Planta (2020) 251:56
https://doi.org/10.1007/s00425-020-03348-8

ORIGINAL ARTICLE

Insights obtained using different modules of the cotton *uceA1.7* promoter

Marcos Fernando Basso[1] · Isabela Tristan Lourenço-Tessutti[1] · Carlos Busanello[2] · Clidia Eduarda Moreira Pinto[1,3] · Elínea de Oliveira Freitas[1,3] · Thuanne Pires Ribeiro[1] · Janice de Almeida Engler[4] · Antonio Costa de Oliveira[2] · Carolina Vianna Morgante[1,5] · Marcio Alves-Ferreira[6] · Maria Fatima Grossi-de-Sa[1,7]

Received: 28 October 2019 / Accepted: 22 January 2020 / Published online: 31 January 2020

Abstract

Main conclusion **The structure of the cotton *uceA1.7* promoter and its modules was analyzed; the potential of their key sequences has been confirmed in different tissues, proving to be a good candidate for the development of new biotechnological tools.**

Abstract Transcriptional promoters are among the primary genetic engineering elements used to control genes of interest (GOIs) associated with agronomic traits. Cotton *uceA1.7* was previously characterized as a constitutive promoter with activity higher than that of the constitutive promoter from the *Cauliflower mosaic virus* (CaMV) 35S gene in various plant tissues. In this study, we generated *Arabidopsis thaliana* homozygous events stably overexpressing the *gfp* reporter gene driven by different modules of the *uceA1.7* promoter. The expression level of the reporter gene in different plant tissues and the transcriptional stability of these modules was determined compared to its full-length promoter and the 35S promoter. The full-length *uceA1.7* promoter exhibited higher activity in different plant tissues compared to the 35S promoter. Two modules of the promoter produced a low and unstable transcription level compared to the other promoters. The other two modules rich in *cis*-regulatory elements showed similar activity levels to full-length *uceA1.7* and 35S promoters but were less stable. This result suggests the location of a minimal portion of the promoter that is required to initiate transcription properly (the core promoter). Additionally, the full-length *uceA1.7* promoter containing the 5′-untranslated region (UTR) is essential for higher transcriptional stability in various plant tissues. These findings confirm the potential use of the full-length *uceA1.7* promoter for the development of new biotechnological tools (NBTs) to achieve higher expression levels of GOIs in, for example, the root or flower bud for the efficient control of phytonematodes and pest-insects, respectively, in important crops.

Keywords Cotton constitutive promoter · Gene expression · New biotechnological tools · Transcriptional core promoter · Transgenic crops

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00425-020-03348-8) contains supplementary material, which is available to authorized users.

✉ Marcos Fernando Basso
marcosbiotec@gmail.com

✉ Maria Fatima Grossi-de-Sa
fatima.grossi@embrapa.br

[1] Embrapa Genetic Resources and Biotechnology, PqEB Final, W5 Norte, PO Box 02372, Brasília, DF 70770-901, Brazil

[2] Federal University of Pelotas, Capão Do Leão, RS 96160-000, Brazil

[3] Federal University of Brasília, Brasília, DF 70910-900, Brazil

[4] UMR Institut Sophia Agrobiotech INRA/CNRS/UNS, Sophia Antipolis, France

[5] Embrapa Semi Arid, Petrolina, PE 56302-970, Brazil

[6] Federal University of Rio de Janeiro, Rio de Janeiro, RJ 21941-901, Brazil

[7] Catholic University of Brasília, Brasília, DF 71966-700, Brazil

Abbreviations

GOIs Genes of interest
NBTs New biotechnological tools

CaMV *Cauliflower mosaic virus*
UTR Untranslated region
TSS Transcriptional start site

Introduction

Plants are constantly challenged by different and concomitant types of abiotic and biotic stresses (Crisp et al. 2016). Crops with new and superior characteristics are intensely demanded by agricultural producers and plant breeders worldwide to overcome these drawbacks (Hickey et al. 2019). Genetic engineering tools have contributed to the development of these elite cultivars and some of them have already been made commercially available (Basso et al. 2019). Several genes of interest (GOIs) have already been characterized in proof-of-concept studies and have been associated with desirable agronomic traits, such as water-deficit tolerance, salinity, heavy metals, flooding, or resistance to nematodes, pest-insects, bacteria, fungi, and viruses (Limera et al. 2017). For the development of new biotechnological tools (NBTs) using these GOIs, transcriptional promoter sequences that stably and adequately control the expression of these genes are highly important (Limera et al. 2017). Promoters that drive high levels of constitutive, stress-induced, organ-specific, or triggered expression in a particular plant growth phase may be required in NBTs to maintain or improve the GOI features or plant phenotype (Lu et al. 2008a).

The constitutive promoter from the *Cauliflower mosaic virus* (CaMV) 35S gene is currently the most commonly used promoter in plant engineering (Odell et al. 1985). Both the wild-type 35S promoter and its enhanced version can drive high transcription levels in dicotyledonous but show reduced and unstable activity in monocots (Benfey et al. 1990; Weeks et al. 1993; Gupta et al. 2001). In contrast, monocot-derived promoters usually have their activity potentiated only in monocot plants, whereas dicot-derived promoters have lower activity in monocots (Park et al. 2010). Ubiquitin-promoter sequences are efficient to drive constitutively high expression levels of transgenes in both monocot and dicot plants. The *Ubi1* from maize (Christensen et al. 1992), *Ubi1* and *3* from rice (Wang and Oard 2003; Lu et al. 2008b), *Ubq10* from *A. thaliana* (Grefen et al. 2010), *uceA1.7* from cotton (Viana et al. 2011), *Ubi4* and *9* from sugarcane (Wei et al. 2003), *GmUbi* from soybean (Hernandez-Garcia et al. 2009), and *Ubi7* from potato (Garbarino et al. 1995) are some species-specific examples. Although there are several plant and virus promoters functionally characterized as constitutive, few of these promoters are validated in other crops or characterized for the transcription level in different tissues or stages of development of the transgenic plant. Additionally, gene stacking in a single transgenic plant requires a greater number of different promoters with high activity to avoid homology-dependent gene silencing when using multiple copies of the same promoter (Park et al. 2010). Furthermore, promoters with higher activity than those already available are required to obtain a greater accumulation of transcripts or proteins in specific plant organs (e.g., entomotoxic protein in cotton flower bud to control the cotton boll weevil) (Ribeiro et al. 2017).

Plant ubiquitin-conjugating (E2) enzymes have a conserved ubiquitin-conjugating domain responsible for modulating the post-transcriptional degradation of target proteins (Moon et al. 2004). In addition, E2 enzymes are also essential for plant immunity (Zhou et al. 2017). The E2 enzyme-related coding gene (Gohir.A11G023700) from cotton (*Gossypium hirsutum*) is highly expressed in different organs and tissues. The promoter sequence of this gene (named *uceA1.7*) was isolated and characterized as a constitutive promoter with activity higher than CaMV 35S in different tissues of transgenic *Arabidopsis thaliana*, such as root, stem, leaf, and flower bud (Viana et al. 2011). In addition, the long 5′-UTR of the gene, which contains an intron sequence, was considered essential to the high expression level of the *gus* reporter gene (*β-glucuronidase*). The *uceA1.7* promoter has been patented (US8227588B2) and recently used to control *Cry10Aa* gene expression in the development of transgenic cotton resistant to cotton boll weevil (*Anthonomus grandis*) (Ribeiro et al. 2017).

In this study, we generated *A. thaliana* homozygous events stably overexpressing the *gfp* (green fluorescent protein) reporter gene driven by the full-length and four different modules of the *uceA1.7* promoter and compared the transcriptional level and stability with the 35S promoter. The module 1 of the *uceA1.7* promoter containing the core promoter and 5′-UTR, module 2 without the 5′-UTR, module 3 without the 5′-UTR and core promoter identified in this study, and module 4 containing only the core promoter sequence predicted previously by Viana et al. (2011) were evaluated. Our findings suggest the minimal promoter sequence, proper location of the core promoter, and reinforce that full-length *uceA1.7* promoter containing the 5′-UTR is essential to drive higher transcriptional stability in different plant tissues.

Materials and methods

In silico analysis of the promoter and adjacent sequences

The genomic sequences corresponding to *uceA1.7* (position A11:2079996..2084242; gene Gohir.A11G023700) were retrieved from *Gossypium hirsutum v1.1* (Zhang et al. 2015) by the Phytozome v.12 database (Goodstein et al. 2012).

Additional sequences, such as constitutive promoters from CaMV 35S (Odell et al. 1985; Somssich 2019), *Cestrum yellow leaf curling virus* (CmYLCV) (Sanger et al. 1990; Stavolone et al. 2003), *Soybean chlorotic mottle virus* (pIV_SoyCMV and pNCR_SoyCMV) (Hasegawa et al. 1989; Conci et al. 1993; Fukuoka et al. 2000), *Figwort mosaic virus* (FMV) (Sanger et al. 1990), *Sugarcane bacilliform virus* (SBCV) (Tzafrir et al. 1998), *Banana streak virus* (BSV) (Harper and Hull 1998), FMV full-length transcript (FLt) (Maiti et al. 1997), and the promoters flower-specific 1 (FS1) from cotton (Artico et al. 2014), petal-specific anthocyanidin synthase 1 (NtANS1) from *Nicotiana tabacum* (Lim et al. 2013) were also included in *cis*-regulatory elements analysis. Initially, some *cis*-acting regulatory DNA elements were identified using PlantCARE (Lescot et al. 2002), PLACE (Higo et al. 1999), and Plant-PAN 3.0 (Chang et al. 2008; Chow et al. 2019) online database with default parameters. All three databases are public and online resources of *cis*-regulatory elements identified in monocot and dicot plants, which are updated periodically. The *Z score* index (Ma et al. 2013) was also used to identify the *cis*-regulatory elements with a *Z score* ≥ 5, while their descriptions were retrieved from the PLACE database (Higo et al. 1999) (https://www.dna.affrc.go.jp/PLACE/place_seq.shtml). Based on the *cis*-regulatory elements identified in this study and previously by Viana et al. (2011), four modules for the *uceA1.7* promoter were planned (Fig. 1).

Agrobacterium-mediated plant transformation

Binary vectors were synthesized by the company Epoch Life Science (Sugar Land, TX, EUA) and subsequently transfected into the *A. tumefaciens* strain GV3101. The *gfp* reporter gene was cloned under control of the full-length *uceA1.7*, CaMV 35S enhanced, and the four modules of the *uceA1.7* promoter (Fig. 1b). *Hygromycin phosphotransferase II* (hptII) was used as a selection marker gene under control of the ubiquitin-ribosomal protein (ubi3) promoter from tomato (Fig. 1b). *A. thaliana* ecotype Col-0 was genetically transformed by the floral dip method (Clough and Bent 1998) and selected in vitro using 25 mg L^{-1} hygromycin B (Invitrogen) as described by Harrison et al. (2006). Resistant

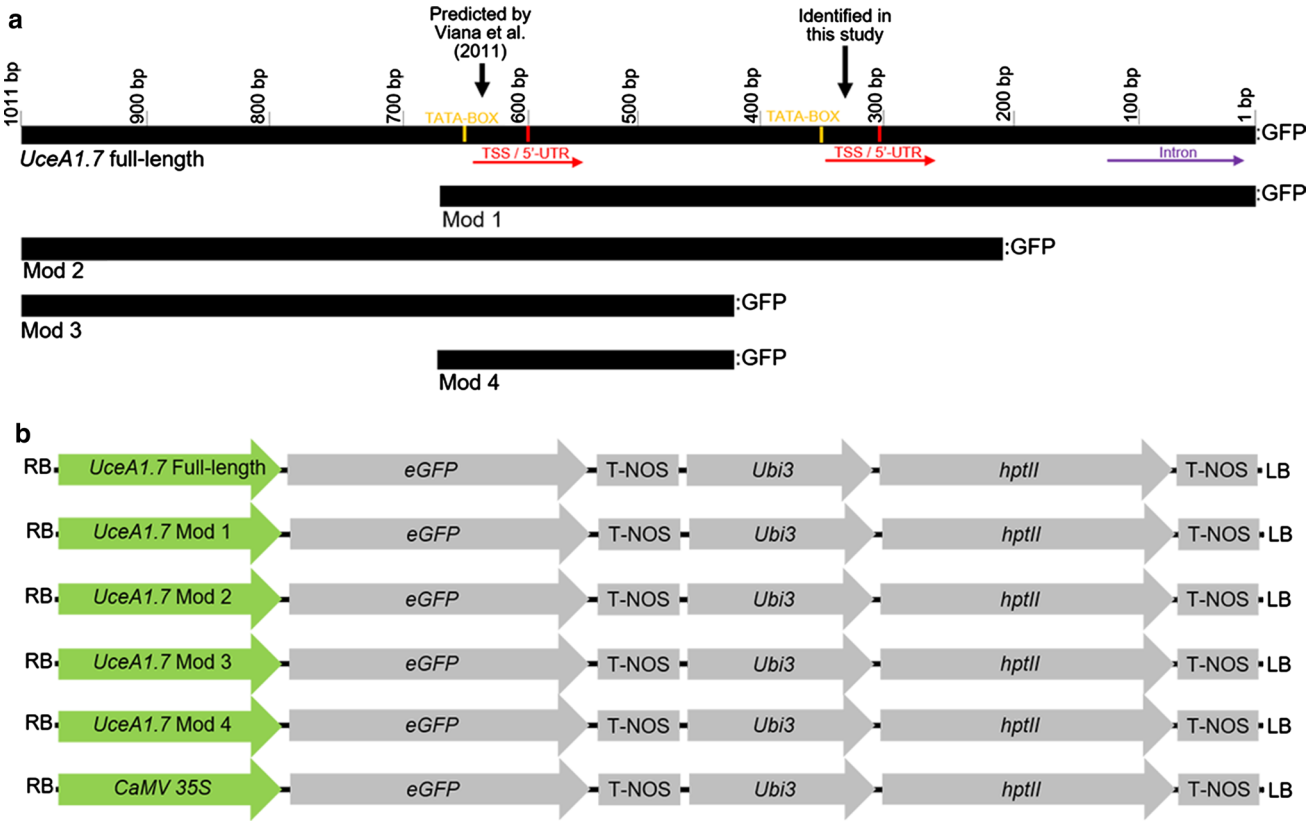


Fig. 1 **a** Schematic representation of the *uceA1.7* promoter regions controlling the expression of the *GFP* reporter gene and binary vectors used in plant transformation. Overview of the four *uceA1.7* modules aligned with the full-length *uceA1.7* promoter. The module 1 containing 681 bp, module 2 containing 792 bp, module 3 containing 581 bp, and module 4 containing 251 bp were planned around the core promoter elements predicted in this study or by Viana et al. (2011). TSS, transcription start site; 5′ UTR, 5′ untranslated region. **b** Overview of the binary vectors for overexpression of *gfp* reporter gene (*green fluorescent protein*) driven by full-length *uceA1.7*, the four *uceA1.7* modules engineered from full-length *uceA1.7* (Mod 1 to 4), and CaMV *35S* promoters

plants were screened by conventional PCR using specific primers (Suppl. Table S1). The generations (T_1–T_4) were advanced with independent events and six homozygous T_4 events were chosen for the molecular characterization of the expression profile of the reporter gene in different plant tissues.

Tissue-specific expression level in transgenic events

Seeds from wild-type and transgenic *A. thaliana* lines were surface sterilized with 75% ethanol for 1 min, sterilized again with 2% sodium hypochlorite solution plus Tween-20 for 10 min, and subsequently rinsed six times with sterile water. Seeds were placed on plates containing half-strength MS medium (Murashige and Skoog 1962) supplemented with 0.8% agar and maintained in the dark for 3 days at 4 °C for stratification. Plants were maintained in a climate-controlled growth room at 22 ± 2 °C, 70–80% relative humidity, with the light intensity 100 µmol photons m^{-2} s^{-1}, and 16/8-h light/dark photoperiod. Root, stem, leaf, silique, and flower bud tissues were harvested, ground in liquid nitrogen using mortar and pestle, and stored at − 80 °C. The total RNA was isolated using TRIzol Reagent (Invitrogen) according to the manufacturer's instructions. The RNA concentration was estimated using a NanoDrop 2000 spectrophotometer (Thermo Scientific), and its integrity was evaluated with 1% agarose gel electrophoresis. RNA samples were treated with RNase-free RQ1 DNase I (Promega) according to the manufacturer's instructions. Then 2 µg of DNase-treated RNA was used as a template for cDNA synthesis using Oligo-$(dT)_{20}$ primer and SuperScript III RT (Life Technologies), according to the manufacturer's instructions. The cDNA was quantified by spectrophotometry and diluted with nuclease-free water to 400 ng μL^{-1}. RT-qPCR assays were performed in an Applied Biosystems 7500 Fast Real-Time PCR System (Applied Biosystems) using 400 ng of cDNA, 0.2 µM of each gene-specific primer (Suppl. Table S1) and GoTaq® qPCR Master Mix (Promega). Conditions for RT-qPCR included an initial 95 °C for 10 min, then 40 cycles of 95 °C for 15 s and 60 °C for 1 min, followed by melting curve analysis. The expression level was normalized using the Actin 2 endogenous reference gene. In a preliminary RT-qPCR assay using three reference genes (GAPDH, EF1, and Actin 2) and some cDNA samples, the Actin 2 reference gene showed the highest stability (Suppl. Table S1). Three biological replicates for each independent event and four plants for each biological replicate were used, while all samples were carried out in technical triplicate reactions. Primer efficiencies and target-specific amplification were confirmed by a single and distinct peak in melting curve analysis. The relative expression level was calculated using the delta–delta cycle threshold (ΔΔCT) method (Schmittgen and Livak 2008).

Transcriptional stability over generations

The transcriptional stability of the *gfp* reporter gene driven by full-length *uceA1.7*, module 1, module 2, and 35S promoters was evaluated over the T_5 and T_6 generations from *A. thaliana* homozygous events. Three independent events from each promoter were selected, germinated in vitro, and grown in the growth chamber. Leaf, flower bud, and root tissues were harvested from adult plants, RNA was isolated, and cDNA was synthesized as described above. The expression profile was evaluated using real-time RT-qPCR, normalized with the Actin 2 reference gene, and represented as fold-change values calculated using the delta–delta cycle threshold (ΔΔCT) method (Schmittgen and Livak 2008). Three biological replicates for each independent event and four plants for each biological replicate were used, while all biological replicates were carried out in technical triplicate reactions.

Results

In silico analysis of the promoter and adjacent sequences

The TATA-box, YR Rule, TSS, some other *cis*-regulatory elements, and a long 5′-UTR containing an intron sequence of the cotton *uceA1.7* promoter were previously predicted by Viana et al. (2011). In this study, using in silico analysis, it was verified that this long 5′-UTR contains the core promoter and a short 5′-UTR with an intron sequence, both of which are rich in *cis*-regulatory elements (Suppl. File S1). Using the PlantCARE, Plant-PAN, and PLACE databases, several *cis*-regulatory elements were identified in the upstream region (843 bp length) of the *uceA1.7* promoter, in the full-length *uceA1.7* promoter sequence, in its 5′-UTR sequence, in its primary transcript (exon and intron sequences), and in its 3′-UTR and additional (378-bp length) downstream sequences (Suppl. Table S2). In the full-length *uceA1.7* promoter sequence, 542 *cis*-regulatory elements were identified, while in its short 5′-UTR, 209 *cis*-regulatory elements were identified, compared to 428 *cis*-regulatory elements identified in the 35S promoter. In contrast, using a *Z score* index supported by high statistical significance (*P* value less than 0.05), 17 *cis*-regulatory elements (8 unique) were identified in the full-length *uceA1.7* promoter sequence and 4 *cis*-regulatory elements (2 unique) in its 5′-UTR, while in the 35S promoter, 47 *cis*-regulatory elements (14 unique) were identified (Suppl. Tables S3, S4). A high number of *cis*-regulatory elements, 1083 and 326, respectively, were also identified in the *uceA1.7* transcript sequence and in its 3′-UTR using the online databases (Suppl. Table S2). Similarly, using the *Z score* index 143 (40 unique) and 3 *cis*-regulatory

elements were identified in the transcript sequence and its 3′-UTR, respectively (Table 2). The common *cis*-acting element in promoter and enhancer regions (PlantCARE), binding site for bZIP transcription factors (TFs), binding site for DOF zinc finger protein (Plant-PAN), core site required for binding of DOF proteins and tetranucleotide of mesophyll expression module 1 (PLACE) were the most represented in the full-length *uceA1.7* promoter. In contrast, the binding sites for bZIP, MYB, AP2, and bHLH TFs were predominant in the 35S promoter (Suppl. Table S2). In comparison with other viral constitutive or tissue-specific promoters, using the *Z score* index, it was observed that full-length *uceA1.7* promoter has a similar amount of *cis*-regulatory elements (8 unique or 17 redundant), but with slightly different characteristics (Suppl. Table S2). The following number of *cis*-regulatory elements were identified in other promoters as CmYLCV (eight unique or nine redundant), pIV_SoyCMV (eight unique or eight redundant), pNCR_SoyCMV (three unique or four redundant), FMV (11 unique or 13 redundant), SCBV (20 unique or 50 redundant), BSV (15 unique or 18 redundant), FLt (12 unique or 30 redundant), cotton flower-specific FS1 (seven unique or 11 redundant), and tobacco petal-specific NtANS1 (14 unique or 37 redundant). The most representative *cis*-regulatory elements of the *uceA1.7* promoter sequence were the binding sites to TFs involved in hormonal signaling, while in the viral promoters analyzed were the binding sites to TFs involved in the plant defense response.

From this previous analysis of *cis*-regulatory elements and core promoters, four modules from the full-length *uceA1.7* promoter were analyzed (Fig. 1a; Suppl. Fig. S1a to S1d; Suppl. File S2). In the modules 1 to 4, 395, 448, 343, and 169, *cis*-regulatory elements were identified based on three online databases (Suppl. Table S2). In contrast, using the *Z score* index, 13 (six unique) *cis*-regulatory elements were identified in modules 1 and 2, while in modules 3 and 4, only one and no *cis*-regulatory elements were identified, respectively (Suppl. Tables S3, S4). The minimal promoter sequence, the core promoter, the contribution of 5′-UTR in enhancing and stability of the transcription, putative *uceA1.7* promoter module with higher tissue-specific activity were evaluated. Module 1 containing the core promoter and 5′-UTR, module 2 without the 5′-UTR, module 3 without the 5′-UTR and core promoter identified in this study, and module 4 containing only the core promoter sequence predicted previously by Viana et al. (2011) were synthesized and cloned into binary vectors to drive expression of the *gfp* reporter gene.

Plant transformation

Around 15 independent events of *A. thaliana* from each module were obtained overexpressing the *gfp* reporter gene driven by the full-length *uceA1.7* promoter, its four modules, and the 35S promoter (Fig. 1b; Suppl. Fig. S1a). Transgenic plants were screened in MS medium containing hygromycin B and confirmed by conventional PCR (Suppl. Fig. S2a, b). Plant generations were advanced to achieve homozygous events. Six homozygous events from each promoter were selected for molecular characterization and the clone plants from each event were propagated in a growth chamber.

Tissue-specific expression level

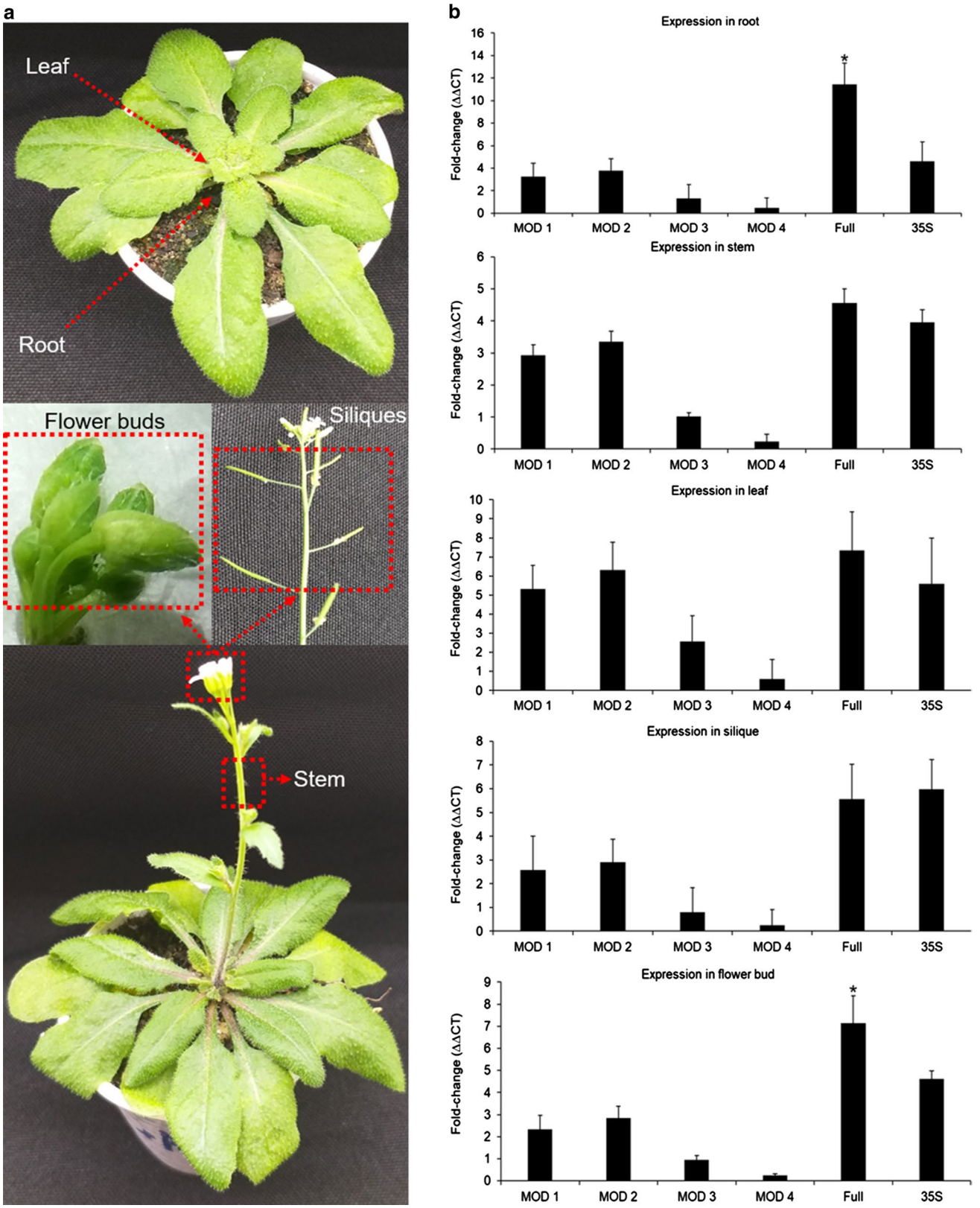
Root, stem, leaf, silique, and flower buds were harvested from adult plants, and the expression profile of the *gfp* reporter gene was monitored using real-time RT-qPCR (Fig. 2a). The full-length *uceA1.7* promoter showed higher activity in the root and flower buds than did the 35S promoter, while in the stem, leaf, and silique tissues, no differences were found (Fig. 2b). In contrast, module 4 presented the lowest activity among the four evaluated modules in both plant tissues evaluated, while modules 1 and 2 had the highest activity. However, both modules showed lower activity than the full-length *uceA1.7* and 35S promoters in almost all tissues evaluated.

Transcriptional stability over generations

Three T_5 and T_6 independent events (homozygous) from module 1, module 2, full-length *uceA1.7*, and 35S promoter were evaluated with respect to transcriptional stability over these two generations conferred by the presence of 5′-UTR. Leaf, flower bud, and root tissues were harvested from adult plants and screened using real-time RT-qPCR (Fig. 3a–c). Module 1 showed greater transcriptional stability, similar to the full-length *uceA1.7* promoter, while module 2 was more unstable over these two generations.

Discussion

Time-, tissue- and amount-specific transcriptional gene regulations play an important role in all biological processes and in plant development. These fine-tuning mechanisms are controlled mainly by interactions between promoter sequences and the numerous TFs (Liu and Stewart 2016). In addition, other proteins (co-activators) also interact with the operon sequences and act to modulate the transcriptional level (initiation, repression, or regulation of transcription). Typical encoding-protein promoters often contain the TATA-box domain located ~30–100 bp upstream of the transcription start site (TSS), which is considered essential to transcriptional initiation (Burley and Roeder 1996; Smale and Kadonaga 2003). Other motifs may also be present in this promoter region near the TATA-box (e.g., YR Rule,



Y-Patch, and REGs) (Yamamoto et al. 2007). The region containing these motifs and the binding sites for the basic transcriptional machinery (e.g., RNA polymerase II and its subunits) is denominated the core promoter (Molina and Grotewold 2005). Proximal and distal regions of the core promoter contain several enhancers, repressors, insulators,

◂**Fig. 2** Tissue-specific expression driven by the full-length *uceA1.7* promoter (Full) and its four modules (modules 1–4; MOD1–4) compared to the CaMV 35S promoter. **a** Root, stem, leaf, silique, and flower bud tissues were harvested and **b** the expression profile of the reporter gene was evaluated. **b** Expression profile of the *gfp* reporter gene (green fluorescent protein) in different tissues of the *A. thaliana* homozygous events was performed using real-time RT-qPCR. The expression levels are represented as fold-change values calculated using the delta–delta cycle threshold (ΔΔCT) method and non-transgenic plants as reference. Error bars represent confidence intervals corresponding to the average of six independent events, three biological replicates per event, while each biological replicate was composed of four plants. Asterisks indicate significant differences from the 35S promoter based on Tukey's test at 5%

and *cis*-regulatory element sequences that contribute to the regulation of gene expression (Bulger and Groudine 2011). These *cis*-regulatory elements are binding sites to TFs, and the amount, their features, and spacing between them define the temporal and spatial expression levels. In addition, these *cis*-regulatory elements are also found in the 5′-UTR, exons, introns, and 3′-UTR conferring more transcriptional stability of the gene or even acting in the initiation of transcription (Hernandez-Garcia and Finer 2014; Biłas et al. 2016). This genome region upstream of the gene containing all these transcriptional regulatory elements, which may or may not include the 5′-UTR, is named the minimal promoter sequence (or full-length promoter) and considered essential for correct transcription (local, temporal, and amount specific) of the target gene. In some cases, partial sequences from the full-length promoter (named promoter modules) containing specific *cis*-regulatory elements may confer greater constitutive, plant tissue- or stage-specific transcriptional activity (Wang et al. 2015). Smaller (compact) promoter sequences that have a high transcriptional activity that is constitutive, induced, tissue- or developmental stage-specific are of extreme relevance for the development of NBTs (Limera et al. 2017; Basso et al. 2019). Promoters with high activity in flower buds, roots and leaves are required for efficient control of pests that specifically attack these tissues, for example, cotton boll weevil, nematodes, and caterpillars in important crops (e.g., cotton, soybean, maize, and sugarcane) (Ribeiro et al. 2017; Wang et al. 2017).

The *uceA1.7* promoter was isolated from cotton (Viana et al. 2011) and used to drive the stable expression of the Cry10Aa entomotoxic protein against the cotton boll weevil (*Anthonomus grandis*) (Ribeiro et al. 2017). Viana et al. (2011) verified that this promoter had a constitutive and higher activity than the viral 35S promoter, in different tissues of transgenic *A. thaliana*. Ribeiro et al. (2017) confirmed that this promoter has constitutive and high activity in cotton flower buds. In this work, we generated four modules from this promoter to identify its core promoter region, minimal promoter sequence, the relevance of its 5′-UTR, and the constitutive and tissue-specific expression levels in *A. thaliana*. Viana et al. (2011) identified by in silico prediction a TATA-box and the putative TSS sequence in the *uceA1.7* promoter, highlighting the importance of the long 5′-UTR for the high expression of the *gus* reporter gene in *A. thaliana* events. However, supported by the recent cotton genome and transcriptome sequencing data available from the Phytozome database (Goodstein et al. 2012; Zhang et al. 2015), new TATA-box and TSS for the *uceA1.7* promoter are suggested in this work. Consequently, the long 5′-UTR predicted initially corresponds, in fact, to the core promoter and a short 5′-UTR rich in *cis*-regulatory elements. In addition, the distribution of *cis*-regulatory elements associated with the transcriptional level conferred by each module of this promoter indicated the minimal promoter sequence and strong importance to the 5′-UTR for high expression level and stability. Curiously, overexpression of GOIs driven by the full-length *uceA1.7* promoter in transgenic cotton showed high transcriptional stability in different plant tissues, development stages, and in flower buds of different sizes (unpublished data). Similar results were obtained using modules from the 35S promoter, indicating that the deletion of its core promoter strongly decreases transcriptional activity (Benfey and Chua 1990). In addition, we identified a transcriptional enhancer sequence downstream of the core promoter, and verified that tissue-specific gene expression depends on synergistic interactions among *cis*-regulatory elements. Our data also confirm that the full-length *uceA1.7* promoter has higher activity in root and flower bud tissues compared to the 35S promoter, whereas it was equivalent in the other tissues (stem, leaf, and silique), supporting its use in the development of NBTs focused on these plant tissues.

Some promoters of viral origin have been successfully used in plant genetic engineering, primarily because they are compact sequences, rich in *cis*-regulatory elements, and already validated in several plant species (Acharya et al. 2014). However, at present, the level of target gene expression conferred by these promoters may not yet be high enough in the desired tissue (Artico et al. 2014; Lambret-Frotte et al. 2016; Ribeiro et al. 2017). Additionally, new promoters with high tissue-specific activity are currently demanded for the gene stacking associated with multiple agronomic traits in several crops (Dong and Ronald 2019). In silico analysis of *cis*-regulatory elements of the *uceA1.7* promoter compared to flower-specific (NtANS1 and FS1) and viral promoters (35S, CmYLCV, SoyCMV, FMV, SBCV, and BSV) shows that these elements are not fully conserved and clearly related to tissue-specific or constitutive activity of the promoter. Thus, the overall contribution (and not specific or only specific) of these *cis*-regulatory elements may be more related to the typical regulation of gene expression (Hernandez-Garcia and Finer 2014). The use of synthetic promoters that combine classical promoters with enhancer sequences, partial sequences from tissue-specific

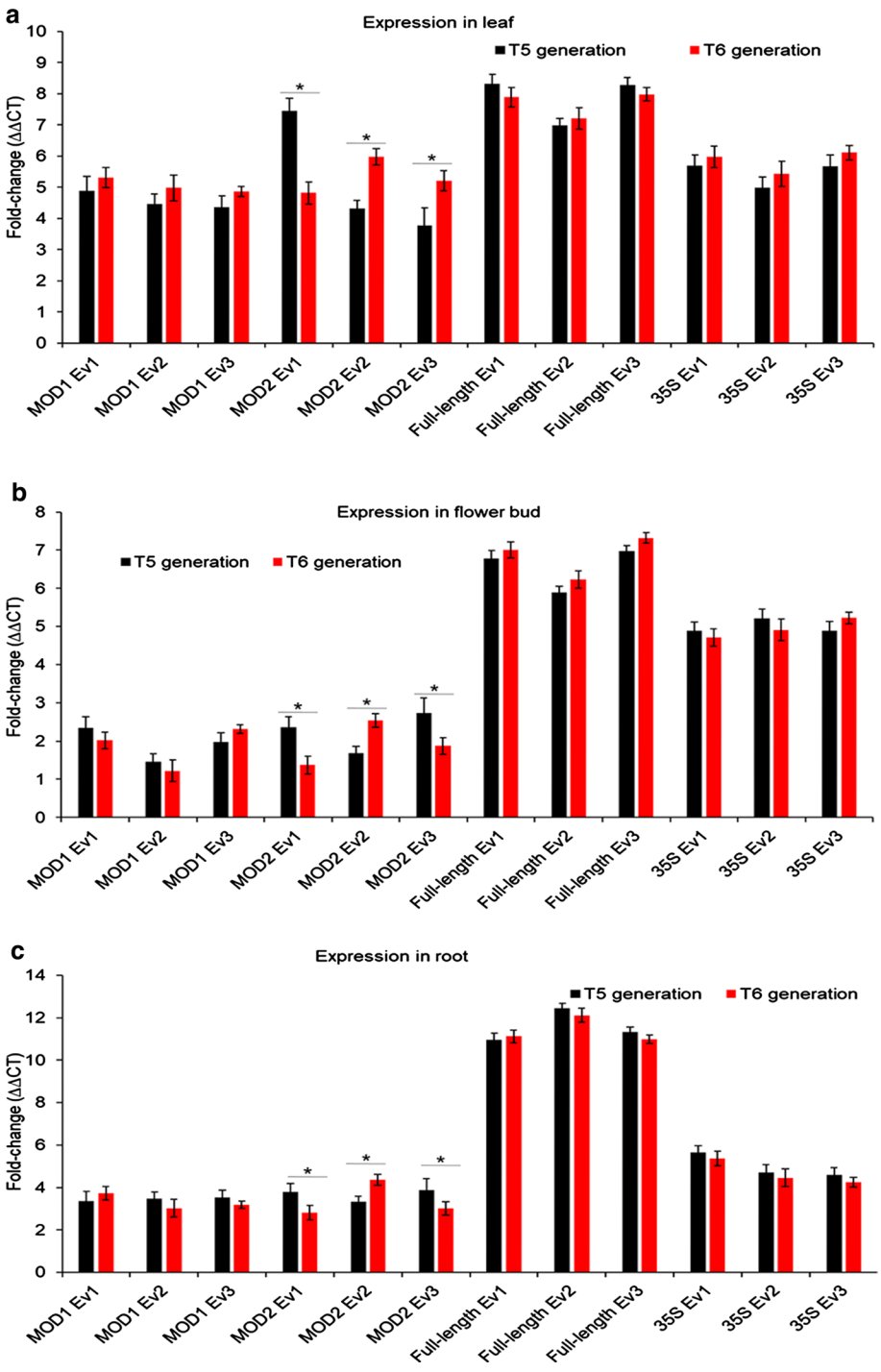


Fig. 3 Transcriptional stability of the *gfp* reporter gene (green fluorescent protein) in leaf (**a**), flower bud (**b**), and root (**c**) of the *A. thaliana* homozygous events over the T_5 and T_6 generations. The expression profile was evaluated using real-time RT-qPCR and represented as fold-change values calculated using the delta–delta cycle threshold (ΔΔCT) method and non-transgenic plants as a reference. Error bars represent confidence intervals corresponding to the average three biological replicates composed of four plants each. Asterisks indicate significant differences from the T_5 to T_6 generations based on Tukey's test at 5%

promoters, regions rich in *cis*-regulatory elements from other promoters, introns, 5′-UTR, and double 3′-UTR is an interesting alternative that has shown promising results (van der Meer et al. 1990, 1992; Wang et al. 2015; Liu and Stewart 2016; Diamos and Mason 2018). Therefore, the use of the full-length *uceA1.7* promoter in combination with enhancer sequences may lead to an increased tissue-specific expression. Since the cotton crop is the target of numerous insect pests and nematodes, the use of this promoter in the original culture for the development of NBTs to overcome these drawbacks is an additional possibility.

Conclusions

In the present study, we further characterized the cotton *uceA1.7* promoter at both the structural level and the ability to control the tissue-specific expression of GOIs. Our

data provided evidence of the correct location of a minimal portion of the *uceA1.7* promoter required to initiate transcription properly (core promoter). The full-length *uceA1.7* promoter containing the 5′-UTR was considered to be essential for higher transcriptional stability in different plant tissues. These findings highlighted the potential use of the full-length *uceA1.7* promoter for the development of NBTs, mainly for cotton crops.

Author contribution statement MFGS was the lead researcher for all the work and provided intellectual input and financial support. ITL-T planned the four *uceA1.7* modules and designed the binary vectors. MFB performed the plant transformation, the advancement of generations and, helped by ITL-T, performed the real-time RT-qPCR assays. EOF performed the CaMV 35: green fluorescent protein (GFP) plant transformation and the advancement of generations. CEMP and TPR helped with the screening of plants and generation advancement. MFB, CB, and ACO performed all analyses of *cis*-regulatory elements from promoter sequences. MFB wrote the manuscript, while JAE, MAF and MFGS provided intellectual input. All authors read and approved the final version.

Acknowledgements We are grateful to EMBRAPA, CAPES, CNPq, INCT PlantStress Biotech, and FAP-DF for providing financial support for this scientific research.

Funding MFB is grateful to Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for a post-doctoral research fellowship (PDJ 150936/2018-4). This work was supported by grants from CAPES, CNPq, FAP-DF, INCT, and EMBRAPA.

Compliance with ethical standards

Conflict of interest The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Acharya S, Ranjan R, Pattanaik S, Maiti IB, Dey N (2014) Efficient chimeric plant promoters derived from plant infecting viral promoter sequences. Planta 239(2):381–396. https://doi.org/10.1007/s00425-013-1973-2

Artico S, Lambret-Frotté J, Nardeli SM, Oliveira-Neto OB, Grossi-de-Sa MF, Alves-Ferreira M (2014) Isolation and characterization of three new promoters from *Gossypium hirsutum* that show high activity in reproductive tissues. Plant Mol Biol Rep 32:630–643

Basso MF, Ferreira PCG, Kobayashi AK, Harmon FG, Nepomuceno AL, Molinari HBC, Grossi-de-Sa MF (2019) MicroRNAs and new biotechnological tools for its modulation and improving stress tolerance in plants. Plant Biotechnol J 17:1482–1500. https://doi.org/10.1111/pbi.13116

Benfey PN, Chua NH (1990) The cauliflower mosaic virus 35S promoter: combinatorial regulation of transcription in plants. Science 250(4983):959–966. https://doi.org/10.1126/science.250.4983.959

Benfey PN, Ren L, Chua NH (1990) Tissue-specific expression from CaMV 35S enhancer subdomains in early stages of plant development. EMBO J 9(6):1677–1684

Biłas R, Szafran K, Hnatuszko-Konka K, Kononowicz AK (2016) *Cis*-regulatory elements used to control gene expression in plants. Plant Cell Tissue Organ Cult 127(2):269–287. https://doi.org/10.1007/s11240-016-1057-7

Bulger M, Groudine M (2011) Functional and mechanistic diversity of distal transcription enhancers. Cell 144(3):327–339. https://doi.org/10.1016/j.cell.2011.01.024

Burley SK, Roeder RG (1996) Biochemistry and structural biology of transcription factor IID (TFIID). Annu Rev Biochem 65:769–799. https://doi.org/10.1146/annurev.bi.65.070196.004005

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 Genom 9:561. https://doi.org/10.1186/1471-2164-9-561

Chow CN, Lee TY, Hung YC, Li GZ, Tseng KC, Liu YH, Kuo PL, Zheng HQ, Chang WC (2019) PlantPAN3.0: a new and updated resource for reconstructing transcriptional regulatory networks from ChIP-seq experiments in plants. Nucleic Acids Res 47(D1):D1155–D1163. https://doi.org/10.1093/nar/gky1081

Christensen AH, Sharrock RA, Quail PH (1992) Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. Plant Mol Biol 18(4):675–689

Clough SJ, Bent AF (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J 16(6):735–743

Conci LR, Nishizawa Y, Saito M, Date T, Hasegawa A, Miki K, Hibi T (1993) A strong promoter fragment from the large noncoding region of soybean chlorotic mottle virus DNA. Jpn J Phytopath 59(4):432–437. https://doi.org/10.3186/jjphytopath.59.432

Crisp PA, Ganguly D, Eichten SR, Borevitz JO, Pogson BJ (2016) Reconsidering plant memory: intersections between stress recovery, RNA turnover, and epigenetics. Sci Adv 2(2):e1501340. https://doi.org/10.1126/sciadv.1501340

Diamos AG, Mason HS (2018) Chimeric 3' flanking regions strongly enhance gene expression in plants. Plant Biotechnol J 16(12):1971–1982. https://doi.org/10.1111/pbi.12931

Dong OX, Ronald PC (2019) Genetic engineering for disease resistance in plants: recent progress and future perspectives. Plant Physiol 180(1):26–38. https://doi.org/10.1104/pp.18.01224

Fukuoka H, Ogawa T, Mitsuhara I, Iwai T, Isuzugawa K, Nishizawa Y, Gotoh Y, Nishizawa Y, Tagiri A, Ugaki M, Ohshima M, Yano H, Murai N, Niwa Y, Hibi T, Ohashi Y (2000) *Agrobacterium*-mediated transformation of monocot and dicot plants using the NCR promoter derived from soybean chlorotic mottle virus. Plant Cell Rep 19(8):815–820. https://doi.org/10.1007/s002990000191

Garbarino JE, Oosumi T, Belknap WR (1995) Isolation of a polyubiquitin promoter and its expression in transgenic potato plants. Plant Physiol 109(4):1371–1378. https://doi.org/10.1104/pp.109.4.1371

Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, Rokhsar DS (2012) Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res 40((Database issue)):D1178–D1186. https://doi.org/10.1093/nar/gkr944

Grefen C, Donald N, Hashimoto K, Kudla J, Schumacher K, Blatt MR (2010) A ubiquitin-10 promoter-based vector set for fluorescent protein tagging facilitates temporal stability and native protein distribution in transient and stable expression studies. Plant J 64(2):355–365. https://doi.org/10.1111/j.1365-313X.2010.04322.x

Gupta P, Raghuvanshi S, Tyagi AK (2001) Assessment of the efficiency of various gene promoters via biolistics in leaf and regenerating seed callus of millets, *Eleusine coracana* and *Echinochloa crus-galli*. Plant Biotechnol J 18(4):275–282. https://doi.org/10.5511/plantbiotechnology.18.275

Harper G, Hull R (1998) Cloning and sequence analysis of banana streak virus DNA. Virus Genes 17(3):271–278. https://doi.org/10.1023/a:1008021921849

Harrison SJ, Mott EK, Parsley K, Aspinall S, Gray JC, Cottage A (2006) A rapid and robust method of identifying transformed *Arabidopsis thaliana* seedlings following floral dip transformation. Plant Methods 2(1):19. https://doi.org/10.1186/1746-4811-2-19

Hasegawa A, Verver J, Shimada A, Saito M, Goldbach R, Van Kammen A, Miki K, Kameya-Iwaki M, Hibi T (1989) The complete sequence of soybean chlorotic mottle virus DNA and the identification of a novel promoter. Nucleic Acids Res 17(23):9993–10013

Hernandez-Garcia CM, Finer JJ (2014) Identification and validation of promoters and *cis*-acting regulatory elements. Plant Sci 217–218:109–119. https://doi.org/10.1016/j.plantsci.2013.12.007

Hernandez-Garcia CM, Martinelli AP, Bouchard RA, Finer JJ (2009) A soybean (*Glycine max*) polyubiquitin promoter gives strong constitutive expression in transgenic soybean. Plant Cell Rep 28(5):837–849. https://doi.org/10.1007/s00299-009-0681-7

Hickey LT, Hafeez AN, Robinson H, Jackson SA, Leal-Bertioli SCM, Tester M, Gao C, Godwin ID, Hayes BJ, Wulff BBH (2019) Breeding crops to feed 10 billion. Nat Biotechnol 37(7):744–754. https://doi.org/10.1038/s41587-019-0152-9

Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant *cis*-acting regulatory DNA elements (PLACE) database. Nucleic Acids Res 27(1):297–300. https://doi.org/10.1093/nar/27.1.297

Lambret-Frotte J, Artico S, Muniz Nardeli S, Fonseca F, Brilhante Oliveira-Neto O, Grossi-de-Sa MF, Alves-Ferreira M (2016) Promoter isolation and characterization of *GhAO-like1*, a *Gossypium hirsutum* gene similar to multicopper oxidases that is highly expressed in reproductive organs. Genome 59(1):23–36. https://doi.org/10.1139/gen-2015-0055

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(1):325–327. https://doi.org/10.1093/nar/30.1.325

Lim S-H, Kim JK, Lee J-Y, Kim Y-M, Sohn S-H, Kim D-H, Ha S-H (2013) Petal-specific activity of the promoter of an anthocyanidin synthase gene of tobacco (*Nicotiana tabacum* L.). Plant Cell Tissue Organ Cult 114(3):373–383. https://doi.org/10.1007/s11240-013-0332-0

Limera C, Sabbadini S, Sweet JB, Mezzetti B (2017) New biotechnological tools for the genetic improvement of major woody fruit species. Front Plant Sci 8:1418. https://doi.org/10.3389/fpls.2017.01418

Liu W, Stewart CN Jr (2016) Plant synthetic promoters and transcription factors. Curr Opin Biotechnol 37:36–44. https://doi.org/10.1016/j.copbio.2015.10.001

Lu J, Sivamani E, Azhakanandam K, Samadder P, Li X, Qu R (2008a) Gene expression enhancement mediated by the 5' UTR intron of the rice rubi3 gene varied remarkably among tissues in transgenic rice plants. Mol Genet Genom 279(6):563–572. https://doi.org/10.1007/s00438-008-0333-6

Lu J, Sivamani E, Li X, Qu R (2008b) Activity of the 5' regulatory regions of the rice polyubiquitin *rubi3* gene in transgenic rice plants as analyzed by both *GUS* and *GFP* reporter genes. Plant Cell Rep 27(10):1587–1600. https://doi.org/10.1007/s00299-008-0577-y

Ma S, Shah S, Bohnert HJ, Snyder M, Dinesh-Kumar SP (2013) Incorporating motif analysis into gene co-expression networks reveals novel modular expression pattern and new signaling pathways. PLoS Genet 9(10):e1003840. https://doi.org/10.1371/journal.pgen.1003840

Maiti IB, Gowda S, Kiernan J, Ghosh SK, Shepherd RJ (1997) Promoter/leader deletion analysis and plant expression vectors with the *Figwort mosaic virus* (FMV) full length transcript (FLt) promoter containing single or double enhancer domains. Transgenic Res 6(2):143–156. https://doi.org/10.1023/a:1018477705019

Molina C, Grotewold E (2005) Genome wide analysis of *Arabidopsis* core promoters. BMC Genom 6:1–12. https://doi.org/10.1186/1471-2164-6-25

Moon J, Parry G, Estelle M (2004) The ubiquitin-proteasome pathway and plant development. Plant Cell 16(12):3181–3195. https://doi.org/10.1105/tpc.104.161220

Murashige T, Skoog F (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiol Plant 15(3):473–497. https://doi.org/10.1111/j.1399-3054.1962.tb08052.x

Odell JT, Nagy F, Chua NH (1985) Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. Nature 313(6005):810–812. https://doi.org/10.1038/313810a0

Park S-H, Yi N, Kim YS, Jeong M-H, Bang S-W, Choi YD, Kim J-K (2010) Analysis of five novel putative constitutive gene promoters in transgenic rice plants. J Exp Bot 61(9):2459–2467. https://doi.org/10.1093/jxb/erq076

Ribeiro TP, Arraes FBM, Lourenco-Tessutti IT, Silva MS, Lisei-de-Sa ME, Lucena WA, Macedo LLP, Lima JN, Santos Amorim RM, Artico S, Alves-Ferreira M, Mattar Silva MC, Grossi-de-Sa MF (2017) Transgenic cotton expressing Cry10Aa toxin confers high resistance to the cotton boll weevil. Plant Biotechnol J 15(8):997–1009. https://doi.org/10.1111/pbi.12694

Sanger M, Daubert S, Goodman RM (1990) Characteristics of a strong promoter from figwort mosaic virus: comparison with the analogous 35S promoter from cauliflower mosaic virus and the regulated mannopine synthase promoter. Plant Mol Biol 14(3):433–443. https://doi.org/10.1007/bf00028779

Schmittgen TD, Livak KJ (2008) Analyzing real-time PCR data by the comparative CT method. Nat Protoc 3:1101. https://doi.org/10.1038/nprot.2008.73

Smale ST, Kadonaga JT (2003) The RNA polymerase II core promoter. Annu Rev Biochem 72:449–479. https://doi.org/10.1146/annurev.biochem.72.121801.161520

Somssich M (2019) A short history of the CaMV 35S promoter. PeerJ Prepr 7:e27096v27093. https://doi.org/10.7287/peerj.preprints.27096v3

Stavolone L, Kononova M, Pauli S, Ragozzino A, de Haan P, Milligan S, Lawton K, Hohn T (2003) Cestrum yellow leaf curling virus (CmYLCV) promoter: a new strong constitutive promoter for heterologous gene expression in a wide variety of crops. Plant Mol Biol 53(5):663–673. https://doi.org/10.1023/B:PLAN.0000019110.95420.bb

Tzafrir I, Torbert KA, Lockhart BE, Somers DA, Olszewski NE (1998) The sugarcane bacilliform badnavirus promoter is active in both monocots and dicots. Plant Mol Biol Rep 38(3):347–356

van der Meer IM, Spelt CE, Mol JNM, Stuitje AR (1990) Promoter analysis of the chalcone synthase (*chsA*) gene of *Petunia hybrida*: a 67 bp promoter region directs flower-specific expression. Plant Mol Biol 15(1):95–109. https://doi.org/10.1007/bf00017727

van der Meer IM, Brouwer M, Spelt CE, Mol JN, Stuitje AR (1992) The TACPyAT repeats in the chalcone synthase promoter of *Petunia hybrida* act as a dominant negative *cis*-acting module in the control of organ-specific expression. Plant J 2(4):525–535

Viana AA, Fragoso RR, Guimaraes LM, Pontes N, Oliveira-Neto OB, Artico S, Nardeli SM, Alves-Ferreira M, Batista JA, Silva MC, 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. https://doi.org/10.1186/1472-6750-11-115

Wang J, Oard JH (2003) Rice ubiquitin promoters: deletion analysis and potential usefulness in plant transformation systems. Plant Cell Rep 22(2):129–134. https://doi.org/10.1007/s00299-003-0657-y

Wang R, Zhu M, Ye R, Liu Z, Zhou F, Chen H, Lin Y (2015) Novel green tissue-specific synthetic promoters and *cis*-regulatory elements in rice. Sci Rep 5:18256. https://doi.org/10.1038/srep18256

Wang WZ, Yang BP, Feng XY, Cao ZY, Feng CL, Wang JG, Xiong GR, Shen LB, Zeng J, Zhao TT, Zhang SZ (2017) Development and characterization of transgenic sugarcane with insect resistance and herbicide tolerance. Front Plant Sci 8:1535–1535. https://doi.org/10.3389/fpls.2017.01535

Weeks JT, Anderson OD, Blechl AE (1993) Rapid production of multiple independent lines of fertile transgenic wheat (*Triticum aestivum*). Plant Physiol 102(4):1077–1084. https://doi.org/10.1104/pp.102.4.1077

Wei H, Wang ML, Moore PH, Albert HH (2003) Comparative expression analysis of two sugarcane polyubiquitin promoters and flanking sequences in transgenic plants. J Plant Physiol 160(10):1241–1251. https://doi.org/10.1078/0176-1617-01086

Yamamoto YY, Ichida H, Matsui M, Obokata J, Sakurai T, Satou M, Seki M, Shinozaki K, Abe T (2007) Identification of plant promoter constituents by analysis of local distribution of short sequences. BMC Genom 8:67. https://doi.org/10.1186/1471-2164-8-67

Zhang T, Hu Y, Jiang W, Fang L, Guan X, Chen J (2015) Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. Nat Biotechnol 33(5):531–537. https://doi.org/10.1038/nbt.3207

Zhou B, Mural RV, Chen X, Oates ME, Connor RA, Martin GB, Gough J, Zeng L (2017) A subset of ubiquitin-conjugating enzymes is essential for plant immunity. Plant Physiol 173(2):1371–1390. https://doi.org/10.1104/pp.16.01190