Transgenic Research **11:** 347–359, 2002.

Cotton α-globulin promoter: isolation and functional characterization in transgenic cotton, *Arabidopsis*, and tobacco

Ganesan Sunilkumar[1,2], James P. Connell[3], C. W. Smith[2], Avutu S. Reddy[3] & Keerti S. Rathore[1,2,*]

[1]*Institute for Plant Genomics & Biotechnology, Texas A&M University, College Station, TX 77843-2123, USA*
[2]*Department of Soil & Crop Sciences, Texas A&M University, College Station, TX 77843, USA*
[3]*Dow AgroSciences, Indianapolis, IN, USA*

Received 5 June 2001; revised 15 September 2001; accepted 5 October 2001

Key words: *Agrobacterium*, *Arabidopsis*, cotton, α-globulin, promoter, tobacco

Abstract

Globulins are the most abundant seed storage proteins in cotton and, therefore, their regulatory sequences could potentially provide a good source of seed-specific promoters. We isolated the putative promoter region of cotton α-globulin B gene by gene walking using the primers designed from a cotton staged embryo cDNA clone. PCR amplified fragment of 1108 bp upstream sequences was fused to *gus*A gene in the binary vector pBI101.3 to create the test construct. This was used to study the expression pattern of the putative promoter region in transgenic cotton, *Arabidopsis*, and tobacco. Histochemical GUS analysis revealed that the promoter began to express during the torpedo stage of seed development in tobacco and *Arabidopsis*, and during cotyledon expansion stage in cotton. The activity quickly increased until embryo maturation in all three species. Fluorometric GUS analysis showed that the promoter expression started at 12 and 15 dpa in tobacco and cotton, respectively, and increased through seed maturation. The strength of the promoter expression, as reflected by average GUS activity in the seeds from primary transgenic plants, was vastly different amongst the three species tested. In *Arabidopsis*, the activity was 16.7% and in tobacco it was less than 1% of the levels detected in cotton seeds. In germinating seedlings of tobacco and *Arabidopsis*, GUS activity diminished until it was completely absent 10 days post imbibition. In addition, absence of detectable level of GUS expression in stem, leaf, root, pollen, and floral bud of transgenic cotton confirmed that the promoter is highly seed-specific. Analysis of GUS activity at individual seed level in cotton showed a gene dose effect reflecting their homozygous or hemizygous status. Our results show that this promoter is highly tissue-specific and it can be used to control transgene expression in dicot seeds.

Introduction

Seed-specific transgene expression is required for a number of applications utilizing genetic engineering. These include transgenic means to improve seed nutritional quality by manipulating flux through metabolic pathways (Hitz et al., 1995; Kinney, 1996; Shintani & DellaPenna, 1998; Goto et al., 1999) and for the production of novel compounds of industrial or pharmaceutical value (Cahoon et al., 2000) in a convenient package, the seed. Some of these transgenic traits may require expression of more than one transgene in the developing seed (Ye et al., 2000). In other cases, metabolic engineering to improve seed quality may require over-expression and/or suppression of various genes during seed development. Each step will require a promoter of appropriate strength depending on the desired degree of over-expression or suppression. In addition, a promoter with appropriate developmental timing may also be required. Even if the same degree of expression of more than one gene is required, it is not advisable to use the same promoter for multiple introduced genes. In some cases of high copy number integration of transgenes, promoter homology can lead

* Author for correspondence: *E-mail:* rathore@tamu.edu

to gene silencing (Vaucheret, 1993; Brusslan & Tobin, 1995; Park et al., 1996). Thus, it is clear that there will be an increasing need for promoters of varying strengths from more than one source to meet the future demands to regulate expression of one or more transgenes in seeds.

Seed storage proteins are expressed at high levels during seed development, and their expression is tightly controlled both spatially and temporally in the developing seed. Therefore, regulatory sequences from genes encoding seed storage proteins represent a valuable source of promoters that can be utilized to drive the expression of transgenes in a seed-specific manner. The promoters from soybean β-conglycinin genes (Barker et al., 1988; Chen et al., 1988; Lessard et al., 1993), French bean phaseolin gene (Bustos et al., 1989, 1991; Kawagoe et al., 1994), sunflower helianthinin gene (Bogue et al., 1990; Nunberg et al., 1995), and the carrot *Dc3* promoter (Seffens et al., 1990; Kim et al., 1997) are examples of some of the well-characterized seed-specific promoters from dicots.

We have isolated and functionally characterized a 1108 bp promoter sequence from a cotton α-globulin gene. Globulins are principal seed storage proteins of cotton and constitute about 60% of total proteins at seed maturity (Dure, 1989). In cotton, two α-globulin genes, gene A and gene B, encode proteins of molecular weight 48 and 51 kDa, respectively (Chlan et al., 1987). The promoter region of α-globulin gene B was isolated by gene walking and was then fused to the reporter gene, β-glucuronidase (*gus*A) to create the test construct. The construct was used to evaluate the expression pattern controlled by this α-globulin promoter (AGP) sequence in transgenic plants. Earlier studies involving seed-specific transgene expression were carried out mostly in heterologous model species such as tobacco. We studied AGP-driven reporter gene expression in cotton as well as two other heterologous dicot species, tobacco, and *Arabidopsis*. Results presented in this report suggest that the pattern of transgene expression, in terms of tissue-specificity and developmental timing, is similar in these three species, however, there is a several fold difference in the level of transgene expression amongst homologous and heterologous species.

Materials and methods

Isolation of cotton α-globulin promoter

Following the method described by Siebert et al. (1995), the 5′ flanking promoter region was cloned using the sequence information from an α-globulin clone from a cotton staged embryo cDNA library. The sequence resulting from the first walk matched that of the published 5′ flanking sequence of the genomic clone of α-globulin B gene from cotton (Chlan et al., 1987). Further walks extended the promoter sequence which includes a 336 nucleotides of 5′ unpublished sequence. Based on the promoter sequence information, primers were designed to amplify a 1144 nucleotide long fragment containing the combined promoter and the untranslated leader region of the α-globulin B gene from cotton (cv. Coker 312) genomic DNA. The primers used were: AGP5 = 5′-aag-ctt-gca-tgc-ctg-cag-CTA-TTT-TCA-TCC-TAT-TTA-GAA-ATC-3′; AGP3 = 5′-ggg-acg-cgt-atc-GAT-TAC-GAT-AAG-CTC-TGT-ATT-T-TG-3′ (unique restriction sites incorporated into the primers are indicated in lowercase). The amplified PCR product was cloned into the TA cloning vector, pCRII (Invitrogen) resulting in pCRII-AGP. The integrity of the insert was verified by sequencing.

The α-globulin promoter from pCRII-AGP was then introduced as a *Hin*dIII-*Xba*I fragment into the polylinker sequence located upstream of the *gus*A gene in pBI101.3 (Clontech). An out-of-frame ATG (from pCRII polylinker) found upsteam of the GUS coding sequence was removed by deleting the region between the *Not*I and *Sma*I sites to create the test construct pBIAGPGUS. The entire putative promoter and 5′ UTR were sequenced to verify the integrity of the final construct. The binary vector, pBIAGPGUS, which harbors *npt*II as the plant selectable marker gene, was then introduced into *Agrobacterium* strains LBA4404 and GV3101-pMP90 using the method described by An et al. (1988).

Plant materials and transformation

Nicotiana tabacum cv. Havana was transformed with *Agrobacterium* strain LBA4404(pBIAGPGUS) using the leaf disc transformation method (Horsch et al., 1988). Transformants were selected on regeneration medium (MS salts, 100 mg/l myo-inositol, 0.4 mg/l thiamine HCl, 4 μM BAP, 0.5 μM NAA, 3% sucrose, pH 5.6, and solidified with 0.8% Difco-Bacto agar) containing 100 mg/l kanamycin and 500 mg/l carbenicillin. Regenerated shoots were excised and grown on MSO medium (MS salts, B-5 organics, 2% sucrose, pH 5.7, and solidified with 0.8% Difco-Bacto agar) containing 100 mg/l kanamycin and 500 mg/l carbeni-

cillin. Plants with good root systems were transferred to soil and grown to maturity in the greenhouse.

Arabidopsis thaliana C24 plants were transformed by the vacuum infiltration method (Bechtold & Pelletier, 1998) using the *Agrobacterium* strain GV3101-pMP90(pBIAGPGUS). Transformed seeds (T1) were selected on MSO medium containing 50 mg/l kanamycin. The kanamycin-resistant plants were transferred to soil and grown to maturity in a growth room (23°C, 65% humidity, 14 h/10 h photoperiod).

The hypocotyl segments of cotton (*Gossypium hirsutum* cv. Coker 312) seedlings were transformed with *Agrobacterium* [LBA4404(pBIAGPGUS)] by following the method described by Sunilkumar and Rathore (2001). Plants were regenerated from kanamycin-resistant transgenic calli and grown to maturity in the greenhouse.

GUS activity in seeds from transgenic plants

GUS assays were performed on T2 seeds of *Arabidopsis* and T1 seeds of tobacco and cotton as described by Jefferson et al. (1987). These generations of seeds will segregate for transgenes. In order to minimize the variation introduced by null segregants and homozygous seeds, the assay was performed with a large number of seeds. Assays were carried out as triplicates for each transgenic line with 25 (∼2 gm), 150 (∼15 mg), and 300 (∼5 mg) seeds in each replicate for cotton, tobacco, and *Arabidopsis*, respectively. Total protein was measured by using the method of Bradford (1976). GUS activity was normalized to the total protein and the results are presented as GUS specific activity (nanomole of 4-MU released per miligram protein per minute).

GUS activity during various stages of seed and plant development

The histochemical and fluorometric analyses of GUS activity during seed development or seed germination were performed with homozygous plants. The T1 generation in cotton and tobacco, and T2 generation in *Arabidopsis* segregate for the transgene. The T2 embryos from the seeds of homozygous plants of cotton and tobacco, and T3 embryos from the seeds of a homozygous *Arabidopsis* plant were isolated microscopically and stained for GUS as described by Jefferson et al. (1987). Fluorometric GUS assay in tobacco was performed in developing seeds collected from three individual capsules (three replicates each consisting of 35 mg of seeds) for each time point. In cotton, the embryos from one boll for each time point were dissected and individually assayed for GUS activity. GUS expression analyses were also carried out in leaf, root, stem, and floral tissues of a plant that showed GUS activity in seeds.

The seeds of T1 homozygous tobacco and T2 homozygous *Arabidopsis* were germinated on MSO medium, and the seedlings at different days after germination were histochemically assayed (in the case of *Arabidopsis*) or fluorometrically assayed (in the case of tobacco) for GUS activity. After GUS staining, the seedlings were treated in ethanol to clear chlorophyll. In case of tobacco and *Arabidopsis*, the embryos/seedlings were photographed using Kodak Elitechrome Tungsten 160T film. The slides were then scanned and digitally enlarged. The cotton embryo images were captured using a Zeiss AxioCam digital camera coupled to a Zeiss M^2BIO Zoom Stereo/Compound microscope. Figure 2 was compiled using Canvas 7.0 software.

GUS activity in individual embryos isolated from seeds of a homozygous and a hemizygous cotton plant

Seeds from a T0 transgenic cotton plant showing seed-specific GUS expression were first germinated on 100 mg/l kanamycin to eliminate null segregants. Those seeds that germinated and grew in the presence of kanamycin were transferred to soil and grown to maturity in the greenhouse. The zygosity status of these T1 plants was determined by GUS histochemical analysis on the seeds. One homozygous plant and one hemizygous plant were selected for quantitative analysis of GUS activity in their seeds. Embryos isolated from the seeds were analyzed individually for GUS activity using the fluorometric procedure described earlier.

Results

Cotton α-globulin promoter sequence

Figure 1(A) shows the 1108 bp sequence that was isolated and functionally characterized in this study. The 772 bp of the clone toward the 3′ end of the sequence presented here matched that of the published, 5′ genomic flanking sequence for α-globulin gene B (Chlan et al., 1987). Further PCR walks resulted in an additional 336 nucleotides of upstream sequence (underlined). The TATA box and CAAT box are shown in bold letters, and the 5′ untranslated

(A)

```
-1108 CTATTTTCAT CCTATTTAGA AATCCAAGTT GACACCTAAA ATTTAGTTGG ACTGCCATGT AGGATTATCG
-1038 TTAGAGAGAT AACGGAGCTT AACGGTAGAG TGATCACTTT GTAACAAAAT AATAACAAAA GTGACTAAAG
 -968 TGTAACATTT CAAACATAAA TGATTAAAAT ATAACCTGAG GCAAACAAAA ATGACTATTT TTATAGATTA
 -898 CCCTAAAATT AAAGAGTCAT GGCCCTAGCC CCTCGCCTAC TTGTTTGTTT TTAATAAACT AACATAGTAT
 -828 AATATATTGT TAGGATTATA TAAAATTATT AATAAATAGT ATAATTAATT TAAAATTTAT GAAAAATAAA
 -758 TTACCATATT TCTTAAATAC GTGGCACCTT ATGTTGGATT GGACTGTATA ACTTATATAC TATTATCTAT
 -688 ATTGAATCCA AATCCTTACT TTTAAGCGTT TTTAGTGAAA CATTTTATTT TCCATTCTTA TTATATAAAT
 -618 TTATATAATG ATATAATATG TAATACTTAG ATAATATTAT TGAAAAAGAA TAAAAATACC TCAAACTTTG
 -548 AAAGGACTAA TTTGTATGAG CATCAAACGT ACAGGATACC AAAAGTATAC ATATCTGAAT TTGTTCATAT
 -478 CTCCTGCAAC TCATAGATCA TCACCATGCA CAGCAACATG TGTACACTTG ACTTGTCCTC TATCAACTCA
 -408 ACCCTTAACT CAGTGAATCG GGACATCTCT GTCTCACTTT AAAACCCTTC CCAGTTTCAA CACTCTTTGA
 -338 ATTCAACTGA GTTCACATAC AACACAACAC AGTCCATCAT CTTTCTGCTG TTAAAGCATC ATCATTTCGC
 -268 CCCTTCCAGT TACAGATGCA ACATGAACCC CCCTGCAACA AAGTTTGTCC GAACCTTGCT AGTACCATGT
 -198 GAAGGGATGT GGCATCTCGA TATCTACCCA CCACTATACA AAAAAAAAAA AAAGAGACAA TATTTCGTCT
 -128 TCTTTAATTT GCACACTCGT CATCTTGCAT GTCAATGTCT TCAACACGTT GATGAAGATT TGCATGCAAA
                                                                    +1
  -58 AATATCACCT TCCACAGCTC CACCTTCTAT AAATACATTA CCACTCTTTG CTATTACCAT CACACAGTAA
  +13 CAAAATACAG AGCTTATCGT AATC
                                  ↳ATG of α-globulin B gene
```

(B)

Figure 1. α-globulin promoter sequence and reporter gene construct. (A) Nucleotide sequence of the promoter region. Putative *cis*-acting elements are shown inside boxes. The transcription initiation site is indicated with +1. The 5′ untranslated region is shown in italics. The additional, unpublished 336 bp sequence determined in this study is underlined. (B) T-DNA of the binary vector pBIAGPGUS.

region is shown in italics. Visual analysis of the promoter sequence revealed a number of putative DNA motifs that may be involved in tissue-specific transcriptional regulation of the α-globulin gene B. There are four CANNTG motifs (Kawagoe & Murai, 1992), one CATGCACA (RY repeat, Dickinson et al., 1988), and two AACACA motifs (Goldberg, 1986). These *cis*-elements are believed to confer seed-specific expression to the promoter. Our preliminary results on transient expression assays suggested a high degree of tissue (seed)-specificity for the 1108 bp sequence. Therefore, we carried out functional analysis of this sequence by stable transformation of three different species with a binary vector construct containing the reporter gene, β-glucuronidase, under the control of the α-globulin promoter (Figure 1(B)).

Histochemical localization of GUS activity during seed development

Expression of β-glucuronidase gene, under the control of α-globulin promoter, was first tested using transient

expression assays following particle bombardment-mediated transformation of developing embryos, endosperm, and leaves of sorghum and cotton. The results (not shown) indicated that the promoter was active only in developing embryos of cotton. On the basis of these results, we carried out stable transformation of tobacco, *Arabidopsis*, and cotton for more detailed characterization of the cotton α-globulin promoter activity.

AGP:*gus*A expression was evaluated in the seeds from three T1-homozygous tobacco plants. Histochemical analysis results for GUS activity in the embryos isolated from seeds at various stages of development are shown in Figures 2(A)–(E). Isolation of embryos from seeds of several capsules and their microscopic visualization indicated that embryos reached the heart stage around 9 dpa (days post anthesis) under the conditions in our greenhouse. No visible GUS activity was detected in embryos at heart or late heart stages. However, GUS activity was observed in embryos at late torpedo and older stages of development.

Histochemical analysis of GUS activity was carried out in developing embryos isolated from the seeds of a T1-homozygous cotton plant. GUS staining was first detected in embryos at 16 dpa. At this stage, the cotyledons had just begun to expand and GUS activity appeared just below the cotyledons, at the junction of cotyledons and hypocotyl (Figures 2(F), (G)). The activity increased and spread throughout the embryo as the seed development progressed (Figures 2(H)–(N)). Intense staining was observed in embryos 40 dpa and in mature embryos isolated from dry seeds. Figure 2(O) shows an embryo isolated from a null segregant seed that was negative following histochemical GUS assay.

Seeds from a homozygous T2 generation of *Arabidopsis* transformed with AGP:*gus*A were used for histochemical analyses. Expression of AGP:*gus*A was monitored in the embryos isolated from seeds at various developmental stages. GUS staining was not visible in heart stage and late heart stage embryos (results not shown). A low level of GUS activity was observed in the torpedo stage embryos (Figure 2(P)), and the intensity of blue staining progressively increased as the embryos grew to maturity (Figures 2(Q)–(T)). An intense GUS staining was found in the embryos isolated from dry seeds. Taken together, results from these three dicot species suggest that gene expression driven by AGP is confined to middle to late stages of embryo development.

Quantitative analysis of GUS activity during seed development in tobacco and cotton

Histochemical analysis does not permit detection of low levels of GUS activity and also does not give a precise measure of increase in the level of GUS expression. Therefore, AGP activity during seed development was studied by monitoring GUS expression in developing seeds (tobacco) and developing embryos (cotton) by quantitative, fluorometric GUS assay at various time points after flowering. Measurable GUS-specific activity was first detected at 12 dpa in the seeds from a T1 homozygous tobacco plant. The activity then increased rapidly, finally reaching a maximum at 20 dpa (Figure 3(A)). From this point to maturity, there was no significant increase in GUS activity.

Because of large seed size, the relatively slow process of embryo development, and the ease with which the embryos can be isolated from developing seeds, cotton offered the best system to carry out detailed characterization of AGP:*gus*A expression at the single seed level. The results from quantitative analysis for GUS activity, protein levels, and fresh weight in developing embryos, isolated from the seeds from a T1 homozygous cotton plant, are shown in Figure 4. This plant was grown in the greenhouse in the month of April and at this time of the year, the bolls opened about 43 dpa. GUS expression was first detected at 15 dpa (60 pmoles/mg protein/min). Thereafter, there was a slow increase in GUS activity till 20 dpa, followed by a rapid rise until 40 dpa. From this peak until seed maturity, there was a small but statistically significant decline in activity. During seed development, the protein levels (as measured in the GUS extraction buffer) increased rapidly from 15 to 30 dpa, followed by a slow increase till 40 dpa before leveling off. It was not possible to accurately weigh the embryos before 18 dpa. However, from this point on, embryo fresh weight increased until 40 dpa, followed by a decrease as the seed reached dry state. These results confirm and extend our observations with histochemical analysis. The AGP:*gus*A expression begins in cotton embryos at around 15 dpa, and the activity either levels off or declines beyond 40 dpa.

GUS activity during seed germination and in mature parts of the plant

In order to determine if AGP expression is confined strictly to developing embryos/seeds, we monitored GUS activity in germinating *Arabidopsis* and tobacco

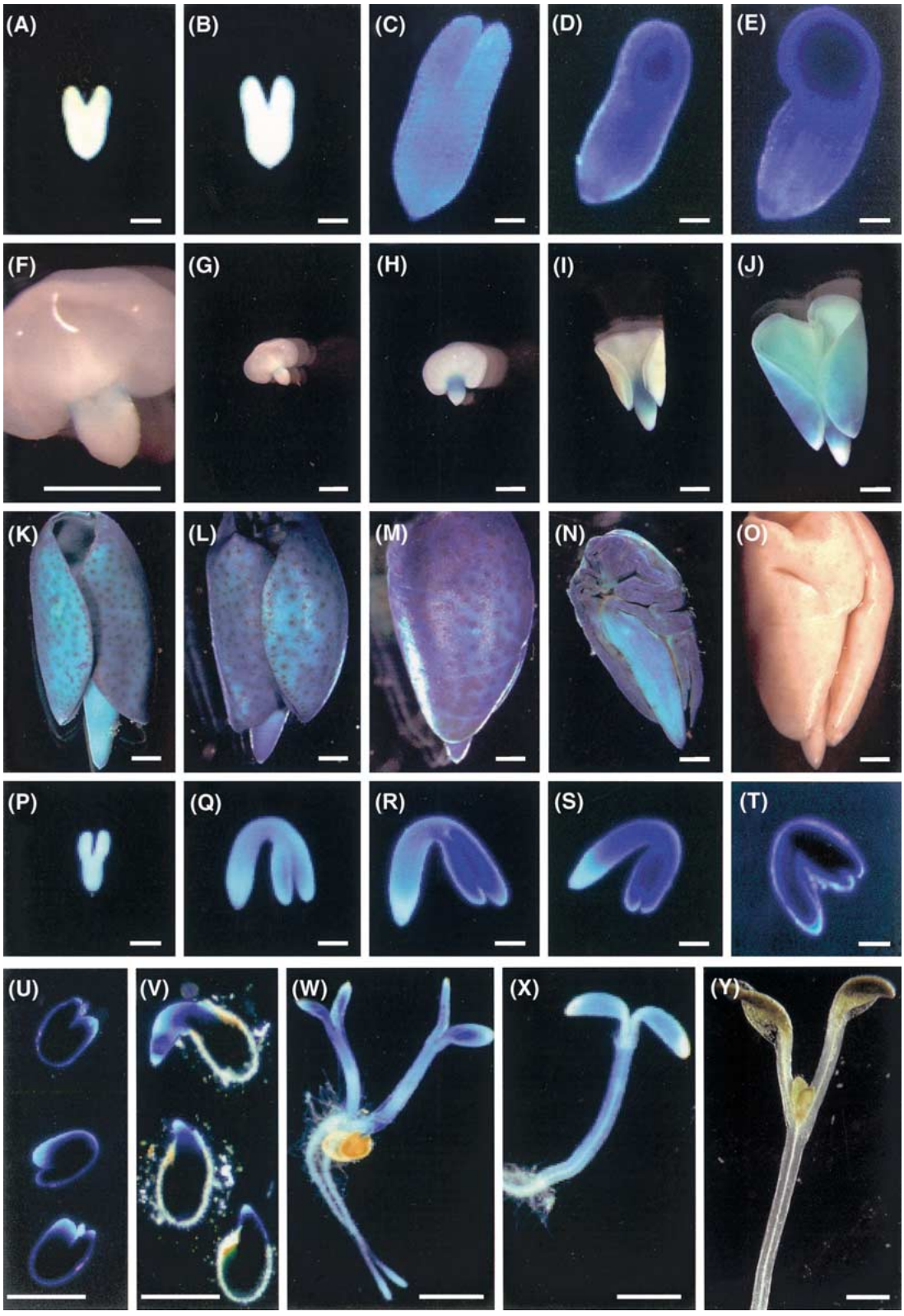

Figure 2.

Figure 2. Histochemical localization of GUS activity in developing embryos from stably transformed tobacco, cotton, and *Arabidopsis* plants and in germinating *Arabidopsis* seedlings. (A)–(E) Activity in embryos from T1-homozygous tobacco plant. (A) embryo 9 dpa, (B) embryo 10 dpa, (C) embryo 13 dpa, (D) embryo 17 dpa, (E) mature embryo from dry seed. (F)–(N) Activity in embryos from T1-homozygous cotton plant. (F) high magnification image of embryo 16 dpa showing the beginning of GUS activity, (G) same embryo at lower magnification, (H) embryo 18 dpa, (I) embryo 19 dpa, (J) embryo 20 dpa, (K) embryo 25 dpa, (L) embryo 30 dpa, (M) embryo 40 dpa, (N) embryo isolated from dry seed that has been cut through the middle showing the staining in the radicle and hypocotyl regions, (O) embryo from a null segregant cotton seed. (P)–(T) Activity in developing embryos from the seeds of T2-homozygous *Arabidopsis thaliana* C24 plant; (P) torpedo stage embryo (4 dpa), (Q) walking-stick stage embryo (5 dpa), (R) upturned-U stage embryo (7 dpa), (S) partially mature embryo (9 dpa), (T) mature embryo from dry seed. (U)–(Y). GUS activity during seed germination in transgenic *Arabidopsis*. (U) embryos isolated from dry seeds, (V) 1-day-old seedling, (W) 3-day-old seedling, (X) 5-day-old seedling, (Y) 8-day-old seedling. Bars: (A)–(E), (P)–(T) = 100 μm; (F)–(O), (U)–(Y) = 1 mm.

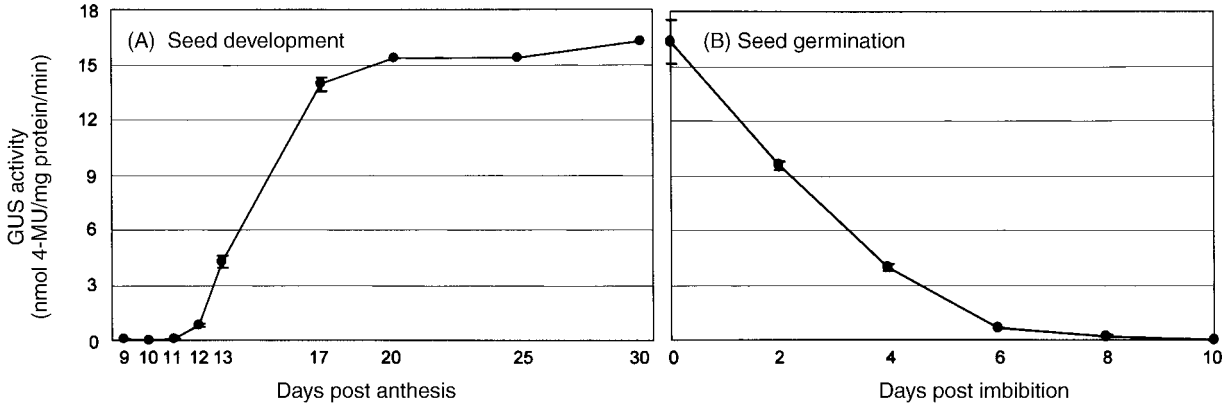

Figure 3. (A) Developmental regulation of GUS expression by the α-globulin promoter in tobacco seeds. GUS specific activity in developing tobacco seeds as a function of days post anthesis. (B) GUS specific activity in tobacco seedlings during germination at different days post imbibition.

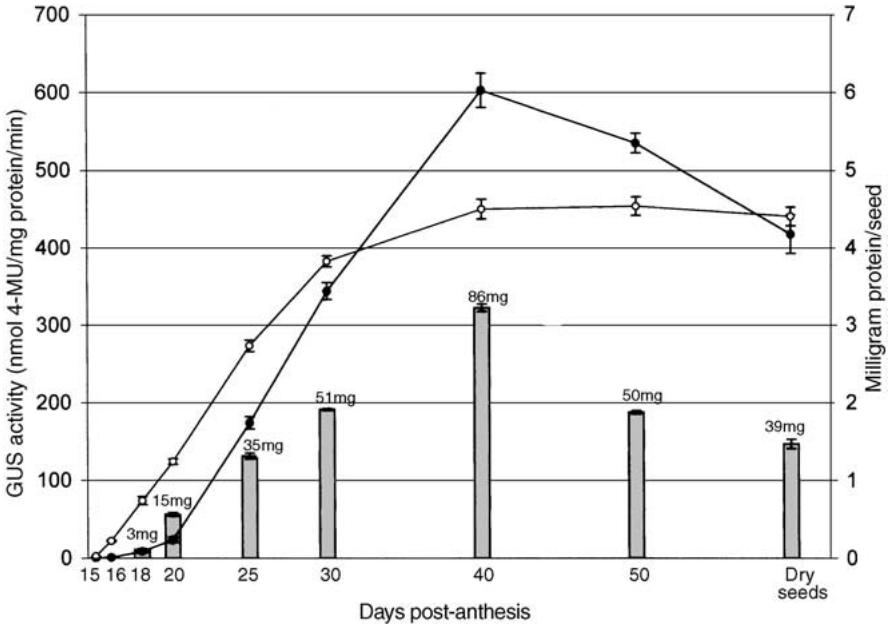

Figure 4. Developmental regulation of GUS expression by the α-globulin promoter in cotton embryos. GUS specific activity (closed circle) and total protein (open circle) in extracts from developing cotton embryos as a function of days post anthesis (dpa). Each point represents mean value (±SE) of 12 individual embryos. Vertical bars represent the mean (±SE) embryo fresh weight of 12 individuals starting from 18 dpa.

seedlings, and in various tissues from a mature cotton plant.

In *Arabidopsis*, GUS activity was analyzed in germinating seedlings using the histochemical method. Results presented in Figures 2(U)–(Y) show that the intensity of GUS staining decreased progressively as the seedlings grew. At 5 days post imbibition, there was still some residual GUS activity visible. However, after 7 days, faint patches of blue staining were observed only at the two ends of the hypocotyl. GUS staining was not visible in cotyledons, root or in the middle portion of the hypocotyl. No GUS activity was observed in seedlings beyond 7 days post imbibition (Figure 2(Y)).

Surface sterilized seeds from a T1 homozygous tobacco plant were germinated on MS medium (Murashige & Skoog, 1962). GUS fluorometric assay was carried out using extracts from whole seedlings at 0, 2, 4, 6, 8, and 10 days post imbibition. GUS activity decreased continuously following seed germination (Figure 3(B)) and only 2% of the initial activity was found after 8 days. No GUS activity was detected in seedlings 10 days after germination.

Histochemical GUS analysis was performed on various parts of three different T0 transgenic cotton plants that expressed the reporter gene in the embryos. GUS activity-dependent histochemical staining was not detected in tissues such as stem, leaf, petiole, flower stock, sepals, petals or square of the transgenic plant. We also carried out more sensitive fluorometric analyses to detect AGP activity in different organs and tissues of one of the transgenic cotton plants. Results from this analysis, presented in Table 1, show clearly that no measurable GUS activity was present in stem, leaf, floral bud, pollen, and root. A high level of GUS activity was detected only in the seeds. These results suggest that AGP-driven transgene activity is tightly controlled and is specific to the seed.

Table 1. GUS specific activity in various tissues of a T1-homozygous transgenic cotton plant and in control seeds

Tissue type	GUS activity[a] (nmol 4-MU/mg protein/min)
Stem	0.018 ± 0.002
Leaf	0.014 ± 0.005
Root	0.12 ± 0.006
Floral bud	0.11 ± 0.05
Pollen	0.024[b]
Transgenic seed[c]	349.9 ± 55
Control seed[c]	0.002 ± 0.0004

[a]Values are mean GUS activity ± SE from three replicates.
[b]The number of replicates were not sufficient to calculate SE (5.7 mg pollen was used in the assay).
[c]Assay was performed in embryos collected from 10 seeds for each replicate.

Certain seed-specific promoters have been shown to be activated in the vegetative parts under water stress conditions (Vivekananda et al., 1992; Siddiqui et al., 1998). In order to explore the possibility of AGP activation under water deficit conditions, we carried out fluorometric GUS analysis in the leaves of three independent T0 transgenic cotton plants that were found to produce GUS-positive seeds. They were subjected to drought stress in the greenhouse by withholding water. Leaf samples were analyzed for GUS activity every 48 h since the last watering until the time they showed complete wilting. No measurable GUS activity was detected in any of the leaf samples from three transgenic plants even after they were completely wilted (results not shown).

GUS activity in seeds from different transgenic lines

Preliminary results had indicated that AGP-driven GUS activity differed greatly amongst these three species. We carried out extensive analyses on seeds from a number of independent transgenic lines (that were positive for GUS activity as tested by histochemical method) from *Arabidopsis*, tobacco, and cotton to confirm this observation. The results from these analyses are presented in Figure 5. GUS activity in the seeds from 11 independent transgenic tobacco lines ranged from 0.6 to 18 nmol 4-MU/mg protein/min. Similar analysis in seeds from 10 independent transgenic *Arabidopsis* lines showed a range of 49–203 nmol 4-MU/mg protein/min. GUS activity in 10 independent transgenic cotton lines ranged from 118 to 1777 nmol 4-MU/mg protein/min. Similar high levels of seed-specific promoter expression has been reported also in maize seeds obtained from glutelin promoter:*gus*A and zein promoter:*gus*A transformants (Russell & Fromm, 1997). The results suggest that the cotton AGP, although recognized in different heterologous systems as a seed-specific promoter, showed the highest level of activity in cotton.

GUS activity in individual seeds from hemizygous and homozygous T1 progeny

The large seed size of cotton allowed analysis of GUS activity at a single seed level. This provided us with an opportunity to get a quantitative measure of GUS

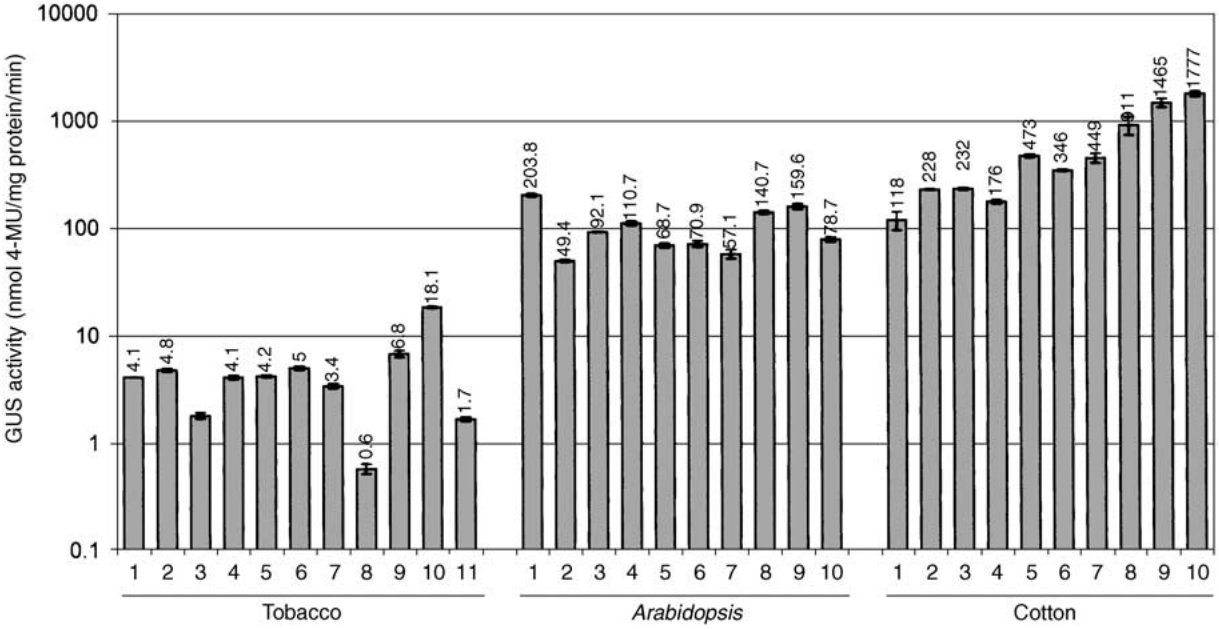

Figure 5. GUS activity in seeds from independent transgenic lines of T0 tobacco (11 lines), T1 *Arabidopsis* (10 lines), and T0 cotton (10 lines). Only the lines that tested positive for GUS activity, as examined by histochemical method, were used for this study. Bars represent mean GUS specific activity (±SE) of three replicate samples. The mean values are shown above each bar. Note that the GUS activity is plotted on a log scale.

activity in individual seeds within the segregating T2 seed population produced by a hemizygous T1 plant and compare these values with activities in individual seeds produced by a homozygous T1 plant. Results from this analysis are shown in Figure 6. All of the T2 seeds from the homozygous T1 parent showed GUS activity (Figure 6, top histograms) suggesting that reintroduction of a native promoter, even under homozygous conditions did not result in transgene silencing in this line. T2 seeds from the hemizygous T1 parent showed clear phenotypic segregation (3:1) for the transgene activity (Figure 6, bottom histograms). Moreover, amongst the seeds showing GUS activity, two different levels of activity was apparent in the majority of the cases. The higher level activity in about one fourth of T2 seeds from the hemizygous parent was similar to the level seen in the T2 seeds from the homozygous parent. Thus, the two different levels of GUS activity in the seeds from the hemizygous plant may be a result of either hemizygous or homozygous transgenic status of the individual seed suggesting a gene dose effect.

Discussion

Globulins are known to be the most prevalent seed storage proteins of dicotyledonous plants (Borroto & Dure, 1987) and their regulatory sequences potentially are a useful source of promoters that can be utilized to confer strong seed-specific expression of transgenes in a wide range of dicot species. In this study, we isolated an 1108 bp long fragment of the regulatory region of an α-globulin B gene from cotton and showed that it provides a high degree of spatial and temporal regulation of a transgene in cotton, *Arabidopsis*, and tobacco.

Histochemical analysis of GUS activity was useful in identifying the timing and localization of α-globulin promoter-regulated expression. Seed storage proteins are known to begin expression in the middle of the cotyledon stage and early maturation stage during embryo development, and their expression increases through seed maturity (Higgins, 1984; Goldberg et al., 1989). Our results showed that AGP expression begins around this stage of seed development (torpedo stage of developing embryos in tobacco and in *Arabidopsis* and embryos with expanding cotyledons in cotton). The increase in the intensity of GUS staining as the embryo grew to maturity corresponded with the increase in the GUS activity as measured by fluorometric GUS assay. In mature cotton embryos, the intensity was higher in cotyledons compared with the hypocotyl and radicle portions (Figure 2(N)). It is known that the presence of discrete promoter

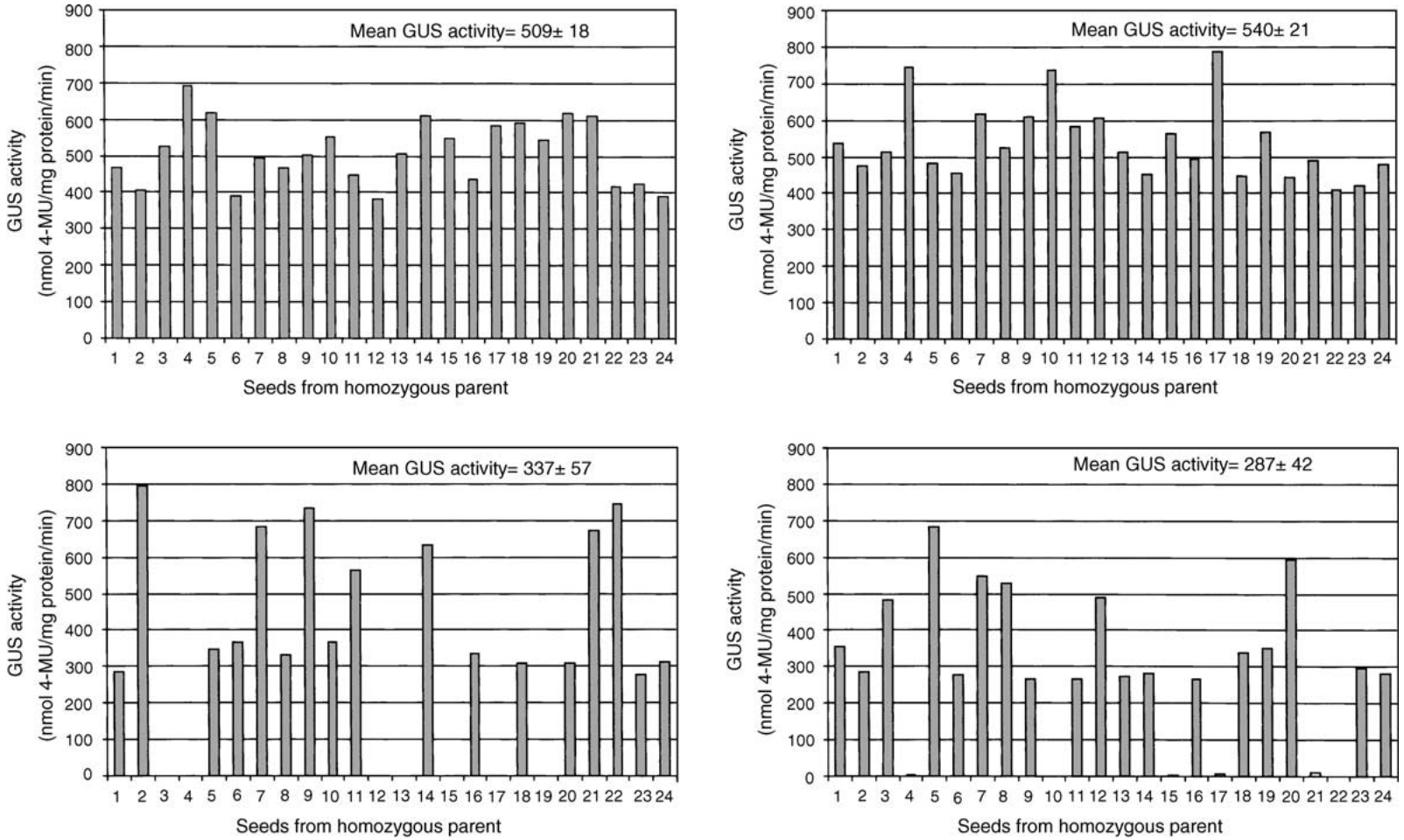

Figure 6. GUS specific activity in individual embryos isolated from two batches of seeds from a single T1 homozygous and two batches of seeds from a single T1 hemizygous cotton plant. Both homozygous and hemizygous T1 plants were derived from a single T0 transgenic line. Mean GUS activity (±SE) of 24 individual seeds is shown in each graph.

elements is important for the expression of the reporter genes in different regions of the embryo (Goldberg et al., 1994). Results from histochemical localization of GUS activity in mature embryos from transgenic plants of three plant species suggests that the 1108 bp promoter region has the required *cis*-acting domains that confer expression in the embryo.

AGP:*gus*A expression during seed development as measured by fluorometric GUS assay corresponded with that of the histochemical analysis. The AGP:*gus*A began to express in tobacco at 12 dpa and increased through seed maturation (Figure 3(A)). This pattern of expression is similar to the expression pattern of *gus*A fusion constructs with the β-phaseolin promoter (Bustos et al., 1989) and sunflower helianthinin promoter (Bogue et al., 1990) in tobacco. In cotton, we found that the promoter began to express as early as 15 dpa and its expression increased thereafter till 40 dpa. Subsequent decline in GUS activity through seed maturity indicates that the promoter activity had either stopped or significantly declined. Because of the relatively stable nature of GUS protein, it is not possible to pinpoint the time at which promoter activity begins to decline or stops. The increase in the expression of AGP:*gus*A in cotton corresponds to the accumulation of the 51 kDa protein encoded by α-globulin B gene in developing cottonseeds, as determined by SDS/PAGE gel analysis (Dure & Chlan, 1981).

To our knowledge, this is the first, report, where a seed-specific promoter has been quantitatively evaluated in the homologous species as well as two heterologous species. Quantitative GUS assays in seeds from individual lines of transgenic plants of cotton, *Arabidopsis*, and tobacco showed that the strength of promoter expression in different species is different. The mean GUS activity ($\pm$ standard error) from all tested transgenic lines in cotton, *Arabidopsis*, and tobacco was 618 ± 183, 103 ± 17, and 5 ± 1.5, respectively. The extent of AGP-driven GUS activity detected in tobacco seeds is within the range reported with the phaseolin promoter:*gus*A (Bustos et al., 1989, 1991) and the helianthinin promoter:*gus*A (Bogue et al., 1990; Nunberg et al., 1994, 1995) expression in tobacco seeds. It is interesting that tobacco showed less than 1% the level of expression detected in cotton; whereas in *Arabidopsis*, the mean level of expression was 16.7% of that in cotton. It is possible that higher level of AGP-driven *gus*A expression in *Arabidopsis* compared to tobacco is due to the fact that *Arabidopsis* (Brassicales) is phylogenetically closer to cotton (Malvales) (Borroto & Dure, 1987; Soltis et al., 1999), and that the *cis*-elements and *trans*-acting factors required for the regulation of AGP are more conserved between these two species. Because it is easy to transform and because it produces large quantities of seeds, tobacco has been used widely as a model dicot system to test various other seed-specific promoters (reviewed by Sun & Larkins, 1993; Goossens et al., 1999). In our studies, a very low-level expression of AGP in tobacco suggests that testing of a promoter in a heterologous system such as tobacco alone does not necessarily provide a realistic indication of the promoter strength.

The large seed size in cotton made it possible to carry out phenotypic segregation analysis at a single seed level in the embryos from a T1 hemizygous line. Null segregants were quite obvious, moreover, the seeds that had GUS activity could be categorized into one of two sets of activity levels, possibly reflecting their hemizygous or homozygous status. Similar analysis on the seed population from a T1 homozygous plant showed that the individual seed GUS activity fell into the higher level category. It is interesting that the average GUS activity of putative hemizygous seeds was close to half the level of the average activity in putative homozygous seeds suggesting a gene dose effect.

The absence of detectable levels of GUS activity in stem, leaf, root, pollen, and floral bud of the cotton plant indicates that the α-globulin promoter does not express in the vegetative or floral organs of the plant. This is a desired quality for a seed-specific promoter. The fact that GUS activity rapidly drops to undetectable levels following seed germination and that it is completely absent in mature seedlings in tobacco (Figure 3(B)) and *Arabidopsis* (Figure 2(Y)) further supports the results from cotton, suggesting that the promoter is highly seed-specific. Certain seed-specific promoters from the *Lea* class of genes have been shown to be induced by ABA as well as by water deficit conditions in the vegetative parts (Seffens et al., 1990; Vivekananda et al., 1992; Siddiqui et al., 1998). Also, the β-phaseolin promoter:*gus*A has been shown to be induced in isolated embryos from a transgenic tobacco by exogenous ABA (Bustos et al., 1998). Thus, the possibility exists that the AGP may be activated in vegetative parts under drought conditions that are known to result in increased levels of endogenous ABA. In order to test this, GUS fluorometric assays were performed in the leaves at different developmental stages from three different transgenic lines of cotton plants that were subjected to water stress

by withholding water. No measurable GUS activity was noticed at any stage of the treatment, even after complete wilting. This property of the promoter is important in cases, where even a low level of promoter activity in the vegetative parts is not acceptable.

It is likely that the 1108 bp sequence analyzed in this investigation may not be the whole promoter. Detailed characterization of the complete sequence and identification of various regulatory *cis*-elements will be part of future studies. Other investigations involving phaseolin and helianthinin promoters showed that partial promoter sequences are sufficient for conferring a high degree of seed-specific expression. It is clear that the 1108 bp α-globulin promoter sequence from cotton, characterized in this study, can confer a strong seed-specific expression in cottonseed as well as in the seeds of two other dicots. We believe that AGP will be useful for any application involving transgene-mediated over-expression or suppression during seed development in dicots, thus adding to the availability of seed-specific promoters.

Acknowledgements

We thank LeAnne Mohr for her excellent technical assistance during all stages of this study. Dr E. Chandrakanth assisted with transient assays, Dr N. Padmavathi with *Arabidopsis* transformation and Janet Alexander with media preparation. This research was supported by funds from Texas Cotton Biotechnology Initiative (TxCOT), Texas Higher Education Coordinating Board (Advanced Research Program), Texas Food & Fibers Commission, and Texas Agriculture Experiment Station.

References

An G, Ebert PR, Mitra A and Ha SB (1988) Binary vectors. In: Gelvin SB and Shilperoort RA (eds), Plant Molecular Biology Manual. (pp. A3: 1–19) Kluwer Academic Publishers, Dordrecht.

Barker SJ, Harada JJ and Goldberg RB (1988) Cellular localization of soybean storage protein mRNA in transformed tobacco seeds. *Proc Natl Acad Sci USA* **85**: 458–462.

Bechtold N and Pelletier G (1998) *In planta Agrobacterium*-mediated transformation of adult *Arabidopsis thaliana* plants by vacuum infiltration. *Meth Mol Biol* **82**: 259–266.

Bogue MA, Vonder Haar RA, Nuccio ML, Griffing LR and Thomas TL (1990) Developmentally regulated expression of a sunflower 11S seed protein gene in transgenic tobacco. *Mol Gen Genet* **222**: 49–57.

Borroto K and Dure III L (1987) The globulin seed storage proteins of flowering plants are derived from two ancestral genes. *Plant Mol Biol* **8**: 113–131.

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.

Brusslan JA and Tobin EM (1995) Isolation of new promoter-mediated co-suppressed lines in *Arabidopsis thaliana*. *Plant Mol Biol* **27**: 809–813.

Bustos MM, Begum D, Kalkan FA, Battraw MJ and Hall TC (1991) Positive and negative *cis*-acting DNA domains are required for spatial and temporal regulation of gene expression by a seed storage protein promoter. *EMBO J* **10**: 1469–1479.

Bustos MM, Guiltinan MJ, Jordano J, Begum D, Kalkan FA and Hall TC (1989) Regulation of β-glucuronidase expression in transgenic tobacco plants by an A/T-rich, *cis*-acting sequence found upstream of a French bean β-phaseolin gene. *Plant Cell* **1**: 839–853.

Bustos MM, Iyer M and Gagliardi SJ (1998) Induction of a β-phaseolin promoter by exogenous abscisic acid in tobacco: developmental regulation and modulation by external sucrose and Ca^{2+} ions. *Plant Mol Biol* **37**: 265–274.

Cahoon EB, Marillia E-F, Stecca KL, Hall SE, Taylor DC and Kinney AJ (2000) Production of fatty acid components of meadowfoam oil in somatic soybean embryos. *Plant Physiol* **124**: 243–251.

Chen Z-L, Pan N-S and Beachy RN (1988) A DNA sequence element that confers seed-specific enhancement to a constitutive promoter. *EMBO J* **7**: 297–302.

Chlan CA, Borroto K, Kamalay JA and Dure III L (1987) Developmental biochemistry of cottonseed embryogenesis and germination. XIX. Sequences and genomic organization of the α globulin (vicilin) genes of cottonseed. *Plant Mol Biol* **9**: 533–546.

Dickinson CD, Evans RP and Nielsen NC (1988) RY repeats are conserved in the 5′-flanking regions of legume seed-protein genes. *Nucl Acids Res* **16**: 371.

Dure III L (1989) Characteristics of the storage proteins of cotton. *J Am Oil Chem Soc* **66**: 356–359.

Dure III L and Chlan C (1981) Developmental biochemistry of cottonseed embryogenesis and germination. XII. Purification and properties of principal storage proteins. *Plant Physiol* **68**: 180–186.

Goldberg RB (1986) Regulation of plant gene expression. *Phil Trans R Soc Lond* **314**: 343–353.

Goldberg RB, Barker SJ and Perez-Grau L (1989) Regulation of gene expression during plant embryogenesis. *Cell* **56**: 149–160.

Goldberg RB, de Paiva G and Yadegari R (1994) Plant embryogenesis: zygote to seed. *Science* **266**: 605–614.

Goossens A, Dillen W, De Clercq J, Van Montagu M and Angenon G (1999) The *arcelin*-5 gene of *Phaseolus vulgaris* directs high seed-specific expression in transgenic *Phaseolus acutifolius* and *Arabidopsis* plants. *Plant Physiol* **120**: 1095–1104.

Goto F, Yoshihara T, Shigemoto N, Toki S and Takaiwa F (1999) Iron fortification of rice seed by the soybean ferritin gene. *Nat Biotechnol* **17**: 282–286.

Higgins TJV (1984) Synthesis and regulation of major proteins in seeds. *Ann Rev Plant Physiol* **35**: 191–221.

Hitz WD, Yadav NS, Reiter RS, Mauvais CJ and Kinney AJ (1995) Reducing polyunsaturation in oils of transgenic canola and soybean. In: Kader J-C and Mazliak (eds), Plant Lipid Metabolism. (pp.506–508) Kluwer Academic Publishers, Dordrecht.

Horsch RB, Fry JE, Hoffmann N, Neidermeyer J, Rogers SG and Fraley RT (1988) Leaf disc transformation. In: Gelvin SB and

Schilperoort RA (eds), Plant Molecular Biology Manual. (pp. A5: 1–9) Kluwer Academic Publishers, Dordrecht.

Jefferson RA, Kavanagh TA and Bevan MW (1987) GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J* **6**: 3901–3907.

Kawagoe Y, Campell BR and Murai N (1994) Synergism between CACGTG (G-box) and CACCTG *cis*-elements is required for activation of the bean seed storage protein β-*phaseolin* gene. *Plant J* **5**: 885–890.

Kawagoe Y and Murai N (1992) Four distinct nuclear proteins recognize *in vitro* the proximal promoter of the bean seed storage protein β-*phaseolin* gene conferring spatial and temporal control. *Plant J* **2**: 927–936.

Kim SY, Chung H-J and Thomas TL (1997) Isolation of a novel class of bZIP transcription factors that interact with ABA-responsive and embryo-specification elements in the *Dc3* promoter using a modified yeast one-hybrid system. *Plant J* **11**: 1237–1251.

Kinney AJ (1996) Development of genetically engineered oilseeds: from molecular biology to agronomics. In: Williams JP, Khan MU and Lem NW (eds), Physiology, Biochemistry and Molecular Biology of Plant Lipids. (pp. 298–300) Kluwer Academic Publishers, Dordrecht.

Lessard PA, Allen RD, Fujiwara T and Beachy RN (1993) Upstream regulatory sequences from two β-conglycinin genes. *Plant Mol Biol* **22**: 873–885.

Murashige T and Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol Plant* **15**: 473–497.

Nunberg AN, Li Z, Bogue MA, Vivekananda J, Reddy AS and Thomas TL (1994) Developmental and hormonal regulation of sunflower helianthinin genes: proximal promoter sequences confer regionalized seed expression. *Plant Cell* **6**:473–486.

Nunberg AN, Li Z, Chung H-J, Reddy AS and Thomas TL (1995) Proximal promoter sequences of sunflower helianthinin genes confer regionalized seed-specific expression. *J Plant Physiol* **145**: 600–605.

Park Y-D, Papp I, Moscone EA, Iglesias VA, Vaucheret H, Matzke AJM et al. (1996) Gene silencing mediated by promoter homology occurs at the level of transcription and results in meiotically heritable alterations in methylation and gene activity. *Plant J* **9**: 183–194.

Russell DA and Fromm ME (1997) Tissue-specific expression in transgenic maize of four endosperm promoters from maize and rice. *Transgenic Res* **6**: 157–168.

Seffens WS, Almoguera C, Wilde HD, Haar RAV and Thomas TL (1990) Molecular analysis of a phylogenetically conserved carrot gene: developmental and environmental regulation. *Dev Genet* **11**: 65–76.

Shintani D and DellaPenna D (1998) Elevating the vitamin E content of plants through metabolic engineering. *Science* **282**: 2098–2100.

Siddiqui NU, Chung H-J, Thomas TL and Drew MC (1998) Abscisic acid-dependent and -independent expression of the carrot late-embryogenesis-abundant-class gene *Dc3* in transgenic tobacco seedlings. *Plant Physiol* **118**: 1181–1190.

Siebert PD, Chenchik A, Kellogg DE, Lukyanov KA and Lukyanov SA (1995) An improved PCR method for walking in uncloned genomic DNA. *Nucl Acids Res* **23**: 1087–1088.

Soltis PS, Soltis DE and Chase MW (1999) Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature* **402**: 402–404.

Sun SSM and Larkins BA (1993) Transgenic plants for improving seed storage proteins. In: Kung S-d and Wu R (eds), Transgenic Plants: Engineering and Utilization. Vol. 1 (pp. 339–372) Academic Press, California.

Sunilkumar G and Rathore KS (2001) Transgenic cotton: factors influencing *Agrobacterium*-mediated transformation and regeneration. *Mol Breed* **8**: 37–52.

Vaucheret H (1993) Identification of a general silencer for 19S and 35S promoters in a transgenic tobacco plant: 90 bp of homology in the promoter sequence are sufficient for *trans*-inactivation. *C R Acad Sci Paris* **316**: 1471–1483.

Vivekananda J, Drew MC and Thomas TL (1992) Hormonal and environmental regulation of the carrot *lea*-class gene *Dc3*. *Plant Physiol* **100**: 576–581.

Ye X, Al-Babili S, Kloti A, Zhang J, Lucca P, Beyer P et al. (2000) Engineering the provitamin A (β-carotene) biosynthetic pathway into (carotinoid-free) rice endosperm. *Science* **287**: 303–305.