATTGGGCAGATTACAGAAGGTCCAAAGAGGATTACATATGGAAATTAATAGGCCAGGCCC
-911   AAAACTAAATCAAACAGAGAAATAAAATACAATTAAAGTATAGTTGATTTAGGGTTG<u>GACACGTA</u>
-846   TATGATACGTGGTCTTCACATTGCCACGCGTCAGATGCAAGCTCCGCCGTTAGTTCGCCGGAAAC
-781   CGCTAAAAACTCCGATGAGCCACCACCGATCAGCTCTGAAGCGGTGAACGTCCTCTTTCCTTCTG
-716   TGGATCCCAATTCCAAACTGTTTCGAAATTCGCTGAACATCACCATCTCCAGGGAGGCTCCACCA
-651   CTGACTACATCGAGAATTGATGTAATCTGTTCTTTGCTCCTATGACATGAGCAAAAATAAGTGAC
-586   GAAGATATAATTCCACAAAAGATTAGAGAGATCAGACGAGATTTGGAAGTAAAACAGAGGGCTTC
-521   GACATCGCAGTTGTAAAAGAGATCTAAACCAGACTTTGCCGGTAAGGAAATCCGATCTAAAGTCG
-456   GAACTTTATTTATCTCGCACAGAGACGAGAATATGTACAAATCCGGACAAACCGGAATCAAAAAG
-391   GATGAACGATGGTGTCTCCTCTCTGAAAGGTTTGGGGAGAGAGAAGACAGGAGAAAACCCCTAAT
-326   TATATTTTGTCTTAGATCTAGTTTTTTTAACACACAAATTCTTCCCACCAAAAAACACACACAAA
-261   TTCCAACAT*TAAAAG*TGAATCCATTATCCATGGACTT*TAAAAG*CCCGAATGAACA**AACCAAAC**TT
-196   AAATCGAAGCCCAAAAACAAACTTTCAGATTCATTTTCAGTTTCTTTCACTCTTCATCTTCTGCA
-131   **AACTAAC**CCACTGGCTTAGCCTAAACCCTAGTAGCCACAGAGACGAAGAAGAAGAAGAAGCTTCA
-66    CCGTTTTACCCTTTTACGA̲TTGTTCTG**ACCTACC**AACCTGGCCCGGTCAATCTCCGTGGAAAGAA**ATG**
                                                                            **+1**
```

Fig. 3 Nucleotide sequence of the *AtTSG1* promoter. The A of the first codon ATG is assigned to +1, the putative translation start site is in bold, underlined letters. The three putative MYB recognition sites (−33 to −39, −125 to −131, and −199 to −206) are in large, bold, and underlined letters, one putative ABRE binding site (−847 to −854) is in large and underlined letters, and one putative DRE binding site (−993 to −1001) is in bold and large letters

region of the vector pCAMBIA1381. This construct was used to transform *Arabidopsis*. Leaf discs, stem, flower petal, and root from primary transformants (T0) were tested for histochemical expression of GUS. Four independent transformation experiments were carried out. From each transformation experiment, 12 transgenic lines were tested for GUS expression pattern. Trichome-specific GUS expression was clearly found in four, five, five, and six of 12 transgenic lines from each of four transformation experiments, respectively. The remaining transgenic lines showed no GUS activity in trichomes and other tissues. Trichome-specific GUS expression was obviously found on stem and leaf discs, but not on flower petal and root (Fig. 4). No GUS activity was detected in non-trichome cells of the leaf epidermis. GUS staining was not even distributed in the trichome cells with intensive blue staining in the bottom and middle parts and less staining in the tips of trichome branches (Fig. 6A)

Determination of the *AtTSG1* Promoter Region for Trichome-specific Expression

To characterize the *cis*-elements that are critical to trichome-specific expression, a series of 5′ deletion constructs were made. First, P-1100 to P-300 constructs was made (Fig. 5). The transgenic plants were generated and leaf discs from each constructs were tested histochemically for GUS expression. Trichome-specific GUS expression was found on leaf discs from all five constructs from promoter deletions from P-1100 to P-300. For each of these five constructs, at least five independent transgenic lines showed distinctive and consistent trichome-specific GUS staining patterns. None of the transgenic lines from each of the five constructs showed GUS
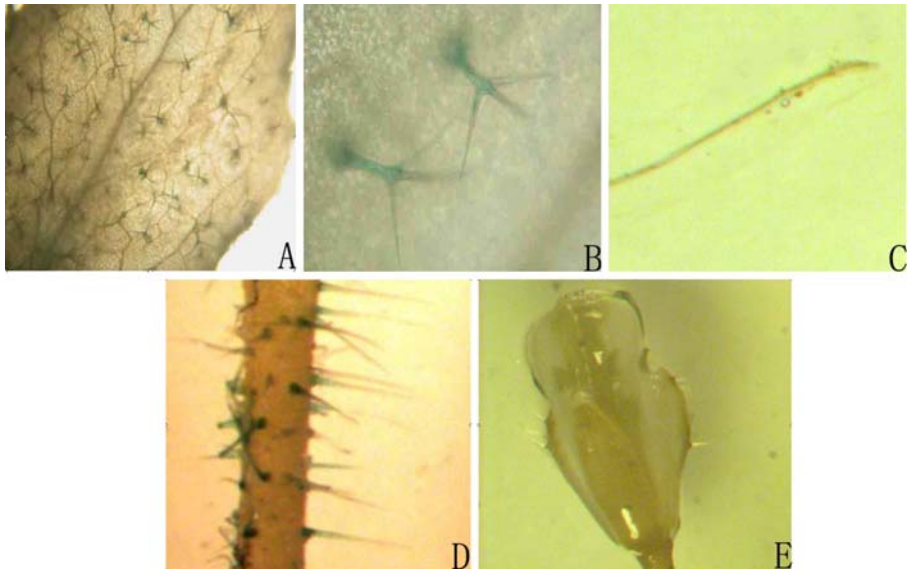

Fig. 4 GUS histochemistry location in *Arabidopsis* plants transformed with *AtTSG1* promoter-GUS fusion vectors. **A**. Trichomes on the leaf; **B**. Enlarged trichomes on the leaf; **C**. Root; **D**. Trichomes on the stem; **E**. Trichomes on the sepal

staining in tissues other than trichomes. These data showed that the region from −300 bp to −1 bp includes the *cis*-elements critical to trichome-specific expression.

To quantitatively determine promoter regions (P-1301 to P-300) of the *AtTSG1* gene that may affect trichome-specific expression, the average GUS activity for each of the five constructs was determined fluorometrically using the MUG assay. For each constructs, five transgenic lines were tested and the average GUS activity of the P-1301 promoter was set at 100% to define relative GUS activities of each promoter deletion. As shown in Fig. 5, the relative GUS activities for P-1100, P-710, and P-300 were significantly lower compared with transgenic plants from P-510, indicating that the region from −710 bp to −510 bp includes a repressive element and the region from −510 bp to −300 bp contains enhancer element.

To map the *cis*-elements controlling trichome-specific expression in the P-300 promoter region, four additional 5′-promoter deletions were made at P-250, P-200,
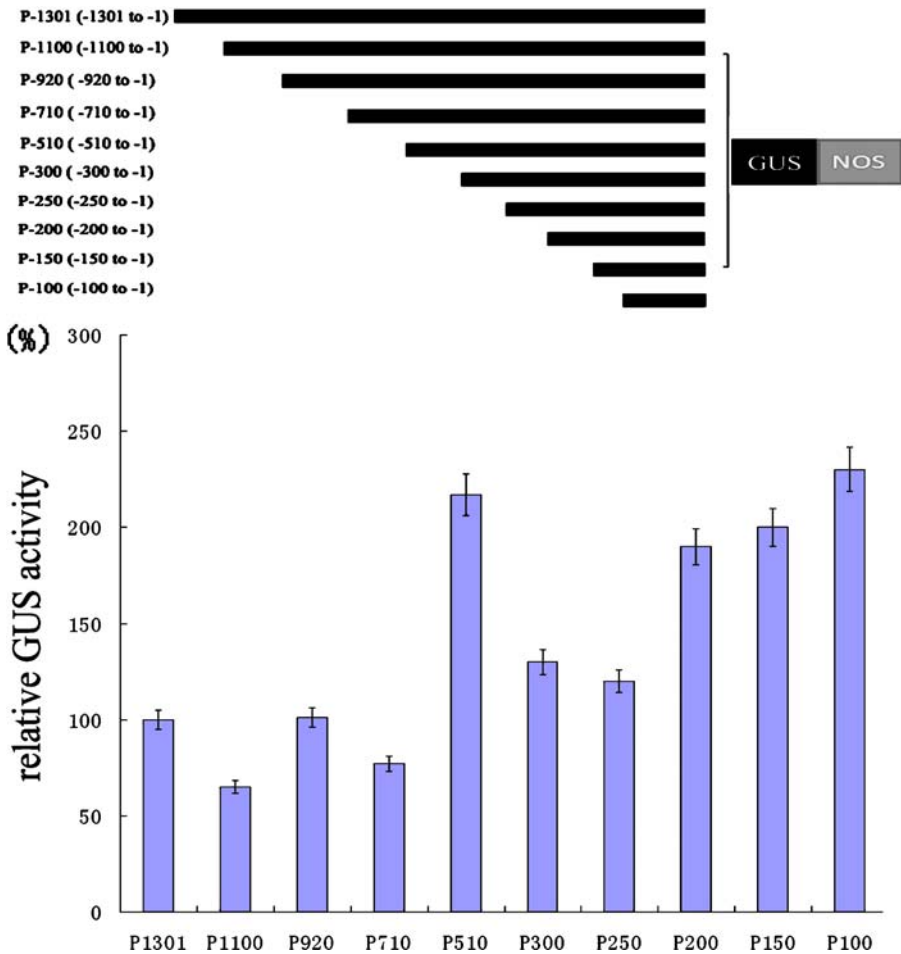

Fig. 5 Comparison of the GUS activity levels controlled by the *AtTSG1* promoter with a series of 5′-deletions. P-1301: longest promoter (−1301 bp); P-1100 to P-100: promoter deletions from −1100 bp to −100 bp

P-150, and P-100, respectively (Fig. 6); and the leaf discs for each construct were examined histochemically for expression of GUS (Fig. 6). The P-250 transformants still exhibited trichome-specific expression of GUS activity. However, the transgenic plants from P-200, P-150, and P-100 showed extensive GUS expression in other leaf cell types than trichomes (Fig. 6). These extensive GUS expression directed by P-200, P-150, and P-100 also resulted in increased levels of GUS expression (Fig. 5). These observations indicate that the 50-bp sequence between −250 and −200 includes *cis*-acting elements involved in the trichome-specific GUS expression.

Discussion

As specialized epidermal cells, wild-type *Arabidopsis* trichomes are unicellular, typically have three branches, and lack glands. Although plant trichomes are not
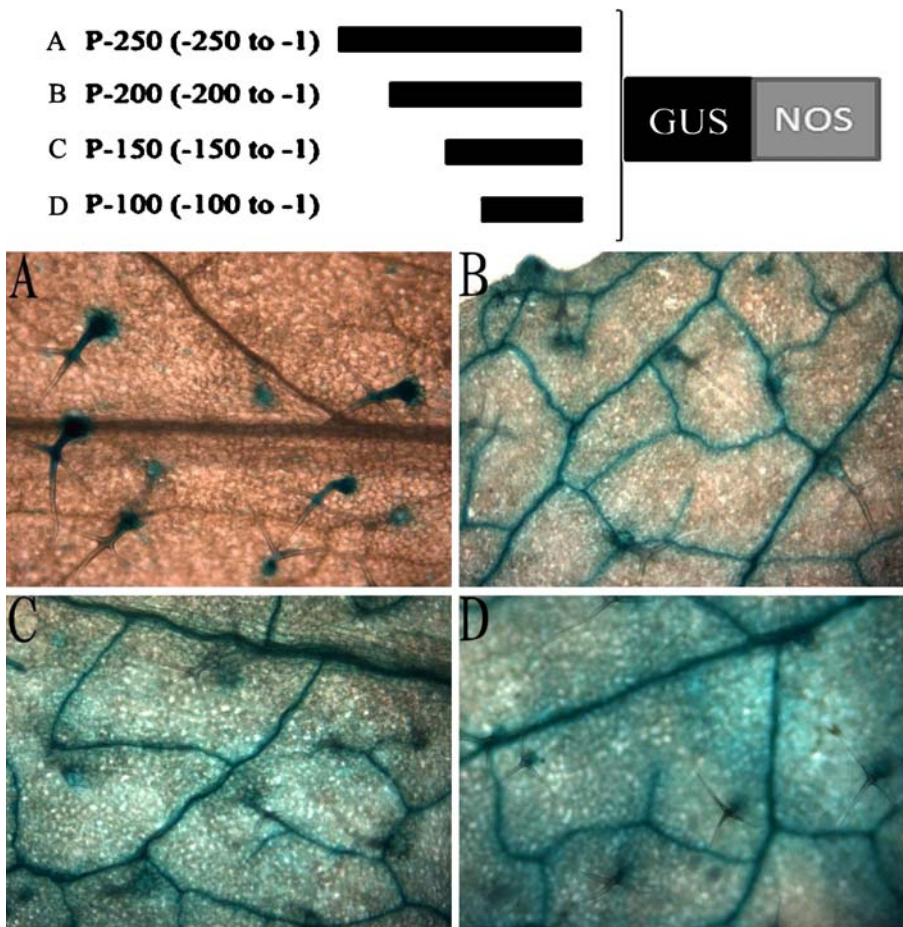
Fig. 6 Histochemical GUS location in transgenic *Arabidopsis* plants with promoter GUS fusion vectors of a series of the *AtTSG1* promoter deletions. (**A**) P-250, (**B**) P-200, (**C**) P-150, (**D**) P-100

essential, they provide a first line of defense against pests and pathogens. Recently, many genes related to trichomes have been found; most of them control the initiation and morphogenesis of the trichomes in *Arabidopsis* and encode several classes of transcriptional factors, such as *GL1*, *GL2*, *GL3*, and *TTG* (Oppenheimer et al. 1991; Larkin et al. 1994; Rerie et al. 1994; Payne et al. 1999). At the same time, the promoters from some genes, such as *CYP71D16, LTP3* and *LTP6*, can direct GUS expression in trichomes, but not in non-trichome cells of leaves (Wang et al. 2002; Hsu et al. 1999; Liu et al. 2000). Also, several plant gene promoters, which direct specific expression, primarily in both epidermal cells and trichomes have been reported, including *Arabidopsis GL1* gene (Larkin et al. 1993), *Brassica oleracea wax9D* gene (Pyee and Kolattukudy 1995), and tobacco *Itp1* gene (Canevascini et al. 1996).

In this paper, we report on cloning and characterization of the promoter of an *Arabidopsis* trichome-specific gene, *AtTSG1*. GUS histochemical localization in *Arabidopsis* transformed with the 1.3-kb *AtTSG1* promoter-GUS revealed that the gene driven by this promoter resulted in trichome-specific expression on both stem and leaves, but not on petals or roots (Fig. 4). These results were consistent with RT-PCR analyses (Fig. 2) as GUS activity detected along the top section of the leaf trichome was lower than that along the bottom and middle sections. This *Arabidopsis* gene encodes a putative translational activator family protein, similar to HsGCN1, and more research is needed to explore its function in trichomes.

To characterize the *cis*-elements that are critical to trichome-specific *AtTSG1* expression, a series of 5′ deletion constructs were made. Analyses of these promoter deletion constructs showed that the region from −300 bp to −250 bp contains the *cis*-elements critical to trichome-specific expression. Promoter activity was analyzed by monitoring MUG hydrolysis in homogenates of leaves from P-1301 to P-300. The relative GUS activities for P-1100 and P-710 were noticeably lower than the plants from P-510 (Fig. 5), suggesting that a repressive element was localized in the promoter region from −1301 bp to −510 bp. The relative GUS activity level and pattern for P-300 were almost the same as those from P-1301, indicating that P-300 contains all necessary elements for normal *AtTSG1* expression.

MYB recognition sites were already identified in the promoter region of a tobacco trichome-specific gene *CYP71D16* (Wang et al. 2002) and in the trichome-specific promoters of both the *LTP3* and *LTP6* genes from cotton (Hsu et al. 1999; Liu et al. 2000). The *MYB* gene family represents one of the largest regulatory factor families in plants, and some members of the *MYB* family control the identity and fate of plant cells (Stracke et al. 2001). Three MYB-like recognition sites (Fig. 3) were present in the full promoter. Analyses of promoter 5′ deletion to positions −250, −200, −150, and −100 indicate that the 50-bp promoter region between −250 and −200 includes *cis*-acting elements involved in the trichome-specific GUS expression. The region between −250 and −200 contains one MYB-like recognition site (AACCAAAC, −199 to −206). Some MYB transcriptional factors (GL1, CAPRICE, TRIPTY-CHON, and WER) control the initiation and morphogenesis of the trichomes in *Arabidopsis*. Further analysis with site-mutagenesis is needed to determine the role of this MYB-like recognition site (−199 to −206) for the trichome-specific expression of the *AtTSG1* promoter.

Regulation of gene expression is mediated by specific interactions of transcription factors with promoter DNA sequences. RNA pol II promoters are conceptually

divided into two domains, upstream regulatory and core promoter regions. Core promoters can be grouped according to the presence of specific DNA sequence elements such as TATA box (Singer et al. 1990; Butler and Kadonaga 2002) and transcriptional initiator (Inr) (Smale et al. 1998; Lo and Smale 1996; Smale et al. 1998). An Inr for eukaryotic RNA polymerase II can be defined as a DNA sequence element that overlaps a transcription start site and mediates the action of at least some upstream activators in the absence of a TATA box (Kollmar and Farnham 1993; Smale 1994; Weis and Reinberg 1992). Although the 'TATA' box was not present in the upstream region of the ATG start codon in the *AtTSG1* promoter, the promoter region (−100 to −1) in the P-100 vector conferred a relatively high level of GUS activity (Fig. 6). Most functional Inr elements contain a CA sequence (where A refers to the transcription start site), and are surrounded by either a few or several pyrimidines (Kollmar and Farnham 1993; Smale 1994; Weis and Reinberg 1992). In general, the larger the number of pyrimidines in these locations was, the greater the activity of the Inr (Javahery et al. 1994). Moreover, if a large number of pyrimidines are present surrounding the start site, low levels of Inr activity can be imparted in the absence of the A at the transcription start site (Javahery et al. 1994). In the *AtTSG1* promoter, the region from −52 to −71 is rich of pyrimidines (17/20), and the upstream of the potential transcription start site A is a pyrimidine C(-68), surrounded by a few pyrimidines, which fits well the Inr definition. These structural features of the *AtTSG1* promoter suggest that an Inr exists in this region to drive gene expression.

More research is needed to delineate the *cis*-elements responsible for trichome-specific expression of the *AtTSG1* gene. The *AtTSG1* promoter with trichome-specific expression may provide a useful genetic element for enhancing trichome-based pest resistance in plants.

References

Benfey PN, Ren L, Chua N-H. The CaMV 35s enhancer contains at least two domains which can confer different developmental and tissue-specific expression patterns. EMBO J. 1989;8:2195–2202.

Butler JE, Kadonaga JT. The RNA polymerase II core promoter: a key component in the regulation of gene expression. Genes Dev 2002;16:2583–92.

Canevascini S, Caderas D, Mandel T, Fleming AJ, Dupuis I, Kuhlemeier C. Tissue-specific expression and promoter analysis of the tobacco Itp1 gene. Plant Physiol 1996;112:513–24.

Duffey SS. Plant glandular trichomes: their partial role in defense against insects. In: Juniper BE, Southwood TRE, editors. Insects and the plant surface. London: Edward Arnold; 1986. p. 151–72.

Gardiner K, Laas W, Patterson D. Fractionation of large mammalian DNA restriction fragments using vertical pulsed-field gradient gel electrophoresis. Somat Cell Mol Genet 1986;12:185–90.

Hsu C, Roy GC, Jenkins JN, Ma D. Analysis of promoter activity of cotton lipid transfer protein gene LTP6 in transgenic tobacco plants. Plant Sci 1999;143:63–70.

Hulskamp M, Misera S, Jurgens G. Genetic dissection of trichome cell development in *Arabidopsis*. Cell 1994;76:555–66.

Javahery R, Khachi A, Lo K, Gregory BZ, Smale ST. DNA sequence requirements for transcriptional initiator activity in mammalian cells. Mol Cell biol 1994;14:116–27.

Jefferson RA. Assaying for chimeric genes in plants: the GUS fusion system. Plant Molecular Biology Reporter 1987;5:387–405.

Johnson HB. Plant pubescence: an ecological perspective. Bot Rev 1975;41:233–58.

Kollmar R, Farnham PJ. Site-specific initiation of transcription by RNA polymerase II. Proc Soc Exp Biol Med 1993;203:127–39.

Kryvych S, Nikiforova V, Herzog M, Perazza D, Fisahn J. Gene expression profiling of the different stages of Arabidopsis thaliana trichome development on the single cell level. Plant Physiol Biochem 2007;46(2):160–73.

Lange BM, Croteau R. Genetic engineering of essential oil production in mint. Curr Opin Plant Biol 1999;2:139–44.

Larkin JC, Oppenheimer DG, Pollock S, Marks MD. *Arabidopsis* GLABROUSI gene requires downstream sequences for function. Plant Cell 1993;5:1739–48.

Larkin JC, Oppenheimer DG, Lloyd A, Paparozzi ET, Marks MD. The roles of *GLABROUS1* and *TRANSPARENT TESTA GLABRA* genes in *Arabidopsis* trichome development. Plant Cell 1994;6:1065–76.

Lewis WJ, van Lenteren JC, Phatak SC, Tumlison JH. A total system approach to sustainable pest management. Proc Natl Acad Sci U S A 1997;94:12243–8.

Liang P, Pardee AB. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 1992;257:967–97.

Lieckfeldt E, Simon Rosin U, Kose F, Zoeller D, Schliep M, Fisahn J. Gene expression profiling of single epidermal, basal and trichome cells of Arabidopsis thaliana. J Plant Phys 2007. (in press).

Liu H, Creech RG, Jenkins JN, Ma D. Cloning and promoter analysis of the cotton lipid transfer protein gene Ltp3. Biochim Biophys Acta 2000;1487:106–11.

Lo K, Smale ST. Generality of a functional initiator consensus sequence. Gene 1996;182:13–22.

Maarit K, Katri K, Myriam G, Geir L, Jon A. Gene, phenotype and function: GLABROUS1 and resistance to herbivory in natural populations of *Arabidopsis lyrata*. Mol Ecol 2007;16:453–62.

Oppenheimer DG, Herman PL, Esch J, Sivakumaran S, Marks MD. A myb-related gene required for leaf trichome differentiation in Arabidopsis is expressed in stipules. Cell 1991;67:483–93.

Payne T, Clement J, Arnold D, Lloyd A. Heterologous myb genes distinct from GL1 enhance trichome production when overexpressed in *Nicotiana tabacum*. Development 1999;126:671–82.

Pyee J, Kolattukudy PE. The gene for the major cuticular wax-associated protein and three homologous genes from broccoli (Brassica oleracea) and their expression patterns. Plant J 1995;7:49–59.

Rerie WG, Feldmann KA, Marks MD. The GLABRA2 gene encodes a homeo domain protein required for normal trichome development in *Arabidopsis*. Genes Dev 1994;8:1388–99.

Schnittger A, Folkers U, Schwab B, Jürgens G, Hülskamp M. Generation of a spacing pattern: the role of TRIPTYCHON in trichome patterning in Arabidopsis. Plant Cell 1999;11:1105–16.

Singer VL, Wobbe CR, Struhl K. A wide variety of DNA sequences can functionally replace a yeast TATA element for transcriptional activation. Genes Dev 1990;4:636–45.

Smale ST. Core promoter architecture for eukaryotic protein-coding genes. In: Conaway RC, Conaway JW, editors. Transcription: mechanisms and regulation. New York: 63–81: Raven Press; 1994.

Smale ST, Jain A, Kaufmann J, Emami KH, Lo K, Garraway IP. The initiator element: a paradigm for core promoter heterogeneity within metazoan protein-coding genes. Cold Spring Harb Symp Quant Biol 1998;63:21–31.

Stracke R, Werber M, Weisshaar B. The R2R3-MYB gene family in *Arabidopsis thaliana*. Curr Opin Plant Biol 2001;4:447–56.

Szymanski DB, Jilk RA, Pollock SM, Marks MD. Control of *GL2* expression in *Arabidopsis* leaves and trichomes. Dev 1998;125:1161–71.

Trudi H, Li SF, Roger W. *AtMYB103* regulates tapetum and trichome development in *Arabidopsis thaliana*. Plant J 2003;35:177–92.

Valvekens D, Van Montagu M, Van Lijsebettens M. Agrobacterium tumefaciens-mediated transformation of Arabidopsis thaliana root explants by using kanamycin selection. Proc Natl Acad Sci. USA 1988;85:5536–40.

Volkov RA, Panchuk II, Schoffl F. Heat-stress-dependency and developmental modulation of gene expression: the potential of house-keeping genes as internal standards in mRNA expression profiling using real-time RT-PCR. J Exp Bot 2003;54:2343–9.

Wagner GJ. Secreting glandular trichomes: more than just hairs. Plant Physiol 1991;96:675–9.

Wang E, Wang R, Deparasis J, Loughrin JH, Gan S, Wagner GJ. Suppression of a P450 hydroxylase gene in plant trichome glands enhances natural-product-based aphid resistance. Nat Biotechnol 2001;19:371–4.

Wang E, Gan S, Wagner GJ. Isolation and characterization of the CYP71D16 trichome-specific promoter from *Nicotiana tabacum* L. J Exp Bot 2002;53:1891–7.

Wang S, Wang JW, Yu N, Li CH, Luo B, Gou JY, Wang LJ, Chen XY. Control of plant trichome development by a cotton fiber MYB gene. Plant Cell 2004;16:2323–34.

Weis L, Reinberg D. Transcription by RNA polymerase II: initiator-directed formation of transcription-competent complexes. FASEB J 1992;6:3300–9.