

Isolation and Characterization of Three New Promoters from *Gossypium hirsutum* that Show High Activity in Reproductive Tissues

Sinara Artico · Julia Lambret-Frotté · Sarah Muniz Nardeli · Osmundo Brilhante Oliveira-Neto · Maria Fátima Grossi-de-Sa · Marcio Alves-Ferreira

Published online: 7 November 2013
© Springer Science+Business Media New York 2013

Abstract Engineering of plant protection requires well-characterized tissue-specific promoters for the targeted expression of insecticidal resistance genes. Herein, we describe the isolation of five different fragments of promoters of three distinct flower-specific cotton (*Gossypium hirsutum*) genes. Expression analyses of the three genes *GhPME-like1*, *GhβGal-like1* and *GhPL-like1* revealed that they are expressed highly in flowers buds ranging from 4 to 12 mm in size. Several putative regulatory cis-elements were identified in the promoter regions, including elements involved in the control of tissue-specific gene expression in pollen grains and fruits. In vivo analyses of these promoters were performed using the heterologous plant system *Arabidopsis thaliana* by fusing them with the gene *uidA* (GUS). GUS staining in *Arabidopsis* tissues revealed that their expression was restricted to anthers, with the majority of expression in pollen grains

and in the upper portion of the carpels and siliques. A comparison between a CaMV35S::GUS constitutive promoter and the promoters isolated in this study revealed that the cotton promoters were more active and were specific to flowers and fruits, which are organs that are preferentially attacked by important pest insects such as the boll weevil (*Anthonomus grandis*). The activity of the promoters was also confirmed using transient expression assays in flower buds of *G. hirsutum*. The promoters of *GhPME-like1*, *GhβGal-like1* and *GhPL-like1* are specific to reproductive tissues and could represent important biotechnological tools for controlling insect pests, in particular the cotton boll weevil, which attacks floral and fruit tissues.

Keywords Cotton · Flower specific genes · Promoter characterization · *GUS* · CaMV35S · Cotton boll weevil

Electronic supplementary material The online version of this article (doi:10.1007/s11105-013-0674-0) contains supplementary material, which is available to authorized users.

S. Artico · J. Lambret-Frotté · S. M. Nardeli · M. Alves-Ferreira (✉)
Department of Genetics, Federal University of Rio de Janeiro (UFRJ), Av. Prof. Rodolpho Paulo Rocco, s/n - Prédio do CCS Instituto de Biologia, 2º andar, sala A2-93, 219410-970 Rio de Janeiro, RJ, Brazil
e-mail: alvesfer@uol.com.br

S. Artico
e-mail: sinaraartico@gmail.com

J. Lambret-Frotté
e-mail: lambret.julia@gmail.com

S. M. Nardeli
e-mail: snardeli@gmail.com

O. B. Oliveira-Neto · M. F. Grossi-de-Sa
Parque Estação Biológica (PqEB), Embrapa Genetic Resources and Biotechnology, Av. W5 Norte (final), Caixa Postal 02372, CEP 70770-900 Brasília, DF, Brazil

O. B. Oliveira-Neto
e-mail: osmundobrilhante@gmail.com

M. F. Grossi-de-Sa
e-mail: fatimasa@cenargen.embrapa.br

O. B. Oliveira-Neto
Department of Biochemistry and Molecular Biology, School of Medicine, FACIPLAC, SIGA AE2, Brasília 72460-000, DF, Brazil
e-mail: osmundobrilhante@gmail.com

M. F. Grossi-de-Sa
University Catholic of Brasilia, Brasília, DF, Brazil

Introduction

Cotton is a commercial crop that provides 90 % of the raw material for the textile industry and contributes 60 % of the world's oilseed requirements. However, cotton production is hampered by several insect pests that attack during the flowering and fructification of the plant, such as the fall armyworm (*Spodoptera frugiperda*), the budworm (*Heliothis virescens*) and the cotton boll weevil (*Anthonomus grandis*). Together, these insect pests are responsible for more than 50 % of insecticide costs borne by the Brazilian cotton crop farmer. Among the major pests, the cotton boll weevil pest is undoubtedly the most devastating (Greenberg et al. 2003; Martins et al. 2007b). The endophytic habit of the *A. grandis* larvae within the floral buds or cotton fruit destroys fiber quality and hampers chemical control (Haynes and Smith 1992). The most promising, cost effective and sustainable method for controlling cotton boll weevil is the development of genetically resistant cotton lines that suppress the development of the insect larvae (Grossi-de-Sa et al. 2007).

The source of the insecticidal toxins produced by commercial transgenic plants is the soil bacterium *Bacillus thuringiensis* (Bt). Bt strains show different specific insecticidal activities and contribute a large reservoir of genes encoding insecticidal proteins (Sanahuja et al. 2011). Silencing genes important for insect viability via the use of RNAi is another example of a promising strategy for insect pest control in agriculture (Mao et al. 2007). A recent study producing transgenic maize dsRNA directed against V-type ATPase of corn rootworm showed suppression of mRNA in the insect and reduction in feeding damage (Baum et al. 2007).

However, to increase the sustainability of genetically modified plants (GMP) in agriculture, it is crucial to develop tools that control more precisely heterogeneous expression of entomotoxic proteins or double-stranded RNA. Promoters that show high activity in tissues attacked by insect pests are highly desirable. Many of the promoters already in use in crop plants show conspicuous and high expression levels, such as the maize ubiquitin (Christensen and Quail 1996), rice actin (Zhang et al. 1991), banana streak virus (Schenk et al. 2001) and cauliflower mosaic virus 35S (CaMV35S) promoters (Guilley et al. 1982). Of these promoters, the CaMV35S is the one used most frequently to drive constitutive transgene expression (Odell et al. 1985; Chitkowski et al. 2003). However, CaMV35S promoter activity in floral tissues of cotton plants has been observed to be variable (Kranthi et al. 2005). Moreover, the expression of the Bt toxin Cry1A in ovaries for example is one of the lowest found in different cotton tissues (Kranthi et al. 2005). This pattern of expression is particularly critical to insects that attack these tissues, such as the cotton bollworm and the cotton boll weevil (Kranthi et al. 2005; Olsen et al. 2005; Martins et al. 2007a). Additionally, mRNA levels of the *Cry1Ac* (Ingard®) gene

and the selectable marker gene *Npt II*, both driven by the CaMV35S promoter, have been shown to decline during cotton plant development (Olsen et al. 2005). Laboratory and glasshouse studies have also suggested that CaMV35S activity is influenced by unknown physiological and possibly also environmental factors (Sunilkumar et al. 2002).

In addition to these negative aspects to the use of the CaMV35S promoter, the identification of new plant promoters is also crucial for the generation of plants with multiple traits (gene stacking)—a typical strategy used in plant biotechnology. If the same promoter is used to drive the expression of more than one transgene, sequence-dependent homologous silencing (transcriptional silencing) may occur. This problem is particularly evident when the promoter is also highly active, and it leads to impaired expression of one or more transgenes (Rocha et al. 2005). Thus, for successful gene stacking, promoters with little or no homology must be used in order to circumvent transcriptional gene silencing.

In this study, we isolated and characterized three new promoters that are active predominantly in cotton flowers, and that may provide more consistent levels of protein expression. The genes were initially identified as being highly expressed within an EST library of whole flowers. The expression patterns of these genes were evaluated in different plant organs, floral verticils and also during flower and fruit development. The activities of the corresponding flower gene promoters were then analyzed qualitatively and quantitatively in transgenic *Arabidopsis* plants using the *GUS* (β -glucuronidase) reporter gene. Their expression in reproductive tissues of cotton was confirmed using transient expression assays. In addition, the activity of the three promoters was compared to the CaMV35S promoter.

Materials and Methods

Plant Material and Growth Conditions

Cotton (*Gossypium hirsutum* variety “BRS Cedro”) plants were grown under controlled temperature (27 ± 2 °C) and natural photoperiod conditions at Embrapa Genetic Resources and Biotechnology (Brasília, DF, Brazil). Different vegetative and reproductive tissues from 3-month-old plants were used for qPCR experiments and for promoter isolation. For transient cotton transformation, 6-mm floral buds were used.

Arabidopsis thaliana ecotype Landsberg erecta (Ler) was transformed using the floral dip method to analyze promoter activity by fusing to the β -glucuronidase (GUS) reporter gene (Desfeux et al. 2000; Clough and Bent 1998).

In silico Analyses and Identification of *G. hirsutum* Flower-Specific Genes

To identify flower-specific genes, two separate databases were used: the PAVE library (<http://www.agcol.arizona.edu/cgi-bin/pave/Cotton/index.cgi>, Udall et al. 2006) and the Cotton Genome Database (US Department of Agriculture, Agricultural Research Service CottonDB, <http://cottondb.org/cdbhome.html>). Initially, 19 genes were selected from the PAVE library of whole flowers of *Gossypium raimondii* at the post-anthesis stage (−3 to +3 days) that showed high expression (seven or more ESTs). The sequences of these 19 genes were analyzed using BLASTN in the *Arabidopsis* database to find putative homologues (see supplementary Table S1). The homologues were scrutinized using the Genevestigator platform (<https://www.genevestigator.com/gv/>) to visualize reproductive-tissue specific expression. Seven genes showed high expression in flower tissues. The sequences of these genes were compared with sequences of *G. hirsutum* genes in the Cotton Genome Database using the tBLASTN program (Altschul et al. 1997). Multiple alignments for each gene family using amino acid sequences obtained from *G. hirsutum* (see Supplementary Table S1), *G. raimondii* and *Arabidopsis* were conducted using the CLUSTALW2 program using default parameters that were then revised manually (Larkin et al. 2007). Phylogenetic trees were constructed after using model test to pick the most appropriate evolutionary model for our analysis in the *Mega 5.2* program (Saitou and Nei 1987; Tamura et al. 2011). The best-fitting amino acid substitution model for all gene families was WAG (Whelan and Goldman 2001) with estimated γ -distribution parameter (G). For β -galactosidase reconstruction, an estimated proportion of the invariable sites (I) was also implemented. The maximum-likelihood analyses were performed with the program PhyML version 3.0 (Guindon et al. 2010; <http://www.atgc-montpellier.fr/phyml/>) and assessment of node confidence was performed using 1,000 bootstrap replicates (see supplementary Fig. S1). The amino acid sequences deduced as *GhPME-like1* (240 amino acid), *Gh β Gal-like1* (879 amino acid), and *GhPL-like1* (275 amino acid) genes were evaluated in the PFAM database (<http://pfam.sanger.ac.uk/search>) to determine protein domain architectures that could provide insight into their function (Finn et al. 2006) (data not shown). The expression pattern of these genes was evaluated in different cotton tissues using qPCR.

Analyses of Expression Patterns of Endogenous Genes in Cotton

The organs used from the cotton plants comprised flower buds, fruits, leaves, stems, branches and roots. We also included seven stages of flower development (flower buds of the following diameters: 2, 4, 6, 7, 8, 10 and 12 mm) and four

stages of fruit development (boll of the following diameters: 10–15, 16–20, 21–30 and >30 mm; the stages of flower and fruit and the respective major events of development are summarized in Artico et al. 2010). In addition, floral organs (sepal, petal, stamen, carpel and pedicel) from 6-mm flower buds were dissected and harvested (because the flower stage is preferred for feeding and oviposition of the cotton boll weevil). Material from at least five different cotton plants was harvested and pooled. Total RNA extractions were performed using Invisorb Spin Plant RNA Mini kit (Invitek), according to the manufacturer's protocol. For amplification of the corresponding genes, cDNAs were synthesized from 1 μ g total RNA using Oligo (dT24V) primers and Superscript III (Invitrogen, Carlsbad, CA), according to the manufacturer's instructions. Primers were designed using Primer 3 software (Rozen and Skaletsky 2002; see supplementary Table S2). Polymerase chain reactions were performed in the 7500 Fast Real-Time PCR detection system (Applied Biosystems, Foster City, CA) using SYBR[®]Green to monitor dsDNA synthesis. PCR efficiencies and optimal quantification cycle threshold or Cq values were estimated using the online Real time PCR Miner (Zhao and Fernald 2005). Two independent biological samples of each experimental condition were evaluated using triplicate technical replicates. The reference genes used to normalize the qPCR data have been discussed previously (Artico et al. 2010; see supplementary Table S2).

Isolation of Promoter Sequences of Flower-Specific Cotton Genes

Genomic DNA was extracted from cotton leaves using NucleoSpin Plant II (Macherey-Nagel, <http://www.mn-net.com>). Based on the EST sequences of the genes *GhPME-like1*, *Gh β Gal-like1* and *GhPL-like1*, reverse primer and reverse-nested primers were designed to isolate promoter sequences using the Genome Walker kit (Clontech, Palo Alto, CA), according to the manufacturer's instructions (see supplementary Table S3). Isolated fragments were cloned into the pGEM-T_Easy vector (Promega, Madison, WI), sequenced and assembled into promoter sequences of approximately 350, 780, 1639 bp in length for *GhPME-like1* and 383 bp and 514 bp in length for *Gh β Gal-like1* and *GhPL-like1*, respectively (see supplementary Fig. S2a,b,c). The transcription initiation site (TIS) was determined experimentally using the GeneRacer Kit (Invitrogen; see supplementary Fig. S2a–c and supplementary Table S3).

Sequence Analysis of Flower-Specific Promoters

The isolated promoter sequences of *GhPME-like1*, *Gh β Gal-like1* and *GhPL-like1* were scrutinized using the PlantCARE (Lescot et al. 2002), PLACE (Higo et al. 1999; Prestridge 1991)

and Plant-PAN (Chang et al. 2008) software packages to determine plant cis-acting regulatory DNA elements.

Construction of Expression Vectors Used for *Arabidopsis* and Cotton Transformation

The promoter regions were amplified and isolated by PCR using promoter-specific primers and Phusion High-Fidelity DNA Polymerase (Finnzymes, Espoo, Finland) from 100 ng gDNA (see supplementary Table S3). The purified PCR products were cloned into the pENTR™ Directional TOPO cloning entry vector (Invitrogen), and subsequently recombined into the pKGWFS7 destination vector bearing the GUS reporter gene, thus generating the recombinant destination clones p300_GhPME-like1::GUS; p760_GhPME-like1::GUS; p1618_GhPME-like1::GUS; pGhβGal-like1::GUS; and pGhPI-like1::GUS. All recombinant plasmids were transformed into *Agrobacterium tumefaciens* GV3101. These constructs were used to transform *Arabidopsis* (Ler) using the floral dip method (Clough and Bent 1998). Transgenic plants were selected on agar plates containing 50 mg/L kanamycin and the integration of promoter-GUS fusions to transgenic plants was confirmed by PCR using primers derived from promoter and GUS sequences (data not shown). More than ten independent lines were produced for each construct, and at least three different transgenic lines of T3 homozygous plants expressing each construct were employed for GUS analyses and quantification.

The promoter regions were tested in transient expression experiments by particle bombardment in 5- to 6-mm *G. hirsutum* flower buds. The preparation of tungsten microparticles was performed as described by Rech et al. (2008). Bombardment used the PDS-1000/He system (Bio-Rad, Hercules, CA) under gas helium pressure, set with 1200 psi He and 25 polHg vacuum. After bombardment, the biological material was assayed using GUS histochemical staining.

Histochemical and Fluorimetric GUS Assays

For the histochemical assays, plant tissues were stained for GUS activity based on the method of Jefferson (1987). For each construct, tissues were collected from at least five different plants. Samples were observed under a Leica Wild Heerbrugg M3Z Stereozoom Microscope (Leica, Wetzlar, Germany).

For quantitative fluorimetric GUS assays, inflorescences, flower buds, open flowers, siliques, stems, leaves and roots were used (see supplementary Table S4). Total protein was quantified according to Bradford (1976) using bovine serum albumin (BSA) as a standard. GUS specific activity was calculated as picomoles of 4-methylumbelliferone (4-MU) produced per minute per microgram of soluble protein. Data

are presented as the mean GUS activity from at least three independent determinations.

Real-Time PCR Analysis of *Arabidopsis* Transgenic Plants

Total RNA was extracted from inflorescences, flower buds, open flower, siliques, stem, leaves and roots of transgenic *Arabidopsis* plants expressing p300_GhPME-like1::GUS; p760_GhPME-like1::GUS; p1618_GhPME-like1::GUS; and pGhβGal-like1::GUS or pGhPI-like1::GUS and CaMV35S::GUS constructions as described by Onate-Sánchez and Vicente-Carbajosa (2008) (see supplementary Table S4). For each construct, three independent transgenic lines were chosen for qPCR analysis using three technical replicates. Expression levels of promoter-*uidA* were normalized to At1g13320 and At1g58050 *Arabidopsis* reference genes (Czechowski et al. 2005; see supplementary Table S2).

Microscopic Examination

To identify the cellular expression patterns, histological sections were obtained from *Arabidopsis* flowers expressing the constructions described above. For sectioning, flower samples processed for GUS histochemical analysis were fixed using 3 % (v/v) glutaraldehyde solution and then dehydrated in a graded ethanol series and embedded in acrylic resin (HISTORESIN Embedding Kit, Leica Microsystems Germany). Resin-embedded samples were sliced into 2- to 5-μm sections using an ultra-microtome (Micron HM 325 Microtime). Tissue sections were observed under a light microscope (Leica DMR).

Results

Identification of Genes with Specific Expression in Cotton Flowers

With the aim of identifying and isolating new tissue-specific promoters in cotton flowers, we first identified genes expressed predominantly in flowers. The search was performed using two databases: the PAVE and the Cotton Genome Database. The PAVE database is particularly informative and contains approximately 185,000 *Gossypium* EST sequences distributed among 30 cDNA libraries from a variety of tissues and organs and also under a range of conditions. The libraries are derived from allopolyploid cotton (*G. hirsutum*; A and D genomes) and its two diploid progenitors *Gossypium arboreum* (A genome) and *Gossypium raimondii* (D genome). Assembly of ESTs of the PAVE databank resulted in 22,030 contigs and 29,077 singletons (51,107 unigenes; Udall et al. 2006). We focused our analysis on the *G. raimondii* flower library. This library includes buds

from -3 dpa to bolls of +3 dpa (days post anthesis), and it contains 35,061 ESTs and 13,344 contigs, which are termed GR-Eb in the PAVE databank. We initially selected genes that were highly expressed with seven or more ESTs in this library. A total of 640 contigs (4.7 %) were identified as being highly expressed in floral tissues. Among this group of contigs, 19 (0.14 %) present ESTs in the flower library only. The sequences of the 19 genes from the flower library of *G. raimondii* were used to identify putative homologs in the *Arabidopsis* genome (listed in supplementary Table S1). The expression of putative *Arabidopsis* homologs was evaluated using the *Genevestigator* platform (Zimmermann et al. 2004; data not shown). *G. raimondii* genes were selected that had putative homologs in *Arabidopsis* of high and specific expression in floral tissue. The seven genes with high and specific expression in floral tissue selected from *G. raimondii* were used to search for *G. hirsutum* homologs in the Cotton Genome Database (listed in supplementary Table S1). The *G. hirsutum* genes were evaluated by qPCR analysis, and three genes showed high expression in cotton flowers (Fig. 1a). The amino acid sequences deduced to these genes show high sequence similarity to genes encoding enzymes involved in the biosynthesis and degradation of the cell wall. The genes encode a pectin methyl esterase (PME) (Family: PME1, Plant invertase/ pectin methylesterase, pfam04043), β -galactosidase (BGAL) (Family: Glycosyl hydrolase 35, pfam01301) and pectate lyase (PL) (Family: Pectate lyase, pfam04431). For clarity, we have termed these genes *GhPME-like1* (gi|193217860), *Gh β Gal-like1* (gi|45758291) and *GhPL-like1* (gi|68131791) (see supplementary Table S1). Phylogenetic analysis showed that the *G. hirsutum GhPME-like1* gene is grouped with the PME clade that is expressed predominantly in flower buds in *Arabidopsis* (Louvet et al. 2006; see supplementary Fig. S1a). Phylogenetic analysis indicated that the *G. hirsutum Gh β Gal-like1* gene was grouped together with the *Arabidopsis* β GAL flower-specific genes β GAL 14, β GAL 11 and β GAL 13 (see supplementary Fig. S1b; Ahn et al. 2007). Finally, phylogenetic analysis of the *GhPL-like1* gene with the PL family of *Arabidopsis* also suggested similarities in the proteins and in tissue expression activity. The *Arabidopsis* genes PLL 8 and 9, which belong to the same clade as *GhPL-like1*, show greater flower-specific expression (Palusa et al. 2007; see supplementary Fig. S1c).

The expression of *GhPME-like1*, *Gh β Gal-like1* and *GhPL-like1* was explored in detail in different cotton organs during flower and fruit development and also in different floral organs. The transcript levels of *GhPME-like1*, *Gh β Gal-like1* and *GhPL-like1* were much greater in flowers than in other plant organs. Moreover, their expression was almost undetectable in leaf, stem or branch tissues (Fig. 1a). *Gh β Gal-like1* is also expressed in cotton fruit, which is a feeding site of the boll weevil. Expression analysis during

fruit development showed that *Gh β Gal-like1* is expressed at a higher level at earlier stages during fruit development (10–15 and 16–20 mm) and decreases dramatically during the later stages of cotton fruit development (21–30 and >30 mm; Fig. 1b). We next evaluated the expression of *GhPME-like1*, *Gh β Gal-like1* and *GhPL-like1* during flower development. The results showed that transcript levels of the three genes were high in flower buds of 4–12 mm (Fig. 1c). In addition, to obtain a detailed portrait of the gene expression profile, we verified the expression level of *GhPME-like1*, *Gh β Gal-like1* and *GhPL-like1* genes in different floral verticils of 6-mm floral buds. *GhPME-like1*, *Gh β Gal-like1* and *GhPL-like1* showed an extremely high relative expression in stamens compared with other floral organs but they were also expressed in petals, sepals and carpels (Fig. 1d).

Isolation of the Promoter Sequences of *GhPME-like1*, *Gh β Gal-like1* and *GhPL-like1*

Because these genes show expression predominantly in floral tissues and/or fruit, we decided to isolate their regulatory regions and evaluate them as potential candidates for driving expression in reproductive tissue. The region immediately upstream of the translation initiation codon (ATG) of *GhPME-like1*, *Gh β Gal-like1* and *GhPL-like1* was isolated using the Universal Genome Walker kit (primers are listed in supplementary Table S3). The promoter sequences isolated, cloned and sequenced of *GhPME-like1* was 350 bp long, 383 bp for *Gh β Gal-like1* and 514 bp for *GhPL-like1*. We decided to evaluate the activity of these fragments in vivo prior to further isolations. This decision was based on the presence of several putative flower-specific regulatory motifs in these short fragments (see below). In addition, the use of short proximal promoters facilitates their manipulation and employment in plant biotechnology (Potenza et al. 2004; Odell et al. 1985; Fang et al. 1989). The transcription initiation sites (TIS) of the three genes studied were determined by cloning the full transcript using 5'-RACE PCR (see supplementary Fig. S2; primers are listed in supplementary Table S3). The promoters regions of *GhPME-like1*, *Gh β Gal-like1* and *GhPL-like1* were analyzed for the presence of putative regulatory motifs using the PLACE, PlantCARE and Plant-PAN databases. Various putative cis-acting regulatory elements were predicted as generic plants cis-elements, such as the TATA-box and CAAT-box (see supplementary Fig. S2). In addition, all promoters isolated present a large number of cis-elements related to pollen- and anther-specific expression. For example, the cis-element POLLEN1LELAT52 (sequence AGAAA) is repeated six times in the *GhPME-like1* gene promoter, four times in the *Gh β Gal-like1* gene promoter and three times in the *GhPL-like1* gene promoter (see supplementary Fig. S2a–c). POLLEN1LELAT52 activity correlates with pollen development in tomato (Bate and Twell 1998). The

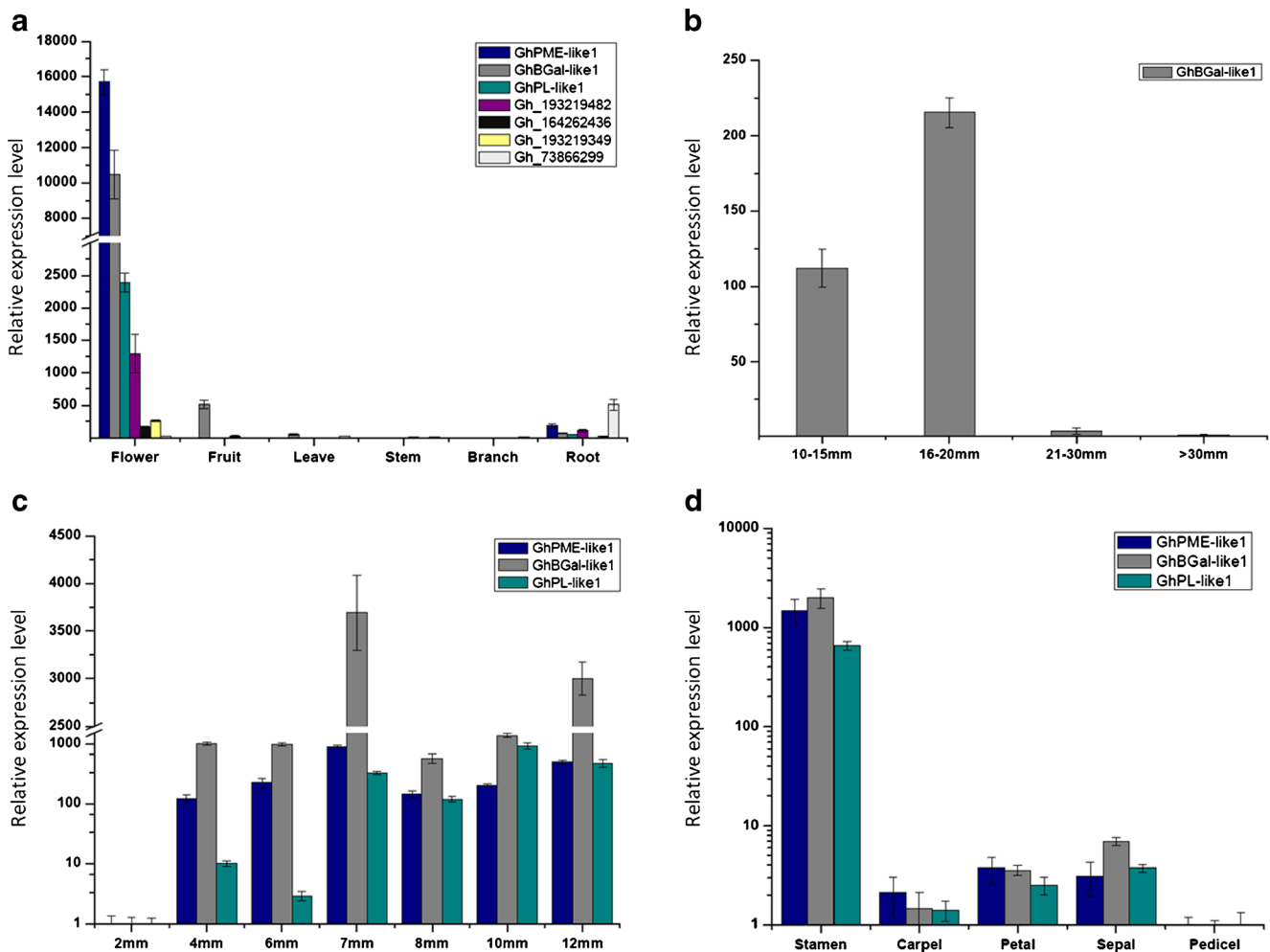


Fig. 1 **a** Relative mRNA level of seven putative flower-specific genes of *Gossypium hirsutum* in different plant organs. **b–d** Relative mRNA level of the three *GhPME-like1*, *GhβGal-like1* and *GhPL-like1* genes show high expression in cotton flowers during fruit (**b**) and flower (**c**) development and in different floral organs (**d**). Cq and amplification efficiency values were processed using qBase software. Normalization

was performed using the best combination of reference genes recommended by Artico et al. (2010). A combination of *GhUBQ14* and *GhPP2A1* genes was used as an internal control for plant organs (**a**), whereas *GhMZA* and *GhPTB* was used for fruit development (**b**), *GhACT4* and *GhUBQ14* for flower buds (**c**), and *GhACT4* and *GhFBX6* for floral organs (**d**)

GhβGal-like1 promoter also contains a fruit-specific element, TGTCACAMCUCUMISIN (sequence TGTCACA), identified in the promoter of the *cucumis* gene of melon (Yamagata et al. 2002) (see supplementary Fig. S2). The full list of cis-acting elements found in the cotton promoters may be viewed in supplementary Tables S5, S6 and S7.

Activity of Promoter:GUS Constructions in Various Tissues of Transgenic *Arabidopsis* Plants

To investigate promoter activity, regulatory sequences of *GhPME-like1* (300 bp), *GhβGal-like1* (356 bp) and *GhPL-like1* (435 bp) were fused to the *uidA* gene in binary vectors and used to transform *Arabidopsis*. All constructs, named pGhPME-like1::GUS; pGhβGal-like1::GUS; and pGhPL-like1::GUS, were transformed into *Arabidopsis* using *Agrobacterium*-mediated transformation. For each construct,

more than 12 homozygous plants with only one T-DNA insertion were evaluated. No differences were observed in the promoter activity of each construct within the transgenic lines evaluated. For *GhPME-like1*, the expression of GUS in flowers was robust and specific to the stamen of young flower buds (stage 11) when the stigmatic papillae are formed (Fig. 2a–g). Staining in the stamens remained intense during later stages of development of the flower bud (Fig. 2f–h). As shown in Fig. 2g and in detail in 2 h, j and l, GUS expression was intense in pollen grains and the inner layer of the anther wall in endothecium cells. Staining was also observed in the style and in the apical region of the stigma during later stages of development and until fruit development (Fig. 2g, h, j, l). The *GhPME-like1* promoter also shows residual activity in *Arabidopsis* seedlings and in the main root (see supplementary Fig. S3a–c). Histochemical staining for GUS from the pGhβGal-like1::GUS construct was robust in buds and open

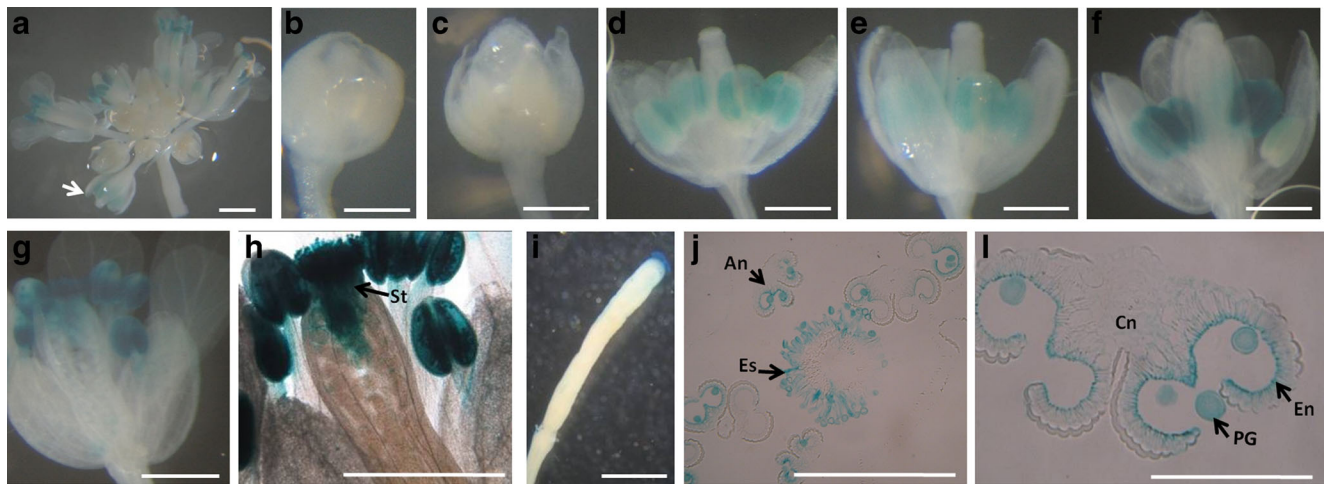


Fig. 2 a–l Histochemical analysis of β -glucuronidase (GUS) expression in flowers buds of transgenic *Arabidopsis* plants expressing the pGhPME-like1::GUS promoter. GUS expression was detected in stamens (d–h), in particular in anthers (An), pollen grains (PG) and endothecium cells (En) (j–l). Robust GUS staining was observed in the apical portion of the stigma

(Es), in stigmatic cells, the stylus (g–h, j) and in the apical portion of the siliques (i). j–l Cross-sections of the flower bud and anther; Cn endothecium tissues. a Arrow Stage 11 of flower development in *Arabidopsis*. b Stage 6 of flower development; c Stage 8; d Stage 11; e Stage 13; f Stage 13; (g, h, j, l) Stage 14, St stylus; (i) Stage 17a. Bar 250 μ m

flower during the later stages of flower development (stage 12; Fig. 3a–f). Strong GUS staining was observed primarily in anthers (Fig. 3d–h) and more specifically in pollen grains, connective tissue and endothecium cells (Fig. 3g, h, j). Staining was also observed in the upper portion of the carpels, in stigmatic cells and the style of the flower (Fig. 3f–g). Promoter activity was also observed in the apical region of the siliques (Fig. 3i). pGh β Gal-like1::GUS seedlings showed slight GUS staining in the cotyledon leaf margin (see supplementary Fig. S3g, h). pGhPL-like1 promoter activity was observed in *Arabidopsis* flowers (Fig. 4a–g), primarily in anthers, connective tissues, endothecium cells and pollen grains (Fig. 4c–h, l). Similar to the promoter construct pGh β Gal-like1::GUS, pGhPL-like1 directs expression of

GUS to stigmatic cells and the style (Fig. 4f, g) and siliques (Fig. 4i, j). GUS staining was also observed in the leaf primordium in *Arabidopsis* seedlings (see supplementary Fig. S3i, j). As expected from qPCR expression data of these genes in cotton tissues, no significant promoter activity was detected in vegetative organs of 5-week-old *Arabidopsis* transgenic plants carrying the constructs pGhPME-like1::GUS, pGh β Gal-like1::GUS; or pGhPL-like1::GUS (see supplementary Fig. S3d, e, f).

Quantitative Analysis of Promoter Activity

To compare activity levels of the cotton promoters in discrete *Arabidopsis* organs, the GUS activity of T3 *promoter:gus*

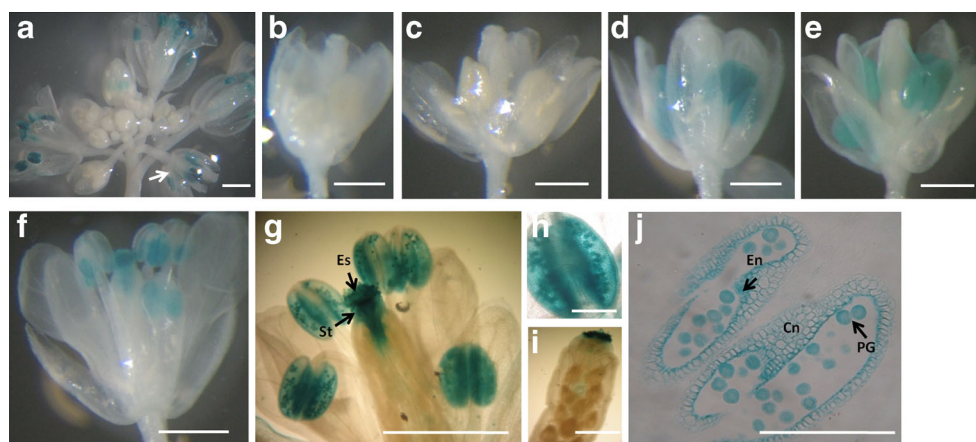


Fig. 3 a–j Histochemical analysis of GUS expression in flowers buds of transgenic *Arabidopsis* plants expressing the pGh β Gal-like1::GUS promoter. Robust GUS staining was observed in the anthers (a, d–h) and in particular in pollen grains (PG) (g, h, j) and endothecium cells (En) (j). Strong GUS staining was observed in the apical portion of the stigma

(Es), in stigmatic cells and the stylus (St) (f–g) and apical portion of the siliques (i). (j) Cross-section of the anther; Cn connective tissue. (a) Arrow Stage 12 of flower development in *Arabidopsis*. b, c Stage 8 of flower development; d Stage 12; e Stage 13; f, g, h, j Stage 14; i Stage 17a. Bar 250 μ m

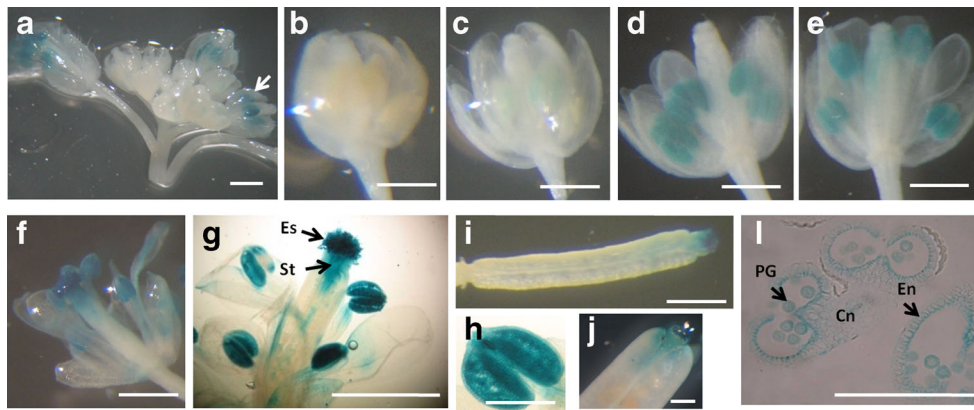


Fig. 4 a–l Histochemical analysis of GUS expression in flowers buds of transgenic *Arabidopsis* plants expressing the pGhPl-like1::GUS promoter. GUS is expressed in stamens of young flower buds (a, c) through to mature flower buds (f, g). GUS staining is restricted to anthers and primarily to endothecium cells (En) and pollen grains (PG) and connective tissues (Cn)

(h, l). GUS blue is expressed in stigmatic cells and the stylus (St) (f–g) and in the apical region of the siliques (i–j). l Cross-section of an anther. Es Stigma. a Arrow Stage 11 of flower development in *Arabidopsis*. b Stage 6 of flower development; c Stage 8; d Stage 11; e Stage 13; f, g, h, and l Stage 14, j Stage 17a. Bar 250 μm

transgenic lines was measured using quantitative fluorimetric analysis of the MUG substrate (Fig. 5). Transgenic plants carrying the promoter constructs were evaluated and compared with those carrying constitutive CaMV35S::GUS promoters and wild-type plants. GUS enzyme activity driven by *pGhPME-like1*, *pGhβGal-like1* and *pGhPl-like1* cotton promoters was determined from reproductive and vegetative organs of the transgenic lines. Quantitative GUS assays of the cotton promoters showed a high level of expression in reproductive tissues compared with vegetative tissues (Fig. 5a). In contrast, as expected, CaMV35S promoter activity was greater in roots, stems and leaves compared with the three cotton promoters (Fig. 5a). GUS activities in inflorescences, when compared with leaves of plants carrying the cotton promoters,

were at least 12.5-fold higher (Fig. 5a). When expression was compared with the 35S promoter, the *pGhPME-like1* promoter showed the highest GUS activity and was approximately 6-fold higher in flower buds.

This analysis showed that the promoter activity of the 300-bp fragment of the *GhPME-like1* (pGhPME-like1::GUS) promoter in flower buds, inflorescences and open flowers was greater than that of the *GhβGal-like1* and *GhPl-like1* promoters (Fig. 5a). Because longer fragments may display even stronger or more specific flower activity, we decided to isolate and evaluate the largest promoter of this gene. From the longer fragment isolated (1,639 bp), we generated two 5' deletions: p1618_GhPME-like1 and the p760_GhPME-like1. In addition to eight repeats of the cis-element POLLENILELAT52,

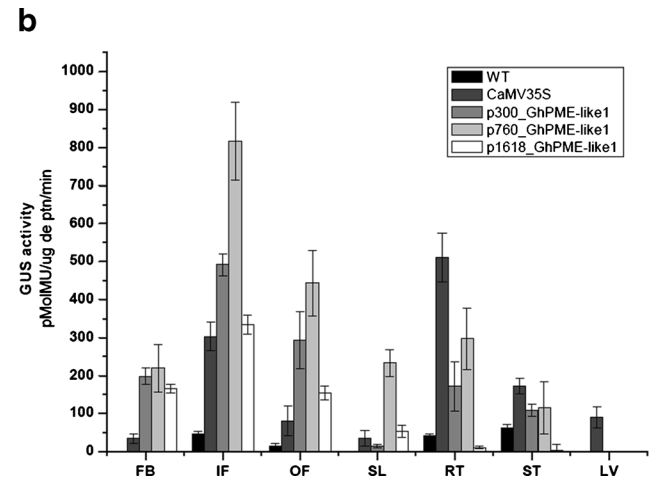
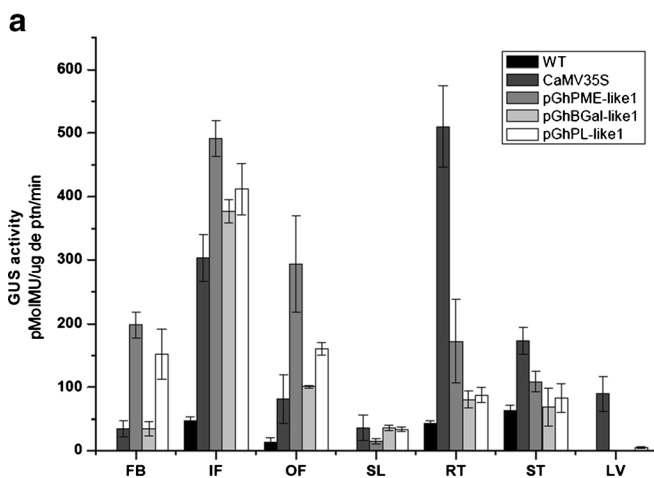


Fig. 5 Comparison of GUS expression via quantitative analysis of the MUG substrate. GUS activity was identified in flower buds (FB), inflorescences (IF), open flowers (OF), siliques (SL), roots (RT), stems (ST) and leaves (LV). The data represent GUS activity ± standard error of the mean (SEM) of three replicates from three independent transgenic lines. a Comparison of GUS expression levels driven by the *pGhPME-like1*,

pGhβGal-like1, *pGhPl-like1* and CaMV35S promoters in different tissues of transgenic *Arabidopsis* plants and wild-type plants via quantitative analysis of the MUG substrate. b Enzyme activity of GUS measured in various 5'-deleted cotton *GhPME-like1* promoters compared with transformed CaMV35S::GUS lines and wild-type plants

these two fragments carry the MYB.ph3 element (sequence AAACAAATA), which is a DNA binding element of a petal epidermis-specific MYB transcription factor, and the AP1-CARbox element (sequence CCAAA), which is present in the APETALA3 gene promoter. The full list of cis-acting elements found in this promoter is shown in supplementary Table S5. These fragments were fused to the GUS reporter gene and used to transform *Arabidopsis*. T3 homozygous seeds from 15–20 independent lines were generated for each construct, and the transgenic plants were analyzed for GUS expression using histochemical and quantitative assays as described in the **Materials and Methods**. The expression of GUS directed by the p760_GhPME-like1::GUS promoter was robust in flowers and particularly in anthers (pollen grains, endothecium cells) and filaments of young flower buds at stage 11 of flower development (Fig. 6a–d, e, o). The p760_GhPME-like1::GUS construct also directed expression of GUS to the vascular bundles of petals (Fig. 6b, c, f), stigmatic cells and the style in the apical portion of the carpel (Fig. 6b–d) and to the apical and basal portion (internodes) of the siliques (Fig. 6g). The 1,618-bp fragment of *GhPME-like1* promoter was capable of directing expression of GUS to petals, anthers (pollen grains, endothecium cells) and also the apical portion of siliques (Fig. 6j–n).

Results from GUS enzyme activity assays of three independent *Arabidopsis* T3 lines expressing each of the promoter constructs of *GhPME-like1* are shown in Fig. 5b and compared with the constitutive 35S promoter. Quantitative GUS assays of both *GhPME-like1* gene promoters showed a high level of expression in flower tissues, such as inflorescences, mature flowers, flower buds and also siliques. The highest

GUS activity was observed in the p760_GhPME-like1::GUS lines, which showed approximately 3- to 6-fold higher activity in reproductive organs compared the CaMV35S promoter, indicating that it may be a strong promoter specific to cotton flower tissues. GUS activity directed by the largest p1618_GhPME-like1 promoter was lower when compared with the other *GhPME-like1* promoter fragments but remained higher than that directed by the CaMV35S promoter in flower buds, inflorescences, mature flowers and siliques (Fig. 5b). The activity of the shorter promoter fragment (p300_GhPME-like1) was dramatically higher compared with the CaMV35S in reproductive tissues. Although GUS activity assays have been used successfully for years to characterize promoter activity, we decided to evaluate the activity of the cotton promoters using qPCR analyses due to its greater sensitivity (Yi et al. 2010; Liang et al. 2011; Koia et al. 2013).

Analysis of *uidA* Transcripts in Transgenic *Arabidopsis* Plants by qPCR

To compare the activity levels of cotton promoters in diverse *Arabidopsis* organs, *uidA* mRNA levels in T3 *promoter:gus* transgenic lines were measured using qPCR. For this purpose, total RNA was extracted from flower buds, inflorescences, open flowers, siliques, leaves, stems and roots of transgenic *Arabidopsis* plants. The levels of promoter-driven *uidA* transcripts were calculated for all lines with cotton promoter fragments fused to *uidA* and also from the CaMV35S::GUS transgenic line. In general, *uidA* mRNA levels driven by cotton promoters were found to be higher, although variable, in all flowers stages and siliques (Fig. 7a). Thus, all cotton

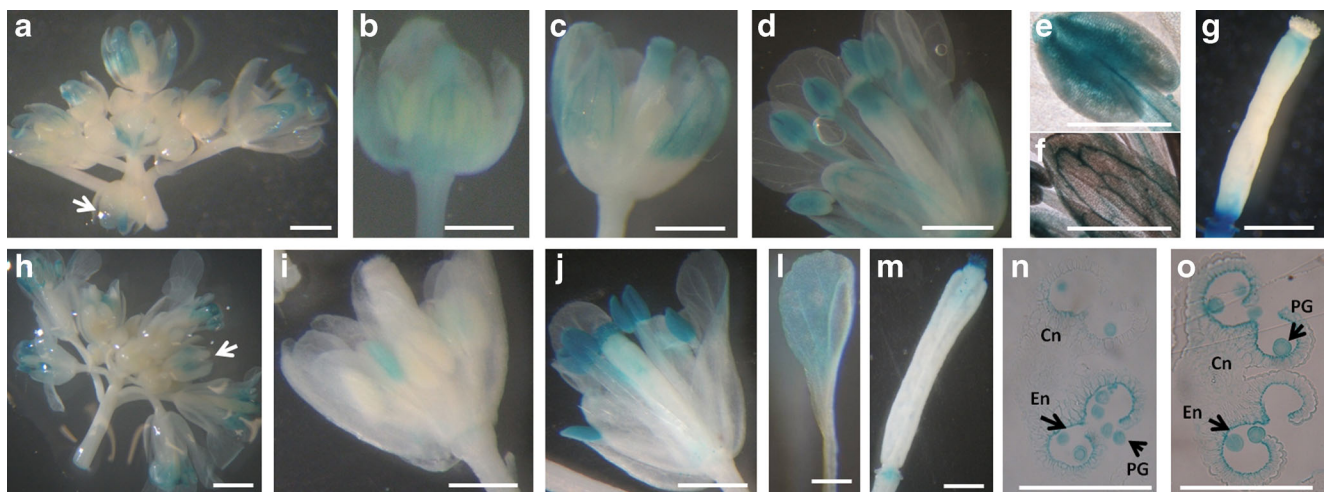


Fig. 6 a–o Histochemical analysis of GUS expression in flowers buds of transgenic *Arabidopsis* plants expressing the p760_GhPME-like1::GUS (a–g, n) and p1618_GhPME-like1::GUS (h–m, o) promoters. For the p760_PME-like1 promoter, GUS expression was detected in stamens and, particularly, in anthers and filaments (d, e) and in apical portions of the stigma (b–d), in petals (b–d, f) and apical and basal portions (internodes) of the siliques (g). Robust GUS

staining was observed directed by the p1618_PME-like1 construct in the apical portion of the stigma (j), in anthers (i–j, n), in petals (l) and siliques (m). n, o Cross-sections of anthers from plants expressing p760_GhPME-like1 and p1618_GhPME-like1 promoters, respectively. Arrow Stage 11 of flower development in *Arabidopsis* (a, h). b Stage 8 of flower development (c) Stage 9; d, e, f, n and j, l, o Stage 14; i Stage 11 and (g, m) Stage 17a. Bar 250 μ m

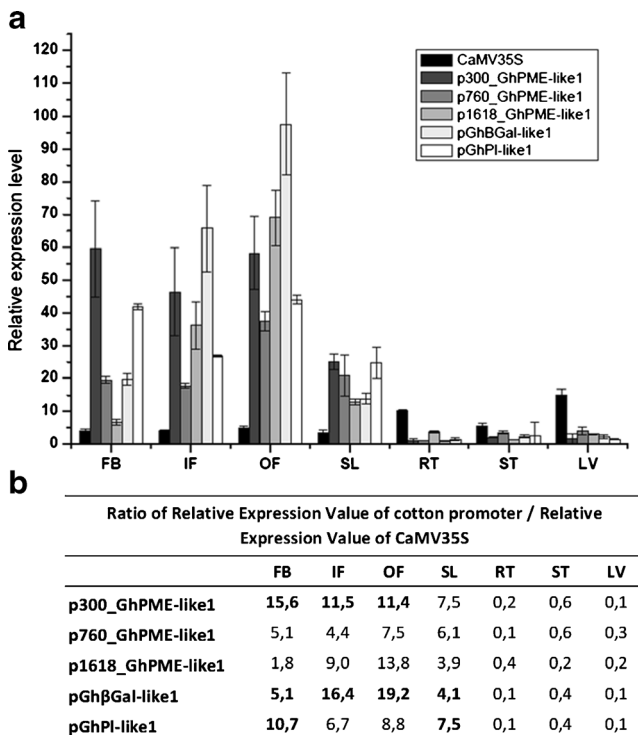


Fig. 7 Relative expression of the cotton promoter/relative expression of CaMV35S. **a** Relative expression of *uidA* transcripts, as measured by qPCR, of each cotton promoter:GUS transgenic plant compared with CaMV35S::GUS transgenic plants. Cq and amplification efficiency values were processed using qBase software. Levels of *uidA* mRNAs were determined in flower buds (FB), inflorescences (IF), open flowers (OF), siliques (SL), roots (RT), stems (ST) and leaves (LV). **b** The ratio of relative expression of cotton promoter/relative expression of the CaMV35S promoter

promoters were more specific and efficient than the CaMV35S in accumulating *uidA* transcripts in flower and fruit tissues. However, as expected, the constitutive promoter CaMV35S::GUS was capable of accumulating a greater amount of *uidA* transcripts in vegetative tissues such as the root, stem and leaves (Fig. 7a). For example, the relative expression of the pGhβGal-like1 promoter was at least 5-fold, 16-fold, 19-fold and 4-fold higher than that of the CaMV35S promoter in flower buds, inflorescences, open flowers and siliques, respectively (Fig. 7b). The relative expression of *uidA* transcripts directed by the pGhPI-like1 promoter was 10.7-fold higher in flower buds and 7.5-fold higher in siliques than the CaMV35S promoter (Fig. 7b). Taking into account only the three fragments, the *GhPME-like1* gene promoter, the smaller p300_GhPME-like1::GUS fragment accumulated greater levels of *uidA* transcript in flower tissues.

Functional Analysis of Promoter Regions Using Transient Expression in Flower Buds of *Gossypium hirsutum*

To evaluate the functionality and specificity of the promoter regions isolated in this study, expression cassettes containing

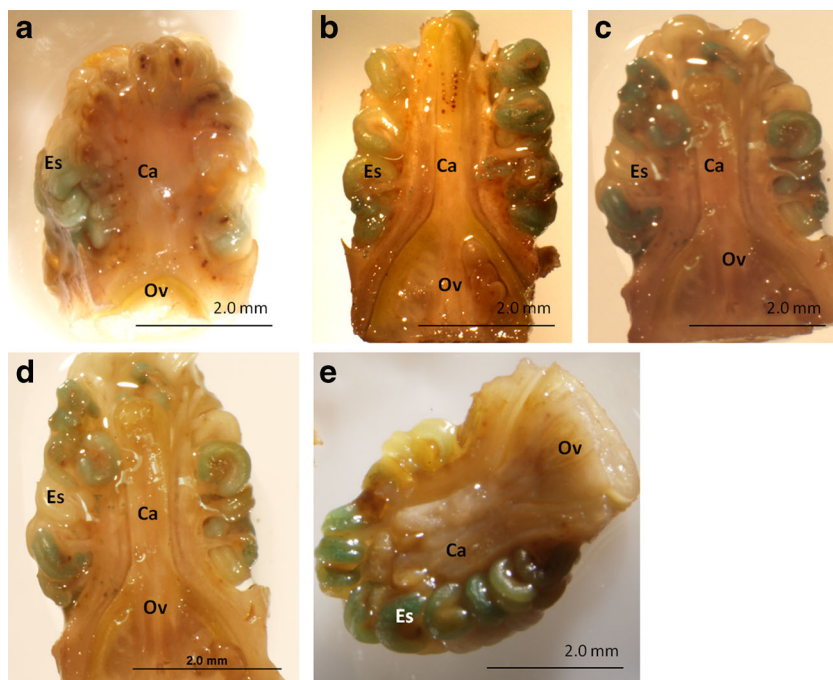
the five different promoters were used in transient expression, via biolistics, in 6-mm cotton flower buds (Fig. 8). All promoters directed GUS expression in stamens of cotton flower buds mirroring the expression pattern observed in qPCR experiments of cotton flower organs (Fig. 1d).

Discussion

Cotton is one of the most important crops cultivated worldwide although its production is hindered by several pests. Of these pests, the cotton boll weevil is the most important insect pest in Brazil—one of the top five producers and exporters—and this pest is considered limiting in other countries. The boll weevil larva and adult feed on reproductive organs including stamens, ovules and fruit. If this insect is not contained, the damage to cotton fiber production can reach 75 % (Martins et al. 2007b). Our major goal was to identify a promoter with high activity in reproductive tissues to fuse to genes encoding an insecticidal protein targeted against the insect-pest *A. grandis*. For this purpose, in our study, we evaluated the expression pattern of three ESTs derived from libraries of whole flowers in vegetative and reproductive tissues of cotton to identify genes expressed exclusively in *G. hirsutum* flowers, with the intention of isolating their promoter regions. These promoters, which were expressed predominantly in cotton flowers, may be useful for many applications in crop biotechnology where expression in reproductive tissues is required.

The genes selected in our search, which show accumulation of transcript in flower verticils and particularly in stamens, encode proteins from families that are highly expressed in the anther. A range of studies has shown that the pollen transcriptome is formed primarily by genes encoding hydrolytic enzymes that loosen and degrade the cell wall, such as pectate lyases, pectin methylesterase, polygalacturonase, glycosyl hydrolases and expansins (Becker et al. 2003; Haerizadeh et al. 2009). These genes are involved not only in the modification of the pollen tube cell wall of pollen grains but are also hydrolytic enzymes that may be important for penetration of flower stigmatic tissue (Jiang et al. 2005). The cotton *GhPME-like1* gene is highly expressed in stamens of flower buds primarily after the 6-mm diameter stage when unicellular microspores are found in anthers (Fig. 1). The *GhPME-like1* shows sequence similarity to the *Arabidopsis* pectin methylesterase genes (PME) *AT4G15980* and *AT5G49180*. These two genes show high and specific expression in the floral buds of *Arabidopsis* (Louvét et al. 2006; see supplementary Fig. S1a). Several studies suggest that PMEs are involved in the development and growth of the pollen tube and in the interaction with female tissues during fertilization (Jiang et al. 2005; Bosch et al. 2005; Tian et al. 2006). The expression pattern of cotton PME indicates that it may have a

Fig. 8 Transient expression assays, via biolistics, in *G. hirsutum* flower buds. Histochemical assays of 6-mm cotton flower buds that had been bombarded with the tissue-specific promoters p300_GhPME-like1 (a), p760_GhPME-like1 (b), p1618_GhPME-like1 (c), pGh β Gal-like1 (d) or pGhPL-like1 (e). *Es* Stamen, *Ca* carpel, *Ov* ovary



similar function to the putative *Arabidopsis* homologues. The *Gh β Gal-like1* gene sequence shows high similarity to the β -galactosidase cotton *GhGal1* expressed in flowers and fibers (Zhang and Liu 2005). The β -galactosidases (BGALs) belong to family 35 of glycosyl hydrolases that catalyze the hydrolysis of terminal b-galactosyl residues from carbohydrates, galactolipids, and glycoproteins. A previous study has shown that *GhGal1* mRNA is abundant in flowers and in 5- to 10-dpa fibers and declines in 20-dpa fibers; thus, it is involved in the early stages of development of cotton fiber (Zhang and Liu 2005). Our qPCR assays of the *Gh β Gal-like1* gene confirmed the temporal expression pattern in fruits (Fig. 1c). The *Gh β Gal-like1* gene sequence was also similar to BGAL sequences from *Arabidopsis* (BGAL 11, 13 and 14) that show high expression in mature pollen grains (Ahn et al. 2007; Hrubá et al. 2005; see supplementary Fig. S1b). In the case of the *GhPL-like1* gene, qPCR results showed that the expression of this gene was flower-specific, primarily in stamens during the later stages of flower development (Fig. 1). The sequence of this gene shows high similarity to pectate lyases-like (PLL)—enzymes that catalyze the eliminative cleavage of de-esterified pectin. This complex polysaccharide is a major component of the primary cell wall of many higher plants. There are 26 genes in the genome of *Arabidopsis* that encode pectate lyase (AtPLLs) proteins, the expression pattern of which vary considerably between different organs. Our phylogenetic analyses comparing the *GhPL-like1* gene sequence of *G. hirsutum* with PLL genes from *Arabidopsis* demonstrates a close relationship of *GhPL-like1* with PLL 8 and 9 genes. These two genes, together with PLL 10 and 11, are flower-specific and show high expression in pollen grains

(see supplementary Fig. S1c; Palusa et al. 2007). The putative flower-specific expression of the cotton genes *GhPME-like1*, *Gh β Gal-like1* and *GhPL-like1* and the previous reports on their putative *Arabidopsis* homologues encouraged us to isolate and characterize their respective promoter regions.

The cotton promoter fragments were qualitatively and quantitatively evaluated in *Arabidopsis*, the best heterologous system for dicotyledons. *A. thaliana* has been used extensively as a heterologous system to characterize promoter regions of dicot species (Koorneef and Meinke 2010; Zhao et al. 2012; Koia et al. 2013; Zhang et al. 2012). All cotton promoters evaluated in this study were found to show activity in various tissues and/or stages of flower and/or fruit development in *Arabidopsis*. Our data were obtained from transgenic lines grown from T₃ seeds, demonstrating that constitutive expression was stable between generations. The *pGhPME-like1*, *pGh β Gal-like1* and *pGhPL-like1* promoters direct significantly higher levels of GUS expression in reproductive tissues and, in particular, in stamens, pollen grains and the inner layer of the anther wall; in endothecium cells; and in stigmatic cells, style and siliques. The expression pattern of the *GhPME-like1*, *Gh β Gal-like1* and *GhPL-like1* genes in cotton agrees in principal with the regulatory pattern of their promoters because lower activity was observed in vegetative tissues and much higher activity was observed in flower buds, inflorescences and open flowers. The *pGhPME-like1*, *pGh β Gal-like1* and *pGhPL-like1* cotton promoters were compared with the uceA1.7 cotton promoter previously described by Viana et al. (2011), and they showed significantly higher activity in *Arabidopsis* reproductive organs (data not shown).

PLACE, PlantCARE and Plant-PAN databases identified several cis-elements involved in expression in reproductive tissues in the promoter regions of the *GhPME-like1*, *Gh β Gal-like1* and *GhPL-like1* genes that correlate with the expression pattern observed (see supplementary Fig. S2a, b and c). Various putative cis-acting regulatory elements required for pollen-specific expression or that are involved in enhancing gene expression in pollen were predicted. Among the putative cis regulatory elements found in the cotton promoters, the POLLEN1LELAT52 (AGAAA) cis-element was highly abundant. This element is repeated 14, 4 or 3 times in the promoter regions of the *GhPME-like1*, *Gh β Gal-like1* and *GhPL-like1* gene, respectively. In tomato, it is responsible for coordinating and specifically activating pollen grain development (Bate and Twell 1998). The POLLEN1LELAT52 cis regulatory element was characterized as being required for specific pollen transcription via a series of deletions of the promoter regions of the *lat52* gene (Bate and Twell 1998). There was also a significant number of copies of the cis-element GTGANTG10 (GTGA) in the cotton promoters. This cis-element is the binding site of transcription factors associated with late expression in pollen (Rogers et al. 2001). It is repeated five times, twice and once in *pGhPME-like1*, *pGh β Gal-like1* and *pGhPL-like1*, respectively. The presence of pollen/anther-specific cis-elements in the different promoter fragments *pGhPME-like1*, *pGh β Gal-like1* and *pGhPL-like1* suggest that they may be key factors in inducing strong anther and pollen promoter activity. Among the promoters evaluated, the *pGhPME-like1* (300 bp) promoter showed the highest number of putative anther/pollen cis-elements comprising six POLLEN1LELAT52 and two GTGANTG10. This unique characteristic may underlie the fact that the *pGhPME-like1* (300 bp) promoter showed the highest activity in reproductive tissues among the promoter fragments tested in this study, as demonstrated both by GUS activity analysis and the expression pattern of the *uidA* gene. Similar results were reported by Chen et al. (2012), who analyzed the *Triticum aestivum* *PSG076* gene promoter and observed that the 0.4-kb promoter fragment was capable of directing pollen-specific expression to a higher level than the longer promoter fragments. Similarly, Ni et al. (2008) observed that GUS activity levels and patterns directed by a 0.3-kb promoter fragment (P-300) were almost identical to that of a 1.3-kb promoter fragment (P-1301), indicating that P-300 contains all necessary elements for normal *AtTSG1* expression. Similar results were reported by Zhao et al. (2012), who analyzed a seed-abundant gene promoter from soybean. The two longer promoter fragments of the *GhPME-like1* gene (*p760_GhPME-like1* and *p1618_GhPME-like1*) showed lower *uidA* expression in flower buds and inflorescences when compared with *pGhPME-like1* (300 bp; Fig. 7a). In addition, the longer fragments were also active in other tissues, such as petals (Fig. 6). This petal expression may result from the activity

of additional cis-elements present in the 760-bp and 1,618-bp fragments, such as the MYB.ph3 element (870 bp) and AP1-CARGBox (1399, 1237, 875 and 741 bp). The *Gh β Gal-like1* promoter also directed the highest level of *uidA* transcripts primarily in inflorescences and open flowers (Fig. 7a). Therefore, the *pGhPME-like1* (300 bp) and *pGh β Gal-like1* (356 bp) promoters may be the best choices when high expression is required in cotton reproductive organs. The *uidA* expression pattern driven by the *GhPL-like1* cotton promoter was similar to that observed by Sun and Nocker (2010) when they analyzed the promoters of the PLL8 and PLL9 genes in *Arabidopsis*. Therefore, in addition to their functional conservation, the expression of these genes may be regulated similarly. Moreover, transient expression analysis via biolistics in 6-mm cotton flower buds confirmed that *GhPME-like1*, *Gh β Gal-like1* and *GhPL-like1* promoters directed GUS expression primarily to stamens. Therefore, these promoters may be effective targeting sequences to increase foreign protein accumulation in transgenic cotton plants and for effective tissue-type-dependent expression.

Arabidopsis transgenic plants harboring cotton promoter constructs were compared with plants expressing the CaMV35S promoter—a well characterized constitutive promoter conferring strong transgene expression in dicot species. All cotton promoters tested showed higher activity and were more specific than CaMV35S in accumulating *uidA* transcripts in flower and fruit tissues (Figs. 5a, 7a). Similar results were reported by Dalal et al. (2010), who analyzed the constitutive CaMV35S promoter in tomato tissues and observed that this promoter drove lower GUS activity in flowers than in leaves. As expected, the CaMV35S promoter showed higher activity in vegetative tissues. Moreover, with regard to the lower expression in reproductive tissues compared with cotton promoters identified in this study, it has been shown that the commercial *Bt*-cotton hybrids in India that use CaMV35S express less than the critical levels of Cry1Ac required for full protection against the bollworm *H. armigera* (Kranthi et al. 2005). The activity of CaMV35S decreases during cotton plant development, particularly in some plant organs, including the boll rind, square bract, bud and flower, which is the main feeding site for bollworm larvae (Kranthi et al. 2005). Therefore, the CaMV35S promoter has proved a poor choice for *Bt*-transgenic technology for efficient and sustainable protection from pests that attack reproductive tissue in cotton (Kranthi et al. 2005). Taking this information into account, the novel promoters that are expressed predominantly in flower and/or fruit characterized in this study may be key tools for enhancing expression of toxin genes in reproductive tissues.

In summary, five different fragments of promoters of three different cotton genes were assessed in transgenic *Arabidopsis* plants throughout various stages of reproductive and vegetative tissue development. These promoters were found to be active in the *Arabidopsis* plants, including in the

flower, stamens and pollen grains, and showed distinct activity patterns. The activity of the promoters was also confirmed by transient expression assays in flower buds of *G. hirsutum*. Due to their high activity in reproductive tissues and their short size, the novel cotton promoters identified by our study may represent an important tool in the generation of crop plants resistant to insect pests such as the cotton boll weevil and bollworm, which damage plants during the reproductive stage.

Acknowledgments We thank the researchers, Fernando Campos de Assis Fonseca and Raquel Sampaio, from Embrapa Genetic Resources and Biotechnology technical assistance with biolistic experiment in cotton. This work was part of S.A.'s PhD research in Genetics, at the Department of Genetics of the Universidade Federal do Rio de Janeiro (UFRJ), and was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq; M. Alves-Ferreira: # 306025/2010-8) and Fundação de Amparo à Pesquisa do Rio de Janeiro (FAPERJ).

References

- Ahn YO, Zheng M, Bevan DR, Esen A, Shiu SH, Peng HP, Miller JT, Cheng LI, Poulton JE, Shih MC (2007) Functional genomics analysis of *Arabidopsis thaliana* glycoside hydrolase family 35. *Phytochemistry* 11:1510–1520
- Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402
- Artico S, Nardeli SM, Oliveira-Neto OB, Grossi-de-Sa MF, Alves-Ferreira M (2010) Identification and evaluation of new reference genes in *Gossypium hirsutum* for accurate normalization of real-time quantitative RT-PCR data. *BMC Plant Biol* 10:12–22
- Bate N, Twell D (1998) Functional architecture of a late pollen promoter: pollen-specific transcription is developmentally regulated by multiple stage-specific and co-dependent activator elements. *Plant Mol Biol* 37:859–869
- Baum JA, Bogaert T, Clinton W, Heck GR, Feldmann P, Ilagan O, Johnson S, Plaetinck G, Munyikwa T, Pleau M (2007) Control of coleopteran insect pests through RNA interference. *Nat Biotechnol* 25:1322–1326
- Becker JD, Boavida LC, Carneiro J, Haury M, Feijo JA (2003) Transcriptional profiling of *Arabidopsis* tissues reveals the unique characteristics of the pollen transcriptome. *Plant Physiol* 133:713–725
- Bosch M, Cheung AY, Hepler PK (2005) Pectin methylesterase, a regulator of pollen tube growth. *Plant Physiol* 138:1334–1346
- Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248–254
- Chang WC, Lee TY, Huang HD, Huang HY, Pan RL (2008) PlantPAN: plant promoter analysis navigator, for identifying combinatorial cis-regulatory elements with distance constraint in plant gene groups. *BMC Genomics* 9:561
- Chen L, Miao Y, Wang C, Su P, Li T, Wang R, Hao X, Yang G, He G, Gao C (2012) Characterization of a novel pollen-specific promoter from wheat (*Triticum Aestivum* L.). *Plant Mol Biol Rep* 30:1426–1432
- Chitkowski RL, Turnipseed SG, Sullivan MJ, Bridges WC (2003) Field and laboratory evaluations of transgenic cottons expressing one or two *Bacillus thuringiensis* var. *kurstaki* Berliner proteins for management of noctuid (Lepidoptera) pests. *J Econ Entomol* 96:755–762
- Christensen AH, Quail PH (1996) Ubiquitin promoter based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Res* 5:213–218.
- Clough SJ, Bent AF (1998) Floral dip: a simplified method for agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J* 16:735–743
- Czechowski TSM, Altmann T, Udvardi MK, Scheible WR (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. *Plant Physiol* 139:5–17
- Dalal M, Chinnusamy V, Bansal KC (2010) Isolation and functional characterization of lycopene β -cyclase (CYC-B) promoter from *Solanum habrochaites*. *BMC Plant Biol* 10:61
- Desfeux C, Clough SJ, Bent AF (2000) Female reproductive tissues are the primary target of agrobacterium-mediated transformation by the arabidopsis floral-dip method. *Plant Physiol* 123:895–904
- Fang RX, Nagy F, Sivasubramaniam S, Chua NH (1989) Multiple cis regulatory elements for maximal expression of the cauliflower mosaic virus 35S promoter in transgenic plants. *Plant Cell* 1:141–150
- Finn RD, Mistry J, Schuster-Bockler B, Griffiths-Jones S, Hollich V (2006) Pfam: clans, web tools and services. *Nucleic Acids Res* 34:D247–D251
- Greenberg SM, Sappington TW, Setamou M, Coleman RJ (2003) Influence of different cotton fruit sizes on boll weevil (Coleoptera: Curculionidae) oviposition and survival to adulthood. *Environ Entomol* 33:443–449
- Grossi-de-Sa MF, Magalhães MQ, Silva MS, Silva SMB, Dias SC, Nakasu EYT, Brunetta PSF, Oliveira GR, Oliveira Neto OB, Oliveira RS, Soares LHB, Ayub MAZ, Siqueira HAA, Figueira ELZ (2007) Susceptibility of *Anthonomus grandis* (cotton boll weevil) and *Spodoptera frugiperda* (fall armyworm) to a CryIIa-type toxin from a Brazilian *Bacillus thuringiensis* strain J. *J Biochem Mol Biol* 40:773–782
- Guilley H, Dudley RK, Jonard G, Balazs E, Richards KE (1982) Transcription of cauliflower mosaic virus DNA: detection of promoter sequences, and characterization of transcripts. *Cell* 30:763–770
- Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 3:307–321
- Haerizadeh F, Wong CE, Bhalla PL, Gresshoff PM, Singh MB (2009) Genomic expression profiling of mature soybean (*Glycine max*) pollen. *BMC Plant Biol* 9:25
- Haynes JW, Smith JW (1992) Longevity of laboratory-reared boll weevils (Coleoptera: Curculionidae) offered honey bee-collected pollen and plants unrelated to cotton. *J Entomol Sci* 27:366–374
- Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Res* 27:297–300
- Hrubá P, Honys D, Twell D, Capková V, Tupý J (2005) Expression of β -galactosidase and β -xylosidase genes during microspore and pollen development. *Planta* 220:931–940
- Jefferson R (1987) Assaying chimeric genes in plants: the GUS gene fusion system. *Plant Mol Biol Report* 5:387–405
- Jiang L, Yang S, Xie LF, Puaah CS, Zhang XQ, Yang WC, Sundaresan VYED (2005) *VANGUARD1* encodes a pectin methylesterase that enhances pollen tube growth in the *Arabidopsis* style and transmitting tract. *Plant Cell* 17:584–596
- Koia J, Moyle R, Hendry C, Lim L, Botella JR (2013) Pineapple translation factor SUI1 and ribosomal protein L36 promoters drive constitutive transgene expression patterns in *Arabidopsis thaliana*. *Plant Mol Biol* 81:327–336
- Koomeef M, Meinke D (2010) The development of *Arabidopsis* as a model plant. *Plant J* 61:909–921

- Kranthi KR, Naidu S, Dhawad CS, Tatwawadi A, Mate K, Patil E, Bharose AA, Behere GT, Wadaskar RM, Kranthi S (2005) Temporal and intra-plant variability of Cry1Ac expression in Bt-cotton and its influence on the survival of the cotton bollworm, *Helicoverpa armigera* (Hubner) (Noctuidae : Lepidoptera). *Curr Sci* 89:291–298
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DJ (2007) Clustal W and clustal X version 2.0. *Bioinformatics* 23:2947–2948
- Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouze P, Rombauts S (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res* 30:325–327
- Liang YS, Jeon YA, Lim SH, Kim JK, Lee JY, Kim YM, Lee YH, Ha SH (2011) Vascular-specific activity of the *Arabidopsis* carotenoid cleavage dioxygenase 7 gene promoter. *Plant Cell Rep* 6:973–980
- Louvet R, Cavel E, Gutierrez L, Guénin S, Roger D, Gillet F, Guéneva F, Pelloux J (2006) Comprehensive expression profiling of the pectin methyltransferase gene family during silique development in *Arabidopsis thaliana*. *Planta* 224:782–791
- Mao YB, Cai WJ, Wang JW, Hong GJ, Tao XY, Wang LJ, Huang YP, Chen XY (2007) Silencing a cotton bollworm P450 monooxygenase gene by plant-mediated RNAi impairs larval tolerance of gossypol. *Nat Biotechnol* 25:1307–1313
- Martins WFS, Ayres CFJ, Lucena WA (2007a) Genetic diversity of Brazilian natural populations of *Anthonomus grandis* Boheman (Coleoptera: Curculionidae), the major cottonpest in the New World. *Genet Mol Res* 6:23–32
- Martins ES, Praça LB, Dumas VF, Silva-Werneck JO, Sone EH, Waga IC, Berry C, Monnerat RG (2007b) Characterization of *Bacillus thuringiensis* isolates toxic to cotton boll weevil (*Anthonomus grandis*). *Biol Control* 40:65–68
- Ni SM, Meng LJ, Zhao J, Wang XC, Chen J (2008) Isolation and characterization of the trichome-specific AtTSG1 promoter from *Arabidopsis thaliana*. *Plant Mol Biol Rep* 26:263–276. doi:10.1007/s11105-008-0036-5
- Odell JT, Nagy F, Chua NH (1985) Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. *Nature* 313:810–812
- Olsen KM, Daly JC, Finnegan EJ, Mahon RJ (2005) Changes in *Cry1Ac* Bt transgenic cotton in response to two environmental factors: temperature and insect damage. *J Econ Entomol* 98:1382–1390
- Onate-Sánchez L, Vicente-Carbajosa J (2008) DNA-free RNA isolation protocols for *Arabidopsis thaliana*, including seeds and siliques. *BMC Res Notes* 1:840–860
- Palusa SG, Golovkin M, Shin SB, Richardson DN, Reddy ASN (2007) Organ-specific, developmental, hormonal and stress regulation of expression of putative pectate lyase genes in *Arabidopsis*. *New Phytol* 174:537–550
- Potenza C, Aleman L, Gopalan CS (2004) Targeting transgene expression in research, agricultural, and environmental applications: promoters used in plant transformation. *In Vitro Cell Dev Biol Plant* 40:1–22
- Prestridge DS (1991) SIGNAL SCAN: a computer program that scans DNA sequences for eukaryotic transcriptional elements. *Comput Appl Biosci* 7:203–206
- Rech EL, Vianna GR, Aragao FL (2008) High-efficiency transformation by biolistics of soybean, common bean and cotton transgenic plants. *Nat Protoc* 3:410
- Rocha PSCF, Sheikh M, Melchiorre R, Fagard M, Boutet S, Loach R, Moffatt B, Wagner C, Vaucheret H, Furner I (2005) The *Arabidopsis* homology-dependent gene silencing gene codes for an S-adenosyl-L-homocysteine hydrolase required for DNA methylation-dependent gene silencing. *Plant Cell* 17:404–417
- Rogers HJ, Bate N, Combe J, Sullivan J, Sweetman J, Swan C, Lonsdale DM, Twell D (2001) Functional analysis of cis-regulatory elements within the promoter of the tobacco late pollen gene *g10*. *Plant Mol Biol* 45:577–585
- Rozen S and Skaletsky HJ (2002) Primer3 on the www for general users and for biologist programmers. In: Misener S, Krawetz SA (eds) *Bioinformatics methods and protocols: methods in molecular biology*. Humana, Totowa, NJ, pp 365–386. Source code available at <http://fokker.wi.mit.edu/primer3/>
- Saitou N, Nei M (1987) The neighbor-joining method—a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425
- Sanahuja G, Banakar R, Twyman RM, Capell T, Christou P (2011) *Bacillus thuringiensis*: a century of research, development and commercial applications. *Plant Biotechnol J* 9:283–300
- Schenk PM, Remans T, Sagi L, Elliott AR, Dietzgen RG, Swennen R, Ebert PR, Grof CPL, Manners JM (2001) Promoters for pregenomic RNA of banana streak badnavirus are active for transgene expression in monocot and dicot plants. *Plant Mol Biol* 47:399–412
- Sun L, Nocker SV (2010) Analysis of promoter activity of members of the PECTATE LYASE-LIKE (PLL) gene family in cell separation in *Arabidopsis*. *BMC Plant Biol* 1:152
- Sunilkumar G, Mohr L, Lopata-Finch E, Emami C, Rathore KS (2002) Developmental and tissue-specific expression of CaMV 35S promoter in cotton as revealed by GFP. *Plant Mol Biol* 50:463–474
- Tamura K, Peterson D, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 10:2731–2739
- Tian GW, Chen MH, Zaltsman A, Citovsky V (2006) Pollen-specific pectin methyltransferase involved in pollen tube growth. *Dev Biol* 294:83–91
- Udall JA, Swanson JM, Haller K, Rapp RA, Sparks ME, Hatfield J, Yu YS, Wu YR, Dowd C, Arpat AB, Sickler BA, Wilkins TA, Guo JY, Chen XY, Scheffler J, Taliencio E, Turley R, Mcfadden H, Payton P, Klueva N, Allen R, Zhang DS, Haigler C, Wilkerson C, Suo JF, Schulze SR, Pierce ML, Essenberg M, Kim H, Llewellyn DJ, Dennis ES, Kudrna D, Wing R, Paterson AH, Soderlund C, Wendel JF (2006) A global assembly of cotton ESTs. *Genome Res* 16:441–450
- Viana AAB, Fragoso RR, Guimarães LM, Pontes N, Oliveira-Neto OB, Artico S, Nardeli SM, Alves-Ferreira M, Batista JN, Silva MCM, Grossi-de-Sa MF (2011) Isolation and functional characterization of a cotton ubiquitination-related promoter and 5'UTR that drives high levels of expression in root and flower tissues. *BMC Biotechnol* 11:115
- Whelan S, Goldman N (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol* 5:691–699
- Yamagata H, Yonesu K, Hirata A, Aizono Y (2002) TGTCACA motif is a novel cis-regulatory enhancer element involved in fruit-specific expression of the *cucumis* gene. *J Biol Chem* 277:11582–11590
- Yi N, Kim YS, Jeong MH, Oh SJ, Jeong JS, Park SH, Jung H, Choi YD, Kim JK (2010) Functional analysis of six drought-inducible promoters in transgenic rice plant throughout all stages of plant growth. *Planta* 232:743–754
- Zhang HM, Liu JY (2005) Molecular cloning characterization of a β -Galactosidase gene expression preferentially in cotton fibers. *J Integr Plant Biol* 47:223–232
- Zhang W, McElroy D, Wu R (1991) Analysis of rice *Act1* 5' region activity in transgenic rice plants. *Plant Cell* 3:1155–1165
- Zhang X, Hu Y, Jiang C, Zhang W, Li Z, Ming F (2012) Isolation of the Chinese rose sHSP gene promoter and its differential regulation analysis in transgenic *Arabidopsis* plants. *Mol Biol Rep* 39:1145–1151
- Zhao S, Fernald RD (2005) Comprehensive algorithm for quantitative real-time polymerase chain reaction. *J Comput Biol* 12:1047–1064
- Zhao Y, Shao S, Li X, Zhai Y, Zhang Q, Qian D, Wang Q (2012) Isolation and activity analysis of a seed-abundant *soyAPI* gene promoter from soybean. *Plant Mol Biol Rep* 30:1400–1407
- Zimmermann P, Hirsch-Hoffmann M, Hennig L, Gruissem W (2004) Genevestigator *Arabidopsis* microarray database and analysis toolbox. *Plant Physiol* 136:2621–2632